

J. Robin Harris
Jon Marles-Wright *Editors*

Macromolecular Protein Complexes V

Structure and Function

Subcellular Biochemistry

Volume 104

Series Editor

J. Robin Harris, Institute of Molecular Physiology, University of Mainz, Mainz,
Germany

The book series SUBCELLULAR BIOCHEMISTRY is a renowned and well recognized forum for disseminating advances of emerging topics in Cell Biology and related subjects. All volumes are edited by established scientists and the individual chapters are written by experts on the relevant topic. The individual chapters of each volume are fully citable and indexed in Medline/Pubmed to ensure maximum visibility of the work.

J. Robin Harris • Jon Marles-Wright
Editors

Macromolecular Protein Complexes V

Structure and Function

 Springer

Editors

J. Robin Harris
Institute of Molecular Physiology
University of Mainz
Mainz, Germany

Jon Marles-Wright
Biosciences Institute
Newcastle University
Newcastle upon Tyne, UK

ISSN 0306-0225

ISSN 2542-8810 (electronic)

Subcellular Biochemistry

ISBN 978-3-031-58842-6

ISBN 978-3-031-58843-3 (eBook)

<https://doi.org/10.1007/978-3-031-58843-3>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

Chapter 13 is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). For further details see license information in the chapter. This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

This, the fifth volume of the *Subcellular Biochemistry* series devoted to Macromolecular Protein Complexes, follows the pattern established in the previous volumes by presenting a broad range of interesting topics. The field of structural biology has continued to expand over recent years, with a particular growth in cryo-electron microscopy owing to the advances in instrumentation seen since the Resolution Revolution. Since the previous volume we have encountered a genuine paradigm shift with the release of AlphaFold2 and other machine-learning based structure prediction algorithms, the availability of computed structure models for most of the known protein universe has accelerated structural biology and discovery in many areas of experimental biology. The achievements of the teams involved in developing these algorithms and their model databases, were built on over 50 years of publicly accessible, and primarily publicly funded, experimental models deposited in the PDB.

This paradigm shift has forced a period of reflection for structural biologists, what can an experimentalist contribute when we have such a wealth of highly accurate models? There will always be a need to validate these computed structure models experimentally; but it is certain that there will certainly be a shift in focus from the discovery of new protein structures. New methods in the analysis of cryo-EM datasets will allow us to better understand the dynamics of protein complexes; and there is still a need for the study and experimental validation of protein-ligand and protein-drug interactions.

Rather than dwelling here upon a few chapters, the reader is simply referred to the immediately following Contents List; where, at a glance, the fascinating range of chapter content can be appreciated. The contributors to this volume are all active research scientists, experts in their respective fields of study.

It can also be readily appreciated that the revolution in cryoEM resolution continues apace, in some instances contributing to dynamic structural studies, where the impact of macromolecular structure on function becomes of increasing significance.

We hope that this volume, along with our four previous volumes and several other relevant volumes in the *Subcellular Biochemistry* series, will help to provide

an encyclopaedic overview, emphasising the importance of macromolecular protein complexes in structural biology. The contribution of electron microscopy to protein structure has been a long-standing research interest of one of us (JRH) since 1968, when the cylindrical 20S proteasome and the ring-like peroxiredoxin-2 protein oligomer were first structurally defined, then as proteins of unknown function. That the overall subject has advanced so significantly in the intervening 55 years is almost beyond belief. CryoEM now contributes in all spheres of molecular and cellular biology and medical science, where a multitude of large protein assemblies are found, including most significantly in recent years, the field of virology. Indeed, the content of this book exemplifies the massive contribution that structural biology has made over recent years, together with its likely impact within the broad field of molecular physiology.

Mainz, Germany
Newcastle, UK

J. Robin Harris
Jon Marles-Wright

Contents

1	Structures and Efflux Mechanisms of the AcrAB-TolC Pump	1
	Zhili Yu, Xiaodong Shi, and Zhao Wang	
2	The Copper Efflux Regulator (CueR)	17
	Yangbo Hu and Bin Liu	
3	The Structural Biology of Catalase Evolution	33
	María Belén López, María Belén Oterino, and Javier M. González	
4	Building the Bacterial Divisome at the Septum	49
	Josiah J. Morrison and Jodi L. Camberg	
5	Telomere C-Strand Fill-In Machinery: New Insights into the Human CST-DNA Polymerase Alpha-Primase Structures and Functions	73
	Ci Ji Lim	
6	The Core Complex of Yeast COMPASS and Human Mixed-Lineage Leukemia (MLL), Structure, Function, and Recognition of the Nucleosome	101
	Uhn-Soo Cho	
7	Oligomerization of Monoamine Transporters	119
	Arnav Paul and Diwakar Shukla	
8	Macromolecular Interactions of Lipoprotein Lipase (LPL)	139
	Anna Wheless, Kathryn H. Gunn, and Saskia B. Neher	
9	Viral Genomic DNA Packaging Machinery	181
	Dorothy E. D. P. Hawkins, Owen C. Godwin, and Alfred A. Antson	
10	Human Transient Receptor Potential Ankyrin 1 Channel: Structure, Function, and Physiology	207
	Viktorie Vlachova, Ivan Barvik, and Lucie Zimova	

11	Structural Variations and Rearrangements in Bacterial Type II Toxin-Antitoxin Systems	245
	Maria Raadkjær Nielsen and Ditlev E. Brodersen	
12	Structures and Functions of the Human GATOR1 Complex	269
	Ilina Ivanova and Kuang Shen	
13	Mitochondrial Alpha-Keto Acid Dehydrogenase Complexes: Recent Developments on Structure and Function in Health and Disease	295
	Eszter Szabo, Balint Nagy, Andras Czajlik, Timea Komlodi, Oliver Ozohanics, Laszlo Tretter, and Attila Ambrus	
14	Structures and Electron Transport Paths in the Four Families of Flavin-Based Electron Bifurcation Enzymes	383
	Xiang Feng, Gerrit J. Schut, Michael W. W. Adams, and Huilin Li	
15	The Prothrombin-Prothrombinase Interaction	409
	Bosko M. Stojanovski, Bassem M. Mohammed, and Enrico Di Cera	
16	Structure and Function of the Glycosylphosphatidylinositol Transamidase, a Transmembrane Complex Catalyzing GPI Anchoring of Proteins	425
	Dianfan Li	
17	Maturation and Assembly of mTOR Complexes by the HSP90-R2TP-TTT Chaperone System: Molecular Insights and Mechanisms	459
	Andrés López-Perrote, Marina Serna, and Oscar Llorca	
18	Valosin-Containing Protein (VCP)/p97 Oligomerization	485
	Guimei Yu, Yunpeng Bai, and Zhong-Yin Zhang	
19	Oligomeric Structure of Yeast and Other Invertases Governs Specificity	503
	Elena Jiménez-Ortega and Julia Sanz-Aparicio	
20	Structure, Dynamics and Functional Implications of the Eukaryotic Vault Complex	531
	María González-Álamos, Pablo Guerra, and Núria Verdaguer	
21	Structure and Dynamics of Type 4a Pili and Type 2 Secretion System Endopili	549
	Nadia Izadi-Pruneyre, Yasaman Karami, and Michael Nilges	

Chapter 1

Structures and Efflux Mechanisms of the AcrAB-TolC Pump



Zhili Yu, Xiaodong Shi, and Zhao Wang

Abstract The global emergence of multidrug resistance (MDR) in gram-negative bacteria has become a matter of worldwide concern. MDR in these pathogens is closely linked to the overexpression of certain efflux pumps, particularly the resistance-nodulation-cell division (RND) efflux pumps. Inhibition of these pumps presents an attractive and promising strategy to combat antibiotic resistance, as the efflux pump inhibitors can effectively restore the potency of existing antibiotics. AcrAB-TolC is one well-studied RND efflux pump, which transports a variety of substrates, therefore providing resistance to a broad spectrum of antibiotics. To develop effective pump inhibitors, a comprehensive understanding of the structural aspect of the AcrAB-TolC efflux pump is imperative. Previous studies on this pump's structure have been limited to individual components or in vitro determination of fully assembled pumps. Recent advancements in cellular cryo-electron tomography (cryo-ET) have provided novel insights into this pump's assembly and functional mechanism within its native cell membrane environment. Here, we present a summary of the structural data regarding the AcrAB-TolC efflux pump, shedding light on its assembly pathway and operational mechanism.

Z. Yu

Verna and Marrs McLean Department of Biochemistry and Molecular Pharmacology,
Baylor College of Medicine, Houston, TX, USA
e-mail: zhiliy@bcm.edu

X. Shi

Jiangsu Province Key Laboratory of Anesthesiology and Jiangsu Province Key
Laboratory of Anesthesia and Analgesia Application, Xuzhou Medical University,
Xuzhou, Jiangsu, China
e-mail: shixd@xzhmu.edu.cn

Z. Wang (✉)

Verna and Marrs McLean Department of Biochemistry and Molecular Pharmacology,
Baylor College of Medicine, Houston, TX, USA
Department of Molecular and Cellular Biology, Baylor College of Medicine,
Houston, TX, USA
e-mail: zhaow@bcm.edu

Keywords Multidrug resistance · RND efflux pump · AcrAB-TolC · Cryo-EM · Cryo-ET

Introduction

Efflux Pumps in Health and Disease

The growing multidrug resistance (MDR) of pathogens poses a significant global health threat (Magill et al. 2014; Antão et al. 2018). With antibiotics being the prevailing treatment for inhibiting bacterial growth, bacteria have gradually developed the ability to resist these drugs. It is estimated that approximately 700,000 people succumb to drug-resistant bacterial infections worldwide every year (Antão et al. 2018). Clinically, most serious bacterial infections observed come from multidrug-resistant gram-negative species (Viale et al. 2015). While the origins of bacterial drug and antibiotic resistance are diverse and intricate, the capacity of bacteria to rapidly expel drugs through their efflux pumps is thought to be a major and common mechanism (Kobylka et al. 2019). The resistance-nodulation-cell division (RND) superfamily comprises a vast majority of efflux pumps existing in both gram-positive and gram-negative bacteria (Alav et al. 2021). These pumps play a pivotal role in MDR in various clinically relevant gram-negative bacterial species (Blair et al. 2014). RND efflux pumps possess the ability to expel a broad spectrum of antibiotics out of the bacterial cell, thereby maintaining their low intracellular concentration and rendering the antibiotics ineffective.

The cell envelope of gram-negative bacteria comprised of three major components: the inner membrane (IM), the peptidoglycan (PG) cell wall, and the outer membrane (OM) (Silhavy et al. 2010). RND pumps form tripartite assemblies spanning the entire bacteria cell envelope, comprising an inner membrane RND transporter, a periplasmic adapter protein, and an outer membrane channel (Ge et al. 2009; Du et al. 2015). Among the extensively studied RND efflux pumps, AcrAB-TolC stands out as the prototypical example (Li et al. 2015). AcrAB-TolC confers resistance to a diverse range of antibiotics (Nikaido 1996; Masuda et al. 2000; Murakami et al. 2006; Schindler and Kaatz 2016), and overproduction of this pump in clinical isolates serves as a biomarker of multidrug resistance (Swick et al. 2011). In the experiment using an in vitro hollow fiber model to simulate drug treatment in humans, AcrAB overproduction preceded the accumulation of target topoisomerase gene mutations and increased the minimal inhibitory concentration (MIC) of quinolone, while the deletion of the pump delayed the emergence of resistance (Singh et al. 2012). Additionally, AcrAB is implicated in quorum sensing (Yang et al. 2006) and production of extended-spectrum β -lactamase in certain bacterial strains (Pitout et al. 2005; Gröbner et al. 2009; Rogers et al. 2011).

Being present in many species of *Enterobacteriaceae*, such as *Escherichia coli* (*E. coli*), *Salmonella*, and *Klebsiella pneumonia*, the AcrAB-TolC efflux pump is composed of the inner membrane transporter AcrB, the periplasmic protein AcrA,

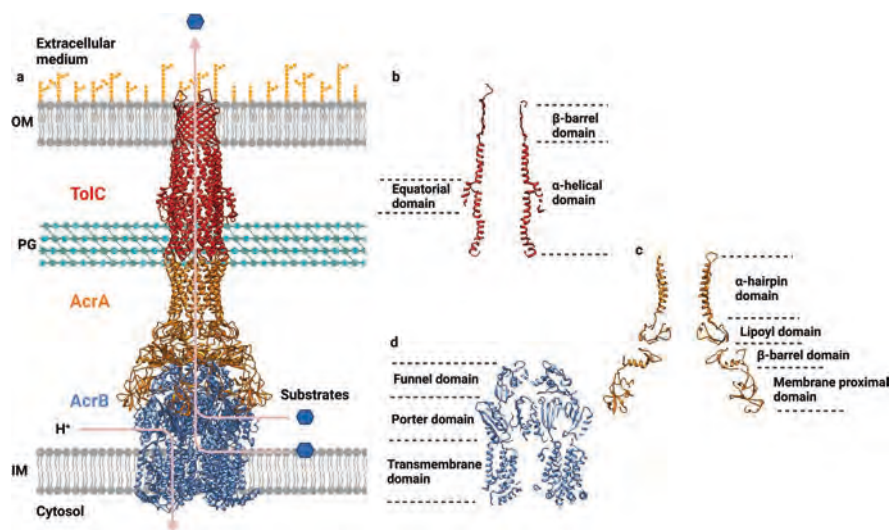


Fig. 1.1 The assembly and function of the AcrAB-TolC pump. (a) The assembly of the full tripartite pump of AcrAB-TolC in the gram-negative cell envelope and the potential pathway for substrate transportation. (b) The domain arrangement of TolC, the outer membrane protein. (c) The domain arrangement of AcrA, the periplasmic protein. (d) The domain arrangement of AcrB, the inner membrane transporter. The structure of in vitro fully assembled pump (PDB: 5NG5) is used for visualization purposes. OM outer membrane; IM inner membrane; PG peptidoglycan. (Adapted from Wang et al. 2017)

and the outer membrane protein TolC (Alav et al. 2021) (Fig. 1.1a). Homologous complexes of the AcrAB-TolC efflux pump are found throughout the other species of gram-negative bacteria, such as MexAB-OprM in *Pseudomonas aeruginosa* (*P. aeruginosa*), CmeABC in *Campylobacter jejuni*, and AdeABC in *Acinetobacter baumannii*, which are often upregulated in clinical isolates (Piddock 2006; Blair et al. 2015; Weston et al. 2018). Given the critical role of the AcrAB-TolC pump in gram-negative bacteria and its relevance to diseases, substantial efforts have been dedicated to studying the structure of this complex. Advancements in understanding the structural aspects of efflux pumps will reveal the mechanism of pump action and future therapeutic targeting, which will facilitate the rational design of effective, selective, and safe efflux pump inhibitors.

The Structural Basis of the AcrAB-TolC Pump

AcrB serves as the key determinant for substrate specificity and energy transduction within the AcrAB-TolC tripartite pump. It exhibits a wide recognition capability for various substrates, including antibiotics, dyes, detergents, and organic solvents (Tsukagoshi and Aono 2000; Blair et al. 2014). The RND transporter AcrB forms a homotrimer on the inner membrane, and each monomer has three domains: the

funnel domain, the porter domain, and the transmembrane domain (Fig. 1.1d) (Murakami et al. 2002, 2006; Seeger et al. 2006b). The transmembrane domain of each AcrB monomer comprises 12 α -helices, named TM1-TM12, along with an additional extra-membrane α -helix attached to the cytoplasmic membrane surface. The funnel domain encompasses the DN and DC subdomains. The porter domain is further divided into four subdomains, namely, PN1, PN2, PC1, and PC2, arranged clockwise when viewed from the top, with the PN1 subdomain closest to the center of the AcrB homotrimer. The access pocket and deep binding pocket for pump substrates are located in the porter domain (Nakashima et al. 2011; Eicher et al. 2012). It has been discovered that the AcrB protomer can adopt three distinct conformations known as access, binding, and extrusion, or alternatively loose (L), tight (T), and open (O) (Murakami et al. 2006; Seeger et al. 2006b; Sennhauser et al. 2007). During substrate transport, the three AcrB protomers cyclically transition between these conformations, operating in an active efflux pattern (Takatsuka and Nikaido 2009).

TolC, the outer membrane protein, forms a homotrimer, with each monomer comprising a β -barrel domain, an α -helical domain, and an equatorial domain surrounding the α -helical domain (Fig. 1.1b). Crystal structures have shown the periplasmic end of TolC to be in a closed conformation, facilitated by interactions between the α -helices (Koronakis et al. 2000; Andersen et al. 2002; Bavro et al. 2008; Pei et al. 2011). The periplasmic adaptor protein AcrA forms a hexamer, contacting with TolC at the top and AcrB at the bottom, respectively. The six AcrA protomers exhibit two distinct conformations known as AcrA protomer I and protomer II (Wang et al. 2017). Each protomer of AcrB interacts with one AcrA protomer I and one AcrA protomer II. AcrA comprises four domains: the α -hairpin domain, the lipoyl domain, the β -barrel domain, and the membrane proximal domain (MPD) (Fig. 1.1c) (Zgurskaya et al. 2009; Symmons et al. 2015). The α -hairpin domain interacts with TolC (Wang et al. 2017), and the MPD interacts with AcrB and is anchored to the inner membrane (Zgurskaya and Nikaido 1999; Alav et al. 2021).

The stoichiometry ratio of the components was determined based on the first cryo-electron microscopy (cryo-EM) structure of the purified fully assembled pump (Wang et al. 2017). The substrates from the periplasm and the inner membrane can enter the pump through the entry channels in the AcrB porter domain (Zwama et al. 2018). The AcrAB-TolC pump is energized by proton motive force (PMF), that is, coupled with substrate transportation, protons are transported across the inner membrane, from the periplasm to the cytoplasm (Kobylka et al. 2019). The pathway of proton transportation is in the AcrB transmembrane domain (Fig. 1.1a) (see details below) (Eicher et al. 2014). However, how the proton transportation is relevant to full pump functioning is unknown. Altogether, TolC, AcrA, and AcrB form a 3:6:3 stoichiometry ratio and form a complex spanning the gram-negative bacterial cell envelope. However, the mechanisms underlying protein transportation, pump assembly, and function remain to be discovered.

Understanding the Pump Assembly and Functioning

Inhibiting efflux pumps represents an attractive approach for combating multidrug resistance (Bhardwaj and Mohanty 2012), and the unique tripartite structure provides multiple opportunities for therapeutic action. The proper pump assembly is crucial for the efficient efflux of drugs from the periplasm and the inner membrane to the outside of bacteria (Yu et al. 2003; Eswaran et al. 2004; Bavro et al. 2008). Therefore, it is imperative to elucidate the assembly pathway and activation mechanisms of the AcrAB-TolC pump for developing strategies to block antibiotic resistance (Nikaido 1998; Poole 2001; Weston et al. 2018). Furthermore, such understanding will also provide insights into the mechanisms of many similar RND tripartite pumps (Sharff et al. 2001; Nikaido 2009; Li et al. 2015; Schuldiner 2018), such as MexAB-OprM (Narita et al. 2003; Llanes et al. 2004; Tsutsumi et al. 2019), MexCD-OprJ (Poole 2001), and MexXY-OprM, all found in *P. aeruginosa* (Masuda et al. 2000; Poole 2001; Hocquet et al. 2003).

Despite long-term study (Koronakis et al. 2000; Murakami et al. 2002, 2006; Koronakis 2003; Yu et al. 2003; Fernandez-Recio et al. 2004; Tamura et al. 2005; Mikolosko et al. 2006; Seeger et al. 2006a; Das et al. 2007; Lobedanz et al. 2007; Sennhauser et al. 2007; Bavro et al. 2008; Drew et al. 2009; Pei et al. 2011; Yao et al. 2020), the whole processes of the AcrAB-TolC tripartite assembly, conformational dynamics, and active antibiotic efflux process remain incompletely understood, thereby limiting our knowledge of their role in bacterial pathogenesis and antibiotic resistance. Targeting the assembly and function of efflux pumps, particularly the RND tripartite pumps, provides a promising avenue to inhibit these clinically significant efflux systems (Eisaku and Hidatoshi 2007; Bhardwaj and Mohanty 2012; Atta-ur-Rahman and Choudhary 2013) and overcome efflux-mediated antibiotic resistance in gram-negative bacteria (Webber and Piddock 2002). Therefore, a structural understanding of the pump assembly and active efflux processes in situ is crucial for the treatment of infectious diseases (Nikaido 1998; Poole 2005; Li et al. 2015; Cao et al. 2018).

Assembling Sequence of Tripartite Pump: Historical Background and In Situ Imaging

Study on the Interactions Between the Components of AcrAB-TolC by Biochemical Methods

After the discovery that AcrA, AcrB, and TolC form a stable complex spanning the bacterial cell envelope (Tikhonova and Zgurskaya 2004; Touzé et al. 2004), several biochemical methods have been employed to investigate the interactions between these three components and unravel the assembly sequence of the AcrAB-TolC

efflux pump. In vivo cross-linking experiments mediated by the chemical cross-linker dithiobis-succinimidyl propionate (DSP) have revealed that AcrA cross-linked with AcrB in the absence of TolC (Zgurskaya and Nikaido 2000; Touzé et al. 2004), and both AcrA and AcrB were independently cross-linked to TolC (Touzé et al. 2004). Isothermal titration calorimetry data showed the interactions between AcrA and AcrB, as well as between AcrA and TolC, but no interaction was found between AcrB and TolC (Touzé et al. 2004). Surface plasmon resonance results supported the assembly of a complex not only between AcrA and AcrB (Tikhonova et al. 2009, 2011) but also between AcrB and TolC (Tikhonova et al. 2011). These biochemical findings highlight the bipartite interactions of AcrA-AcrB and AcrA-TolC, but there is no consensus regarding the interaction between AcrB and TolC. Subsequent in vitro structural studies of the fully assembled AcrAB-TolC pumps confirmed the absence of direct contact between TolC and AcrB (Du et al. 2014; Daury et al. 2016; Wang et al. 2017), suggesting two possible pathways for the AcrAB-TolC efflux pump to assemble. One pathway involves the initial formation of a subcomplex between TolC and AcrA, which recruits AcrB to complete the full pump assembly. The other pathway involves the formation of a subcomplex between AcrA and AcrB followed by the recruitment of TolC to form the full tripartite pump.

In Situ Imaging of the AcrA-AcrB Bipartite in *E. coli* Cells

If the assembly intermediate of the AcrAB-TolC efflux pump can be visualized within the cell, the assembly sequence of the pump becomes instantly clear. In situ electron cryotomography (cryo-ET) is a powerful approach that can visualize macromolecules directly inside cells (Sarkar et al. 2014; Asano et al. 2016; Beck and Baumeister 2016). It allows for the capture of various conformations and assembly intermediates of protein complexes, such as secretion systems, in the native cellular environments (Oikonomou et al. 2016; Hu et al. 2017; O'Reilly et al. 2020; Chmielewski et al. 2021). Furthermore, continuous development of the cryo-ET pipeline, including sample preparation, data collection, and data processing, has pushed the resolution of cryo-ET to the subnanometer and atomic level (Schmid 2011; Oikonomou et al. 2016; Chreifi et al. 2019; Zhang 2019; Chmielewski et al. 2021; Sibert et al. 2021; Tegunov et al. 2021; Chen et al. 2022; Kaplan et al. 2023).

Although the biochemical results support the existence of bipartite complexes AcrA-AcrB and AcrA-TolC (Zgurskaya and Nikaido 2000; Touzé et al. 2004), cellular cryo-ET of *E. coli* cells has only revealed the AcrA-AcrB bipartite complex together with the fully assembled AcrAB-TolC pump. Interestingly, it has been found that the PG layer is located just above the top of AcrA (Shi et al. 2019). The interaction between AcrA and PG in *E. coli* cells has been further confirmed through DTSSP-mediated cross-linking followed by LC-MS analysis, which has revealed several potential binding sites of AcrA with PG (Shi et al. 2019). Two of these binding sites were subsequently validated using the results obtained from molecular dynamic simulations of the AcrAB-TolC complex in the presence of the PG layer (Gumbart et al. 2021).

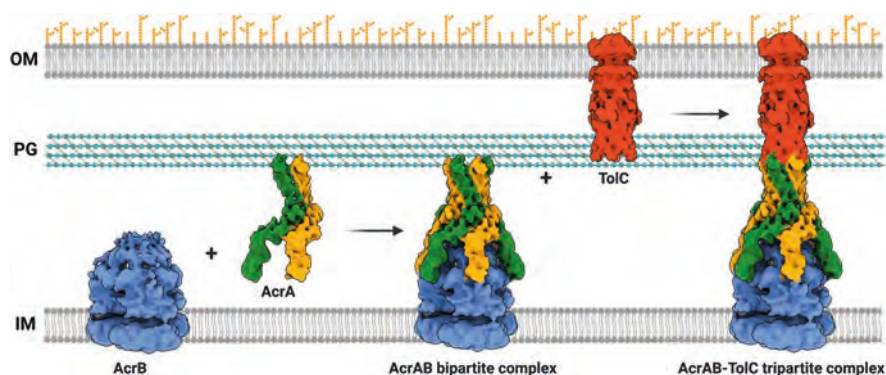


Fig. 1.2 The assembly pathway of the AcrAB-TolC efflux pump. The assembly process of the AcrAB-TolC efflux pump can be summarized as follows: initially, AcrA and AcrB associate with each other to form the AcrA-AcrB bipartite complex. The subcomplex is associated with the PG layer in the cell. Subsequently, TolC is recruited, completing the assembly of the AcrAB-TolC tripartite complex. *OM* outer membrane; *IM* inner membrane; *PG* peptidoglycan

It was discovered many years ago that the purified AcrA presents a specific affinity with the purified PG from *E. coli*, indicating a potential interaction between the AcrAB-TolC efflux pump and the PG layer in the cell (Xu et al. 2012). The in situ structure of the AcrA-AcrB bipartite provides the first direct visual evidence for this subcomplex as a stable entity existing in bacteria. Combining the findings from cryo-ET and in vivo interaction data, a sequential assembly process of the AcrAB-TolC pump has been proposed (Shi et al. 2019). According to this model, the AcrA-AcrB subcomplex forms first, followed by the recruitment of TolC, with the PG layer potentially serving as an anchor to maintain the stability of the AcrA-AcrB subcomplex or facilitate the recruitment of TolC (Fig. 1.2). However, there is currently no evidence suggesting that disrupting the PG layer would impair the efflux function of the previously assembled AcrAB-TolC pumps. Whether the PG layer plays a role in the assembly of the AcrAB-TolC efflux pump still requires experimental validation. Interestingly, recent findings have indicated that the MexAB-OprM efflux pump from *P. aeruginosa* also interacts with the PG layer (Ma et al. 2021), hinting at a universal interaction between the RND efflux pump and the PG layer. If this hypothesis holds true, the interaction interface between these efflux pumps and the PG layer could become a valuable target for the design of novel inhibitors.

Mechanisms of the Efflux Action of AcrAB-TolC

Energization of Efflux Pump

AcrAB-TolC-mediated transport occurs against the substrate concentration gradient, enabling the movement of the substrate from a low-concentration to a high-concentration environment. This transport requires external energy, and the RND

transporters utilize the electrochemical potential of protons, known as the PMF, to drive their substrate-transport processes (Delmar et al. 2014). In the AcrAB-TolC tripartite pump, the D407-D408 pair in AcrB serves as the proton binding site and functions as the energy transduction center. Mutations on these two sites lead to loss of efflux activity of the assembled pump in bacteria (Su et al. 2006; Yue et al. 2017). This proton relay pathway is supported by surrounding lipids (Qiu et al. 2018).

The AcrB transmembrane domain exhibits distinct conformations among its three protomers. In the loose (L) and tight (T) conformations as we mentioned earlier, the side chains of Lys940 from TM10 form salt bridges with side chains of Asp407 and Asp408 from TM4. While in the open (O) conformation, the side chain of Lys940 rotates to the other side, establishing interaction with Asn941 and Thr978 from TM11, coupled with the rigid movements of TM5 and TM8 toward TM10 (Murakami et al. 2006; Seeger et al. 2006b; Sennhauser et al. 2007). Additionally, in the L and T conformations, the upper part of TM8 unwinds and forms loops. In contrast, as in the O conformation, the upper part of TM8 rewinds into an alpha helix, coupled with PC2 movement toward PC1 (Sennhauser et al. 2007). Therefore, the proton transportation induces conformational changes in the AcrB transmembrane domain, which is coupled with large changes in the allostery of the AcrB periplasmic domains.

Studying conformational changes of AcrB and the function of the full pump under its native cell membrane environment, where the PMF exists, is of paramount importance. A recent study employing styrene-maleic acid (SMA) copolymer membrane protein extraction (Lee et al. 2016) has revealed a well-organized lipid-bilayer structure associated with transmembrane domains of the AcrB trimer (Qiu et al. 2018). This SMA purification method offers the advantage of retaining a significant portion of the natural lipids associated with AcrB in contrast to other purification strategies that employ recombination in a nanodisc lipid environment (Bayburt and Sligar 2010), bicelle (Ujwal and Bowie 2011), or saposin (Frauenfeld et al. 2016) that requires the use of detergents at some stage, thereby compromising native membrane environment during the sample preparation. Another promising strategy involves using liposomes to mimic near-physiological conditions (Yao et al. 2020), which could facilitate the future investigation of AcrB in the presence of electrochemical gradients. This approach shows promise in enhancing our understanding of the functional behavior of AcrB within a context that closely mimics its natural environment. However, investigating fully assembled pumps using this approach still presents challenges.

The TolC Periplasmic End Open-Close Switch in the Efflux Pump

As previously mentioned, the AcrB trimer cycles through its three different conformations to facilitate unidirectional substrate pumping from the periplasm or inner membrane into the AcrAB-TolC channel. Consequently, the porter domain of AcrB serves as a natural gate that potentially regulates substrate transportation. In contrast, the periplasmic end of TolC represents another open-close switch that is

thought to regulate substrate transportation. However, in all in vitro structures of purified efflux pumps, TolC exhibits an open conformation regardless of substrate binding (Sharff et al. 2001; Yu et al. 2003; Eswaran et al. 2004).

In the in situ structure of AcrAB-TolC in the presence of the antibiotic puromycin (Shi et al. 2019), a closed TolC periplasmic end was observed (Fig. 1.3a). In this experimental setup, the pump is actively transporting puromycin, as verified by the MIC assay using the same batch of bacteria (Shi et al. 2019), indicating that AcrB is actively changing through its LTO conformations (Murakami et al. 2006; Seeger et al. 2006b; Wang et al. 2017). It was proposed that substrate transportation is a transient process, and the TolC periplasmic end closes after each round so that most of the particles observed have a closed TolC at a certain time point (Shi et al. 2019).

In the in situ structure of AcrAB-TolC in the presence of the AcrB inhibitor MBX3132, TolC adopts an open state, and the AcrAB-TolC assembly shows a shorter length compared to the closed state (Fig. 1.3b) (Shi et al. 2019). In this scenario, the AcrB trimer is expected to exist in its TTT conformation, as the MBX3132 inhibitor binds AcrB in its deep binding pocket (Sjuts et al. 2016). The differences between the open and closed states of TolC are readily observed under cryo-ET imaging, with a sufficiently defined 30–40 Å difference in the chamber diameter in the cryo-ET density maps (Shi et al. 2019). A high-resolution in situ structure of AcrAB-TolC in the presence of MBX3132 has been achieved at a resolution of 7 Å (Fig. 1.4). This structure reveals that the MPD of AcrA protomer II interacts with

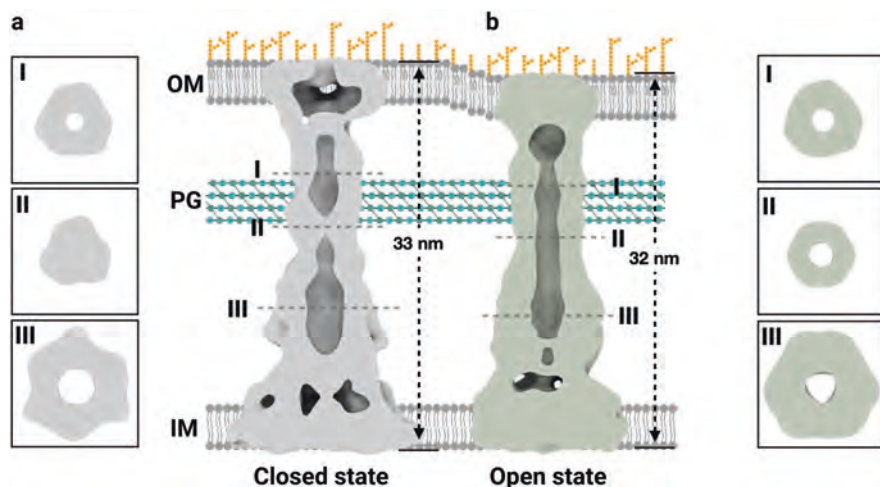


Fig. 1.3 The closed and open states of the AcrAB-TolC efflux pump in situ. **(a)** The closed state of AcrAB-TolC, in the presence of puromycin (EMD-0532). The density map is cross-sectioned at three different positions, indicated by the gray dashed lines, and the cross-sectional views are shown on the left, from the top to the bottom, respectively. **(b)** The open state of AcrAB-TolC, in the presence of MBX3132 (EMD-0533). The density map is cross-sectioned at three different positions, indicated by the gray dashed lines, and the cross-sectional views are shown on the right, from the top to the bottom, respectively. OM outer membrane; IM inner membrane; PG peptidoglycan. (Adapted from Shi et al. 2019)

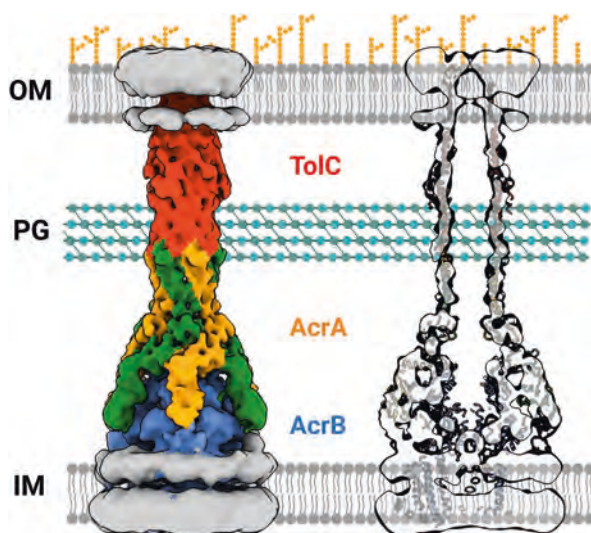


Fig. 1.4 The high-resolution in situ structure of AcrAB-TolC. The left panel shows the overview of the structure (EMD-24609), and the right panel shows a central slice view, with the in vitro structural model fitted (PDB: 5NG5). *OM* outer membrane; *IM* inner membrane; *PG* peptidoglycan. (Adapted from Chen et al. 2022)

the PC2 subdomain of AcrB (Chen et al. 2022), which undergoes significant movement during the transition between the LTO conformations (Wang et al. 2017). Additionally, the α -helical domain of AcrA interacts with the bottom of TolC. These findings provide a possible explanation for the coupling of conformational changes between AcrB and TolC.

Conclusions and Outlook

Future AcrAB-TolC Pump Studies and Cryo-ET In Situ Imaging

Although in situ cryo-ET imaging has provided valuable insights into the assembly process of efflux pumps and their structural variations in different bacterial contexts, the reasons for the constitutive expression of *acrA*, *acrB*, and *tolC* genes and the factors that trigger the assembly of this complex system in living bacteria continue to remain elusive. According to biochemical data, the assembly of AcrAB-TolC is independent of drug presence (Tikhonova and Zgurskaya 2004; Touzé et al. 2004). However, it remains possible that the assembly of AcrAB-TolC relies on unidentified intracellular substrates of this pump. A recent study has proposed oxidized fatty acids as one of the potential native substrates of AcrB (Wang-Kan et al. 2021). The AcrA-AcrB subcomplex was captured in situ when *E. coli* cells were

treated with puromycin, but only a fully assembled AcrAB-TolC complex was observed when the cells were treated with the AcrB inhibitor MBX3132 (Shi et al. 2019). The existence of the AcrA-AcrB subcomplex in the native cellular environment without any drug treatment remains unclear.

The role of the inner membrane potential in regulating the opening and closure of TolC in the AcrAB-TolC tripartite pump within the cell remains a puzzling aspect. The specific local cellular environment of the pump, including PMF condition, ion concentration, surrounding ligands, and other factors in the cell, remains elusive. The in situ structure of AcrAB-TolC upon puromycin and MBX3132 treatment is in a closed or open state, respectively (Shi et al. 2019). But in the native cellular environment without any treatment, the open-close state of TolC itself or TolC in the AcrAB-TolC efflux pump is unknown. Additionally, it remains unclear whether conformational changes in AcrB and the open-close switch of TolC must be coupled to ensure the functionality of the pump. It appears not energetically efficient if TolC opens and closes with each round of the AcrB conformational switch; however, observations showing that TolC remains closed when the in situ pump is treated with puromycin support the notion that TolC actively switches between its open and closed states during substrate transport (Shi et al. 2019).

In summary, the utilization of cryo-ET for in situ imaging has provided a promising avenue to visualize the wide range of intermediate states and conformations of efflux pumps within bacterial cells. This technique holds immense potential for conducting thorough investigations into the assembly and working mechanism of the AcrAB-TolC pump, taking into consideration all the native environmental factors. By further developing techniques to manipulate cell conditions and modify these environmental factors, a wealth of untapped possibilities exists to explore and uncover new insights about the pump's behaviors in vivo.

Acknowledgments The figures are created with BioRender.com. We thank Valerie Dalton for polishing the language of the manuscript. We thank cryo-EM cores at UTHealth Houston and BCM, which includes equipment purchased under the support of CPRIT Core Facility Award RP190602. X.S. was supported by the National Natural Science Foundation of China (No. 82072312) and the Natural Science Foundation of Jiangsu Province (No. BK20211053). Z.W. was supported by the Welch Foundation (Q-1967-20180324; Q-2173-20230405), R01GM143380, and R01HL162842.

References

- Alav I, Kobyłka J, Kuth MS, Pos KM, Picard M, Blair JMA et al (2021) Structure, assembly, and function of tripartite efflux and type 1 secretion Systems in Gram-Negative Bacteria. *Chem Rev* 121(9):5479–5596
- Andersen C, Koronakis E, Bokma E, Eswaran J, Humphreys D, Hughes C et al (2002) Transition to the open state of the TolC periplasmic tunnel entrance. *Proc National Acad Sci* 99(17):11103–11108
- Antão EM, Vincze S, Hanke R, Klimmek L, Suchecka K, Lübke-Becker A et al (2018) Antibiotic resistance, the 3As and the road ahead. *Gut Pathog* 10(1):52

- Asano S, Engel BD, Baumeister W (2016) In situ Cryo-electron tomography: a post-reductionist approach to structural biology. *J Mol Biol* 428(2):332–343
- Atta-ur-Rahman, Choudhary MI (2013) *Frontiers in Anti-Infective Drug Discovery* [Internet]
- Bavro VN, Pietras Z, Furnham N, Pérez-Cano L, Fernández-Recio J, Pei XY et al (2008) Assembly and channel opening in a bacterial drug efflux machine. *Mol Cell* 30(1):114–121
- Bayburt TH, Sligar SG (2010) Membrane protein assembly into Nanodiscs. *FEBS Lett* 584(9):1721–1727
- Beck M, Baumeister W (2016) Cryo-electron tomography: can it reveal the molecular sociology of cells in atomic detail? *Trends Cell Biol* 26(11):825–837
- Bhardwaj AK, Mohanty P (2012) Bacterial efflux pumps involved in multidrug resistance and their inhibitors: Rejuvenating the antimicrobial chemotherapy. *Recent Patents Anti-infective Drug Discov* 7(1):73–89
- Blair JM, Richmond GE, Piddock LJ (2014) Multidrug efflux pumps in gram-negative bacteria and their role in antibiotic resistance. *Future Microbiol* 9(10):1165–1177
- Blair JMA, Smith HE, Ricci V, Lawler AJ, Thompson LJ, Piddock LJV (2015) Expression of homologous RND efflux pump genes is dependent upon AcrB expression: implications for efflux and virulence inhibitor design. *J Antimicrob Chemoth*. 70(2):424–431
- Cao H, Xia T, Li Y, Xu Z, Bougouffa S, Lo YK et al (2018) A multidrug resistant clinical *P. Aeruginosa* isolate in the MLST550 clonal complex: uncoupled quorum sensing modulates the interplay of virulence and resistance. *Biorxiv*:415000
- Chen M, Shi X, Yu Z, Fan G, Serysheva II, Baker ML et al (2022) In situ structure of the AcrAB-TolC efflux pump at subnanometer resolution. *Structure* 30(1):107–113.e3
- Chmielewski D, Schmid M, Simmons G, Jin J, Chiu W. In situ alphavirus assembly and budding mechanism revealed by cellular CryoET. 2021;
- Chreifi G, Chen S, Metskas LA, Kaplan M, Jensen GJ (2019) Rapid tilt-series acquisition for electron cryotomography. *J Struct Biol* 205(Bioinformatics 27 2011):163–169
- Das D, Xu QS, Lee JY, Ankoudinova I, Huang C, Lou Y et al (2007) Crystal structure of the multidrug efflux transporter AcrB at 3.1 Å resolution reveals the N-terminal region with conserved amino acids. *J Struct Biol* 158(3):494–502
- Daury L, Orange F, Taveau JC, Verchère A, Monlezun L, Gounou C et al (2016) Tripartite assembly of RND multidrug efflux pumps. *Nat Commun* 7(1):10731
- Delmar JA, Su CC, Yu EW (2014) Bacterial multidrug efflux transporters. *Biophysics* 43(1):93–117
- Drew D, Klepsch MM, Newstead S, Flaig R, Gier JWD, Iwata S et al (2009) The structure of the efflux pump AcrB in complex with bile acid. *Mol Membr Biol* 25(8):677–682
- Du D, van Veen HW, Murakami S, Pos KM, Luisi BF (2015) Structure, mechanism and cooperation of bacterial multidrug transporters. *Curr Opin Struc Biol*. 33:76–91
- Du D, Wang Z, James NR, Voss JE, Klimont E, Ohene-Agyei T et al (2014) Structure of the AcrAB–TolC multidrug efflux pump. *Nature* 509(7501):512
- Eicher T, Cha H, Seeger MA, Brandstätter L, El-Delik J, Bohnert JA et al (2012) Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc National Acad Sci*. 109(15):5687–5692
- Eicher T, Seeger MA, Anselmi C, Zhou W, Brandstätter L, Verrey F et al (2014) Coupling of remote alternating-access transport mechanisms for protons and substrates in the multidrug efflux pump AcrB. *elife* 3:e03145
- Eisaku Y, Hidatoshi I (2007) Method or agent for inhibiting the function of efflux pump *Pseudomonas aeruginosa* [Internet]. [cited 2007 May 9]
- Eswaran J, Koronakis E, Higgins MK, Hughes C, Koronakis V (2004) Three's company: component structures bring a closer view of tripartite drug efflux pumps. *Curr Opin Struc Biol* 14(6):741–747
- Fernandez-Recio J, Walas F, Federici L, Pratap JV, Bavro VN, Miguel RN et al (2004) A model of a transmembrane drug-efflux pump from gram-negative bacteria. *FEBS Lett* 578(1–2):5–9
- Frauenfeld J, Löving R, Armache JP, Sonnen AFP, Guettou F, Moberg P et al (2016) A saposin-lipoprotein nanoparticle system for membrane proteins. *Nat Methods* 13(4):345–351
- Ge Q, Yamada Y, Zgurskaya H (2009) The C-terminal domain of AcrA is essential for the assembly and function of the multidrug efflux pump AcrAB-TolC. *J Bacteriol* 191(13):4365–4371

- Gröbner S, Linke D, Schütz W, Fladerer C, Madlung J, Autenrieth IB et al (2009) Emergence of carbapenem-non-susceptible extended-spectrum β -lactamase-producing *Klebsiella pneumoniae* isolates at the university hospital of Tübingen, Germany. *J Med Microbiol* 58(7):912–922
- Gumbart JC, Ferreira JL, Hwang H, Hazel AJ, Cooper CJ, Parks JM et al (2021) Lpp positions peptidoglycan at the AcrA-TolC interface in the AcrAB-TolC multidrug efflux pump. *Biophys J* 120(18):3973–3982
- Hocquet D, Vogne C, Garch FE, Vejux A, Gotoh N, Lee A et al (2003) MexXY-OprM efflux pump is necessary for a adaptive resistance of *Pseudomonas aeruginosa* to aminoglycosides. *Antimicrob Agents Ch.* 47(4):1371–1375
- Hu B, Lara-Tejero M, Kong Q, Galán JE, Liu J (2017) In situ molecular architecture of the salmonella type III secretion machine. *Cell* 168(6):1065–1074.e10
- Kaplan M, Chang YW, Oikonomou CM, Nicolas WJ, Jewett AI, Kreida S et al (2023) *Bdellovibrio* predation cycle characterized at nanometre-scale resolution with cryo-electron tomography. *Nat Microbiol* 8(7):1267–1279
- Kobyłka J, Kuth MS, Müller RT, Geertsma ER, Pos KM (2019) AcrB: a mean, keen, drug efflux machine. *Ann N Y Acad Sci* 1459(1):38–68
- Koronakis V (2003) TolC—the bacterial exit duct for proteins and drugs. *FEBS Lett* 555(1):66–71
- Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405(6789):914–919
- Lee SC, Knowles TJ, Postis VLG, Jamshad M, Parslow RA, Lin Y et al (2016) A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nat Protoc* 11(7):1149–1162
- Li XZ, Plésiat P, Nikaido H (2015) The challenge of efflux-mediated antibiotic resistance in gram-negative bacteria. *Clin Microbiol Rev* 28(2):337–418
- Llanes C, Hocquet D, Vogne C, Benali-Baitich D, Neuwirth C, Plésiat P (2004) Clinical strains of *Pseudomonas aeruginosa* overproducing MexAB-OprM and MexXY efflux pumps simultaneously. *Antimicrob Agents Ch.* 48(5):1797–1802
- Lobedanz S, Bokma E, Symmons MF, Koronakis E, Hughes C, Koronakis V (2007) A periplasmic coiled-coil interface underlying TolC recruitment and the assembly of bacterial drug efflux pumps. *Proc National Acad Sci.* 104(11):4612–4617
- Ma M, Lustig M, Salem M, Mengin-Lecreulx D, Phan G, Broutin I (2021) MexAB-OprM efflux pump interaction with the peptidoglycan of *Escherichia coli* and *Pseudomonas aeruginosa*. *Int J Mol Sci* 22(10):5328
- Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA et al (2014) Multistate point-prevalence survey of health care-associated infections. *New Engl J Med* 370(13):1198–1208
- Masuda N, Sakagawa E, Ohya S, Gotoh N, Tsujimoto H, Nishino T (2000) Substrate specificities of MexAB-OprM, MexCD-OprJ, and MexXY-OprM efflux pumps in *Pseudomonas aeruginosa*. *Antimicrob Agents Ch.* 44(12):3322–3327
- Mikolosko J, Bobyk K, Zgurskaya HI, Ghosh P (2006) Conformational flexibility in the multidrug efflux system protein AcrA. *Structure* 14(3):577–587
- Murakami S, Nakashima R, Yamashita E, Matsumoto T, Yamaguchi A (2006) Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* 443(7108):173–179
- Murakami S, Nakashima R, Yamashita E, Yamaguchi A (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* 419(6907):587–593
- Nakashima R, Sakurai K, Yamasaki S, Nishino K, Yamaguchi A (2011) Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket. *Nature* 480(7378):565–569
- Narita S, Eda S, Yoshihara E, Nakae T (2003) Linkage of the efflux-pump expression level with substrate extrusion rate in the MexAB-OprM efflux pump of *Pseudomonas aeruginosa*. *Biochem Bioph Res Co* 308(4):922–926
- Nikaido H (1996) Multidrug efflux pumps of gram-negative bacteria. *J Bacteriol* 178(20):5853–5859
- Nikaido H (1998) Antibiotic resistance caused by gram-negative multidrug efflux pumps. *Clin Infect Dis* 27(s1):S32–S41

- Nikaido H (2009) Multidrug resistance in bacteria. *Biochemistry* 78(1):119–146
- Oikonomou CM, Chang YW, Jensen GJ (2016) A new view into prokaryotic cell biology from electron cryotomography. *Nat Rev Microbiol* 14(4):205–220
- O'Reilly FJ, Xue L, Graziadei A, Sinn L, Lenz S, Tegunov D et al (2020) In-cell architecture of an actively transcribing-translating expressome. *Science* 369(6503):554–557
- Pei XY, Hinchliffe P, Symmons MF, Koronakis E, Benz R, Hughes C et al (2011) Structures of sequential open states in a symmetrical opening transition of the TolC exit duct. *Proc National Acad Sci* 108(5):2112–2117
- Piddock LJV (2006) Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. *Clin Microbiol Rev* 19(2):382–402
- Pitout JDD, Nordmann P, Laupland KB, Poirel L (2005) Emergence of Enterobacteriaceae producing extended-spectrum β -lactamases (ESBLs) in the community. *J Antimicrob Chemoth* 56(1):52–59
- Poole K (2001) Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. *J Mol Microb Biotech* 3(2):255–264
- Poole K (2005) Efflux-mediated antimicrobial resistance. *J Antimicrob Chemoth* 56(1):20–51
- Qiu W, Fu Z, Xu GG, Grassucci RA, Zhang Y, Frank J et al (2018) Structure and activity of lipid bilayer within a membrane-protein transporter. *Proc National Acad Sci* 115(51):12985–12990
- Rogers BA, Sidjabat HE, Paterson DL (2011) *Escherichia coli* O25b-ST131: a pandemic, multiresistant, community-associated strain. *J Antimicrob Chemoth* 66(1):1–14
- Sarkar P, Bosneaga E, Yap EG, Das J, Tsai WT, Cabal A et al (2014) Electron tomography of Cryo-immobilized plant tissue: a novel approach to studying 3D macromolecular architecture of mature plant cell walls in situ. *PLoS One* 9(9):e106928
- Schindler BD, Kaatz GW (2016) Multidrug efflux pumps of gram-positive bacteria. *Drug Resist Update* 27:1–13
- Schmid MF (2011) Single-particle electron cryotomography (cryoET). *Adv Protein Chem Str* 82:37–65
- Schuldiner S (2018) The *Escherichia coli* effluxome. *Res Microbiol* 169(7–8):357–362
- Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, Pos KM (2006a) Asymmetric structure of trimeric AcrB from *Escherichia coli*.
- Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, Pos KM (2006b) Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science* 313(5791):1295–1298
- Sennhauser G, Amstutz P, Briand C, Storchenegger O, Grütter MG (2007) Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS Biol* 5(1):e7
- Sharff A, Fanutti C, Shi J, Calladine C, Luisi B (2001) The role of the TolC family in protein transport and multidrug efflux. *Eur J Biochem* 268(19):5011–5026
- Shi X, Chen M, Yu Z, Bell JM, Wang H, Forrester I et al (2019) In situ structure and assembly of the multidrug efflux pump AcrAB-TolC. *Nat Commun* 10(1):2635
- Sibert BS, Kim JY, Yang JE, Wright ER (2021) Micropatterning transmission electron microscopy grids to direct cell positioning within whole-cell Cryo-electron tomography workflows. *J Vis Exp Jove* (175). <https://doi.org/10.3791/62992>
- Silhavy TJ, Kahne D, Walker S (2010) The bacterial cell envelope. *Csh Perspect Biol* 2(5):a000414
- Singh R, Swick MC, Ledesma KR, Yang Z, Hu M, Zechiedrich L et al (2012) Temporal interplay between efflux pumps and target mutations in development of antibiotic resistance in *Escherichia coli*. *Antimicrob Agents Ch* 56(4):1680–1685
- Sjuts H, Vargiu AV, Kwasny SM, Nguyen ST, Kim HS, Ding X et al (2016) Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proc National Acad Sci* 113(13):3509–3514
- Su CC, Li M, Gu R, Takatsuka Y, McDermott G, Nikaido H et al (2006) Conformation of the AcrB multidrug efflux pump in mutants of the putative proton relay pathway. *J Bacteriol* 188(20):7290–7296
- Swick MC, Morgan-Linnell SK, Carlson KM, Zechiedrich L (2011) Expression of multidrug efflux pump genes *acrAB-tolC*, *mdfA*, and *norE* in *Escherichia coli* clinical isolates as a function of fluoroquinolone and multidrug resistance $\pm$. *Antimicrob Agents Ch* 55(2):921–924

- Symmons MF, Marshall RL, Bavro VN (2015) Architecture and roles of periplasmic adaptor proteins in tripartite efflux assemblies. *Front Microbiol* 6:513
- Takatsuka Y, Nikaido H (2009) Covalently linked trimer of the AcrB multidrug efflux pump provides support for the functional rotating mechanism. *J Bacteriol* 191(6):1729–1737
- Tamura N, Murakami S, Oyama Y, Ishiguro M, Yamaguchi A (2005) Direct interaction of multidrug efflux transporter AcrB and outer Membrane Channel TolC detected via site-directed disulfide cross-linking †. *Biochemistry* 44(33):11115–11121
- Tegunov D, Xue L, Dienemann C, Cramer P, Mahamid J (2021) Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells. *Nat Methods* 18(2):186–193
- Tikhonova EB, Dastidar V, Rybenkov VV, Zgurskaya HI (2009) Kinetic control of TolC recruitment by multidrug efflux complexes. *Proc National Acad Sci*. 106(38):16416–16421
- Tikhonova EB, Yamada Y, Zgurskaya HI (2011) Sequential mechanism of assembly of multidrug efflux pump AcrAB-TolC. *Chem Biol* 18(4):454–463
- Tikhonova EB, Zgurskaya HI (2004) AcrA, AcrB, and TolC of *Escherichia coli* form a stable inter-membrane multidrug efflux complex*. *J Biol Chem* 279(31):32116–32124
- Touzé T, Eswaran J, Bokma E, Koronakis E, Hughes C, Koronakis V (2004) Interactions underlying assembly of the *Escherichia coli* AcrAB–TolC multidrug efflux system. *Mol Microbiol* 53(2):697–706
- Tsukagoshi N, Aono R (2000) Entry into and release of solvents by *Escherichia coli* in an organic-aqueous two-liquid-phase system and substrate specificity of the AcrAB-TolC solvent-extruding pump. *J Bacteriol* 182(17):4803–4810
- Tsutsumi K, Yonehara R, Ishizaka-Ikeda E, Miyazaki N, Maeda S, Iwasaki K et al (2019) Structures of the wild-type MexAB–OprM tripartite pump reveal its complex formation and drug efflux mechanism. *Nat Commun* 10(1):1520
- Ujwal R, Bowie JU (2011) Crystallizing membrane proteins using lipidic bicelles. *Methods* 55(4):337–341
- Viale P, Giannella M, Tedeschi S, Lewis R (2015) Treatment of MDR-gram negative infections in the 21st century: a never ending threat for clinicians. *Curr Opin Pharmacol* 24:30–37
- Wang Z, Fan G, Hryc CF, Blaza JN, Serysheva II, Schmid MF et al (2017) An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump. *elife* 6:e24905
- Wang-Kan X, Rodríguez-Blanco G, Southam AD, Winder CL, Dunn WB, Ivens A et al (2021) Metabolomics reveal potential natural substrates of AcrB in *Escherichia coli* and *salmonella enterica* Serovar typhimurium. *MBio* 12(2):e00109–e00121
- Webber MA, Piddock LJV (2002) The importance of efflux pumps in bacterial antibiotic resistance. *J Antimicrob Chemoth*. 51(1):9–11
- Weston N, Sharma P, Ricci V, Piddock L (2018) Regulation of the AcrAB-TolC efflux pump in Enterobacteriaceae. *Res Microbiol* 169(7–8):425–431
- Xu Y, Moeller A, Jun SY, Le M, Yoon BY, Kim JS et al (2012) Assembly and channel opening of outer membrane protein in tripartite drug efflux pumps of gram-negative bacteria*. *J Biol Chem* 287(15):11740–11750
- Yang S, Lopez CR, Zechiedrich EL (2006) Quorum sensing and multidrug transporters in *Escherichia coli*. *Proc National Acad Sci*. 103(7):2386–2391
- Yao X, Fan X, Yan N (2020) Cryo-EM analysis of a membrane protein embedded in the liposome. *Proc National Acad Sci*. 117(31):18497–18503
- Yu EW, McDermott G, Zgurskaya HI, Nikaido H Jr, DEK. (2003) Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. *Science* 300(5621):976–980
- Yue Z, Chen W, Zgurskaya HI, Shen J (2017) Constant pH molecular dynamics reveals how proton release drives the conformational transition of a transmembrane efflux pump. *J Chem Theory Comput* 13(12):6405–6414
- Zgurskaya HI, Nikaido H (1999) AcrA is a highly asymmetric protein capable of spanning the periplasm 1 | Edited by I. B. Holland. *J Mol Biol* 285(1):409–420
- Zgurskaya HI, Nikaido H (2000) Cross-linked complex between oligomeric periplasmic lipoprotein AcrA and the inner-membrane-associated multidrug efflux pump AcrB from *Escherichia coli*. *J Bacteriol* 182(15):4264–4267

- Zgurskaya HI, Yamada Y, Tikhonova EB, Ge Q, Krishnamoorthy G (2009) Structural and functional diversity of bacterial membrane fusion proteins. *Biochimica Et Biophysica Acta Bba-Proteins Proteom* 1794(5):794–807
- Zhang P (2019) Advances in cryo-electron tomography and subtomogram averaging and classification. *Curr Opin Struc Biol*. 58:249–258
- Zwama M, Yamasaki S, Nakashima R, Sakurai K, Nishino K, Yamaguchi A (2018) Multiple entry pathways within the efflux transporter AcrB contribute to multidrug recognition. *Nat Commun* 9(1):124

Chapter 2

The Copper Efflux Regulator (CueR)



Yangbo Hu and Bin Liu

Abstract The copper efflux regulator (CueR) is a classical member of the MerR family of metalloregulators and is common in gram-negative bacteria. Through its C-terminal effector-binding domain, CueR senses cytoplasmic copper ions to regulate the transcription of genes contributing to copper homeostasis, an essential process for survival of all cells. In this chapter, we review the regulatory roles of CueR in the model organism *Escherichia coli* and the mechanisms for CueR in copper binding, DNA recognition, and interplay with RNA polymerase in regulating transcription. In light of biochemical and structural analyses, we provide molecular details for how CueR represses transcription in the absence of copper ions, how copper ions mediate CueR conformational change to form holo CueR, and how CueR bends and twists promoter DNA to activate transcription. We also characterize the functional domains and key residues involved in these processes. Since CueR is a representative member of the MerR family, elucidating its regulatory mechanisms could help to understand the CueR-like regulators in other organisms and facilitate the understanding of other metalloregulators in the same family.

Keywords Copper homeostasis · Metalloregulator · Transcriptional regulation · Promoter · RNA polymerase · Allosteric activation

Y. Hu (✉)

State Key Laboratory of Virology, Wuhan Institute of Virology, Center for Biosafety
Mega-Science, Chinese Academy of Sciences, Wuhan, China
e-mail: ybhu@wh.iov.cn

B. Liu (✉)

Section of Transcription & Gene Regulation, The Hormel Institute, University of Minnesota,
Austin, MN, USA
e-mail: liu00794@umn.edu

Introduction

Transferring environmental signals into control of gene expression is the key function of transcriptional regulators. Transcription regulators of the MerR family are a large group of dimeric proteins comprised of the N-terminal homogenous helix-turn-helix DNA-binding domain and the C-terminal effector-binding domain; these are specific to various signals, such as oxidative stress, heavy metals, and antibiotics (Brown et al. 2003). They function as transcription repressors under normal conditions, but rapidly transform to transcription activators after sensing corresponding environmental signals (Fang and Zhang 2022). Learning more about the interactions of MerR family regulators with corresponding effectors and target promoters, and their interplay with RNA polymerase (RNAP), the only enzyme in charge of transcription in bacteria (Murakami and Darst 2003; Borukhov and Nudler 2008), is essential for understanding the unique mechanism of such regulators.

MerR family regulators exist in a wide range of bacterial genera (Brown et al. 2003; Fang and Zhang 2022). A representative one, the copper efflux regulator (CueR), is one of the most thoroughly characterized members of the MerR family. Once copper ions are sensed, dimerized CueR acts on the promoters with a long spacer region (19 or 20 bp) between -10 and -35 elements (Stoyanov et al. 2001; Outten et al. 2000; Philips et al. 2015), a distinct characteristic of MerR family target promoters (Outten et al. 2000; Stoyanov et al. 2001). This differs from the 17 bp spacer one recognized by the primary σ^{70} -type RNAP (Paget 2015). Afterward, CueR bends and twists the promoter to facilitate recognition by σ^{70} -type RNAP and subsequent initiation of transcription (Fig. 2.1). In this chapter, we will introduce and summarize what is known about how CueR senses copper ions, how it binds to promoters, and how it influences RNAP binding and actions to initiate transcription.

Copper Homeostasis and CueR

Maintaining cellular metal homeostasis is essential for survival of all cells, whether prokaryotic or eukaryotic. Metal ions play key roles in the structure and function of many enzymes and regulators. On the other hand, high concentrations of metal ions can induce cytotoxicity via Fenton and Haber-Weiss-like reactions to produce biologically lethal hydroxyl radicals and carbonate ions (Waldron et al. 2009; Finney and O'Halloran 2003; Chandrangu et al. 2017; Robinson and Winge 2010). Therefore, intracellular metal ion concentrations must be tightly regulated to maintain normal cell metabolism, in which metalloregulators regulate transcription of genes controlling metal uptake, efflux, or other metal homeostasis genes in response to environmental and intracellular metal ions (Ma et al. 2009; O'Halloran 1993; Summers 2009; Price and Boyd 2020).

Copper is an essential cofactor for multiple key enzymes of important biological processes, such as aerobic respiration or superoxide dismutation process (Andrei et al. 2020). Due to the presence of NADH-linked cupric reductase that reduces Cu^{II}

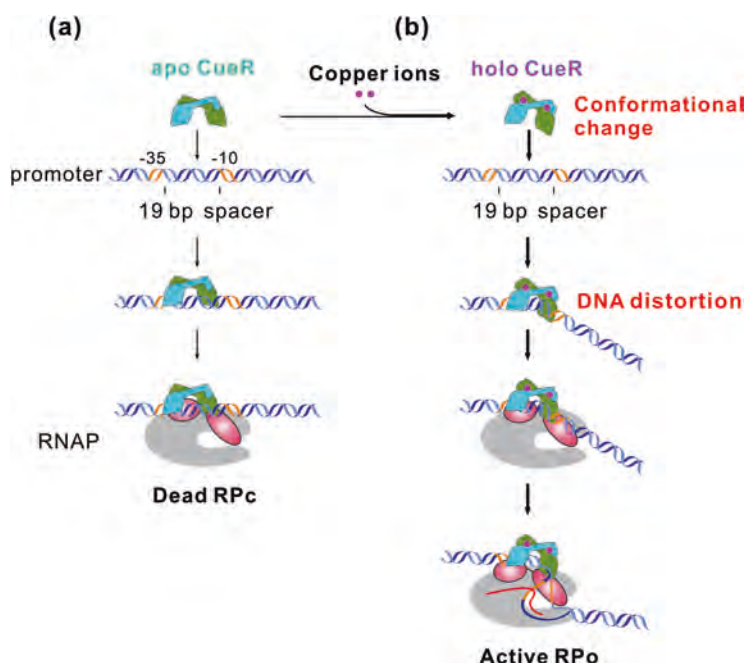


Fig. 2.1 Overview of CueR activating transcription upon copper ion binding. (a) Apo CueR binds promoter DNA to form a dead end complex (RPc). (b) Copper ions bind and modify CueR to form a holo molecule, further twisting promoter DNA to form an active open complex (RPo)

to Cu^I (Rapisarda et al. 1999), intracellular copper is assumed to be present as Cu^I , which could generate hydroxyl radicals and is toxic at high concentrations (Goldstein and Czapski 1986). Bacteria have developed elaborate strategies used for uptake, storage, and export of copper ions. Among them, the P-type ATPase CopA is a copper transporter, whose expression is induced by copper ions and is required for optimum copper tolerance in *E. coli* (Rensing et al. 2000). Three research groups identified the protein encoded by the gene previously known as *ybbI* as the direct activator of *copA* in response to copper stress and renamed it as CueR (or CopR) almost simultaneously (Stoyanov et al. 2001; Outten et al. 2000; Petersen and Moller 2000). Homologs of CueR widely exist in gram-negative bacteria (Yakobov et al. 2022) and control transcription of other genes contributing to copper homeostasis, such as *cueO* (named as Cu efflux oxidase) (Outten et al. 2000) and *cusC* (tripartite copper pump component) (Yamamoto and Ishihama 2005). These data suggest that CueR plays key roles in regulating copper homeostasis in bacteria.

Sequence alignment shows that CueR is a member of the MerR family regulators, as about 70% of its residues are conserved with *Bacillus* MerR protein in DNA-binding domains (Stoyanov et al. 2001) (Fig. 2.2). Further studies confirmed that two monomers of the CueR molecule form a homodimer as the MerR protein (Changela et al. 2003). Additionally, matching the characteristics of the MerR family members (Brown et al. 2003), each monomer of CueR is composed of at least

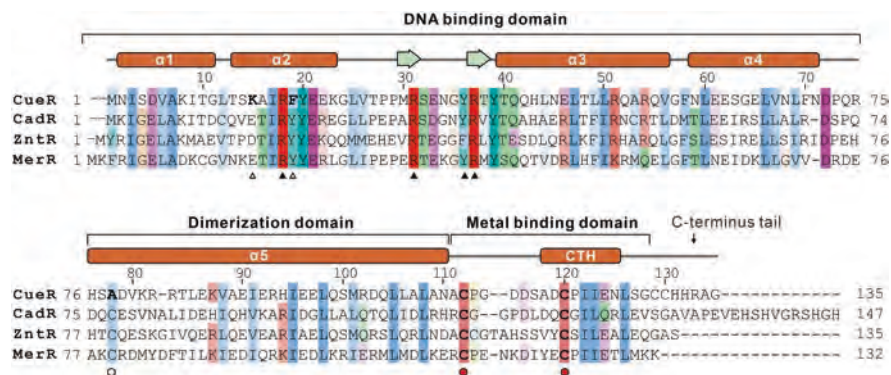


Fig. 2.2 Sequence alignment of four different MerR family members. Individual domains and helices are labeled. Key residues involved in monovalent metal ion binding in CueR are indicated by red circles. Cysteine that specifically contributes to divalent metal ion binding is indicated by a hollow circle. Key residues involved in DNA binding are denoted with triangles

three domains: (i) a N-terminal DNA-binding domain containing two helix-turn-helix motifs, (ii) a long dimerization helix forming an antiparallel coiled-coil structure, and (iii) a C-terminal metal-binding domain characterized by two conserved cysteine residues (Cys112 and Cys120 in *E. coli*) (Fig. 2.2). However, CueR lacks a cysteine residue that is conserved at approximately position 79, which is observed in other MerR family metalloregulators and is essential for divalent metal binding (Outten et al. 2000), but holds two additional cysteine residues at the C-terminus CCHHRAG fragment tail (Fig. 2.2), suggesting that CueR might have different options in metal selection (Balogh et al. 2022).

Metal-Ion Selectivity of CueR

The major function of transcriptional regulators is to transfer environmental signals into gene transcription. In the CueR molecule, the C-terminal metal-binding domain is responsible for copper sensing and binding. When CueR was first characterized, its activity on the *copA* promoter was strongly induced by copper and silver ions, but not by Ni^{II} , Pb^{II} , Cd^{II} , Hg^{II} , or Zn^{II} (Stoyanov et al. 2001), suggesting strong selectivity of CueR in ion specificity. Since copper exists as both Cu^{II} and Cu^{I} forms, and Cu^{II} is easily reduced to Cu^{I} both in vitro and in vivo (Brenner and Harris 1995; Rapisarda et al. 1999), CueR has been supposed to respond to monovalent but not divalent metal ions. Consistent with this hypothesis, CueR-dependent transcription on the *copA* promoter is activated by gold ions (Au^{I}) (Stoyanov and Brown 2003). Further in vitro titration of different metals in a CueR-dependent transcription activation system revealed that *copA* promoter activities rise steadily with increasing amounts of Cu^{I} , Ag^{I} , or Au^{I} (Changela et al. 2003). Consistent with in vivo data (Stoyanov et al. 2001), neither Hg^{II} nor Zn^{II} induces CueR-activated transcription

in vitro (Changela et al. 2003). These studies clearly indicate that CueR can directly distinguish between metal ion charges in gene regulation.

The response of CueR to monovalent metal ions strictly depends on the conserved Cys112 and Cys120 (numbered in *E. coli*) in its C-terminal metal-binding domain. To explore the mechanism of metal ion binding, the structures of Cu^I-, Ag^I-, and Au^I-bound forms of CueR have been determined at 2.2, 2.1, and 2.5 Å resolution, respectively (Changela et al. 2003). Overall structures of these metal-bound CueR molecules are identical (Fig. 2.3). Two CueR molecules form a homodimer sharing the same topology as the MerR family activators: MerR (Chang et al. 2015), BmrR (Heldwein and Brennan 2001), SoxR (Watanabe et al. 2008), and MtaN (Godsey et al. 2001). Coordinate-covalent bonds between the Cu^I ion and the two sulfur atoms of Cys112 and Cys120 exhibit Cu-S bond distances of 2.13 Å (Ag-S, 2.35 Å; Au-S, 2.32 and 2.39 Å, respectively) with an essentially linear S-Cu-S bond angle of 176° (Fig. 2.3). This linear S-Cu-S coordination has also been confirmed in solution by extended X-ray absorption fine structure spectral analyses (Chen et al. 2003).

Structural and biochemical data verified that the surrounding environment does not allow CueR to have other types of interactions with metal ions. Notably, a Zn^{II} ion is bound through Cys79, Cys114, Cys115, His119, and Cys124 in ZntR dimeric

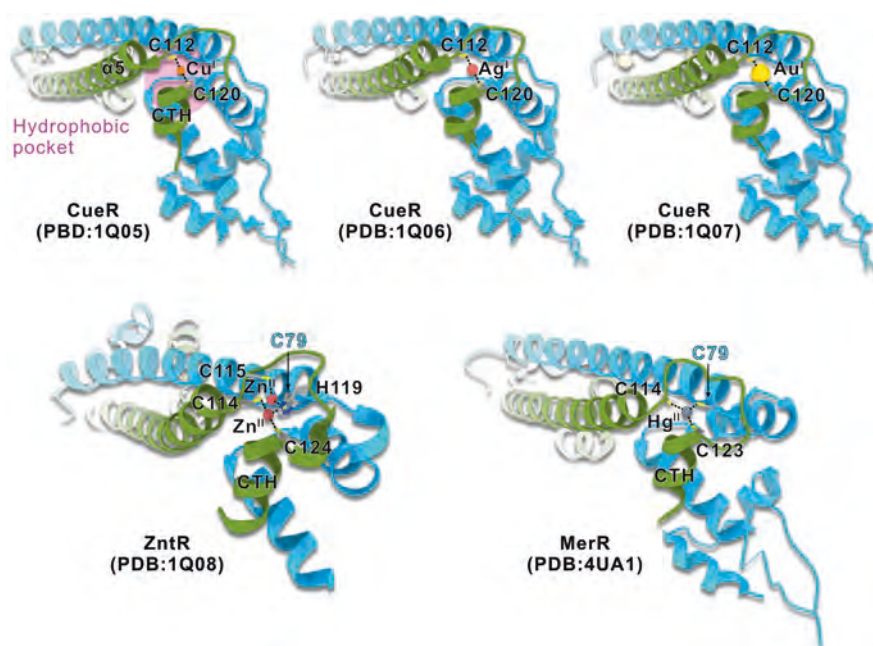


Fig. 2.3 Metal ion recognition of MerR family members. The top three illustrations show forms of CueR that recognize Cu^I, Ag^I, and Au^I. The bottom two illustrations show ZntR and MerR binding of Zn^{II} and Hg^{II}, respectively. Cartoon representations are displayed except for showing metal-binding residues with sticks. The hydrophobic pocket is shown as a pink blot

structures (Fig. 2.3). Similarly, Cys79, Cys114, and Cys123 are coordinated for Hg^{II} binding for MerR (Chang et al. 2015) (Fig. 2.3). Through comparing the structures and sequences of these regulators, it appeared that Cys79 is essential for divalent metal ion binding, and lack of this residue might be an important reason for the strong selectivity of ions for CueR (Fig. 2.2). Additionally, using a computational framework approach, the electrostatic environment of CueR is well-tuned to favor the binding of monovalent metal ions over divalent ions (Rao et al. 2010).

The C-terminus CCHHRAG tail in the CueR molecule also plays a key role in ion selectivity. Despite exhibiting a disorder state in most of structures determined by X-ray diffraction and cryo-electron microscopy (cryo-EM) studies (Changela et al. 2003; Philips et al. 2015; Fang et al. 2021; Shi et al. 2021), the position of the C-terminus tail might place the last C-terminal helix (CTH) domain in the hydrophobic pocket, a key site significant for the binding of CueR with metal ions (Philips et al. 2015) (Fig. 2.3). Although the C-terminus tail is not essential for the function of CueR in activating transcription, the affinity of the metal-binding site for monovalent metal ions is decreased considerably when this tail is deleted (Hobman et al. 2005; Chen et al. 2003), suggesting it is involved in metal binding. In addition, this tail is also implicated in other mechanisms. Intriguingly, Hg^{II} does not activate CueR-associated transcription (Changela et al. 2003), although bis-thiolate metal-binding sites exhibit high affinity for Hg^{II} (Balogh et al. 2019). Based on electrospray ionization mass spectrometry assays, Hg^{II} ions bind to CueR in a trigonal geometry, recruiting a third thiolate ligand from two additional cysteines (Cys129 and Cys130) located at the C-terminus CCHHRAG tail (Balogh et al. 2019). These two cysteines can provide auxiliary ligands that coordinate CueR to Hg^{II} , thereby modifying the structure of the CueR molecule and then preventing activation of transcription (Balogh et al. 2019; Balogh et al. 2022). This process may explain the specificity of CueR for monovalent metal ions and against divalent metal ions.

Promoter Recognition and Allosteric Activation of CueR

DNA binding is the key process for the regulatory roles of CueR. MerR family regulators are characterized by a helix-turn-helix DNA-binding motif close to the N-terminus of the protein as the presumptive DNA-binding domain. In addition, they are also characterized by acting on the noncanonical σ^{70} -RNAP-recognized promoters that have extended spacer regions (19 or 20 bp, compared to the consensus 17 bp) between the -35 and -10 elements (Ansari et al. 1995). Using DNase I footprinting to characterize the binding sequence of CueR, the palindrome region (ACCTTCC-N7-GGAAGGT) in the *copA* promoter between -35 and -10 elements was reported as the CueR operator (Stoyanov et al. 2001; Outten et al. 2000). The same palindrome sequence was identified in the *cueO* promoter (Outten et al. 2000). Additionally, in light of systematic analysis of the CueR target promoter sequences among different strains, operator bases at positions -22 and -26 may be favored by C/G for CueR specificity (ACCTTCC-n-C/G-nnn-C/G-n-GGAAGGT)

(see Fig. 2.4) (Perez Audero et al. 2010). Based on these features of CueR target promoters, especially the extended spacer region between the promoter -35 and -10 elements, potential CueR-associated promoters in the bacterial genome were identified through sequence analyses (Thaden et al. 2010; Yamamoto and Ishihama 2005).

After identifying the binding sequences of the CueR protein, the next question is how CueR recognizes these promoter sequences. Using single-molecule fluorescence resonance energy transfer (smFRET) measurements, both the apo (without

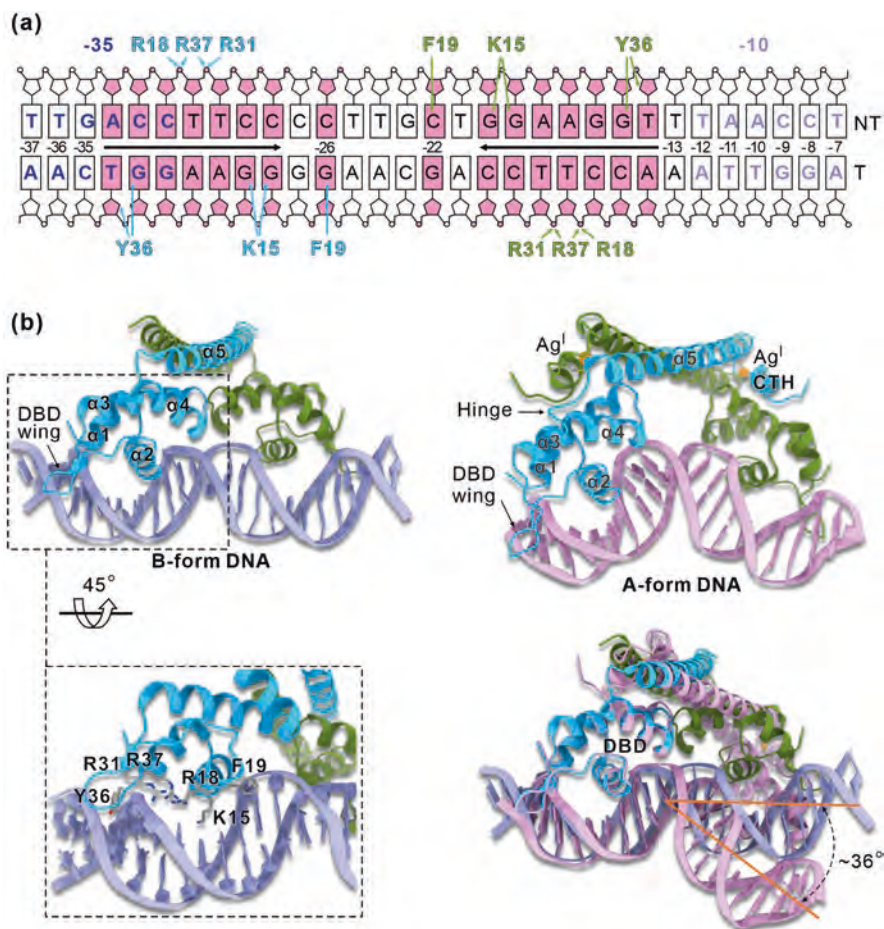


Fig. 2.4 Interaction between CueR protein and target DNA. **(a)** Schematic representation of the *copA* promoter DNA scaffold. The -35 and -10 elements, CueR-binding sites, and residues involved in protein-DNA interactions are denoted. NT, non-template DNA; T, template DNA. **(b)** Views of the structures of apo CueR-DNA and holo CueR-DNA complexes with labeled helices and functional domains. The zoom-in displays the key residues contacting DNA. Bottom right: superimposition of the two structures on one DBD domain. There is a $\sim 36^\circ$ rotation angle between the two DNA conformations

metal ions bound) and holo (with metal ions bound) states of CueR were demonstrated to bind to the specific DNA sequence within the promoter with almost the same affinity (Andoy et al. 2009; Stoyanov et al. 2001). Crystal structures of both CueR-Ag^I (holo form) and the S112-S120-mutated CueR (apo form) in complex with DNA have been determined (Philips et al. 2015). Remarkably, CueR-DNA contacts in both complexes are indistinguishable. In the structures, CueR interacts with phosphate groups at the distal edges of the pseudopalindrome (ACCTTCC-n--C/G-nnn-C/G-n-GGAAGGT) through three arginine residues from helix 2 (α 2) and the loop wing of its N-terminal DNA-binding domain (Arg18, 31, and 37). In addition, Tyr36 inserts into the minor groove to interact with nucleotides located at outer edges of the pseudopalindrome (ACCTTCC-n-C/G-nnn-C/G-n-GGAAGGT), while Lys15 inserts into the major groove and forms H bonds to N7 and O6 of the inner edges of the pseudopalindrome (ACCTTCC-n-C/G-nnn-C/G-n-GGAAGGT). The aromatic ring of Phe19 sits perpendicular to and forms van der Waals contacts with C-22 and C-26 (ACCTTCC-n-C/G-nnn-C/G-n-GGAAGGT) (Fig. 2.4). These CueR-promoter interactions have been further confirmed in the structures of transcription activation complexes in the presence of RNAP (Fang et al. 2021; Shi et al. 2021).

Notably, although apo and holo CueR proteins have similar protein-DNA contacts, the DNA topologies in these two complexes are largely different. In the apo state, the promoter DNA is only slightly bent and belongs to the B-form. Upon binding a metal ion in the allosteric site to form a holo one, CueR maintains the protein-DNA contacts but introduces torsional stresses that kink and twist the promoter, stabilizing an A-form DNA-like conformation (Philips et al. 2015) (Fig. 2.5), first described in the structure of the TBP/DNA complex (Guzikevich-Guerstein and Shakked 1996). Binding of holo CueR alters the trajectory of the helical axis by $\sim 36^\circ$ relative to the repression state bound by apo CueR. In the meantime, apo CueR twists the angle between promoter -10 and -35 elements by about 50° , but the twisted angle by holo CueR is $\sim 72^\circ$, identical to the consensus 17 bp spacer B-form DNA promoter (Fig. 2.5). Binding of apo CueR shortens the distance of the 19 bp spacer region between -10 and -35 elements from 64.2 Å to 59.3 Å, while binding of holo CueR further shortens the distance to 56.4 Å (Fig. 2.5), similar to that of the 17 bp long spacer in B-form (Philips et al. 2015).

The difference in the DNA conformations bound by the apo and holo CueR proteins is a consequence of allosteric change on the conformation of CueR induced by metal ion binding. The allosteric switch starts with conformational change of the C-terminal metal-binding domain induced by metal binding, which triggers displacement of key residues in the hinge between the N-terminal DNA binding and dimerization helix domains (Fig. 2.6). As a result, the hinge residue Arg75 directly interacts with the backbone carbonyl oxygens of Ala118 located at the metal-binding domain as well as Asp72 and Phe70 located at the helix 4 (α 4) of the DNA-binding domain, stabilizing the new dimerization conformation (Philips et al. 2015) (Fig. 2.6). Using smFRET and double electron-electron resonance measurements,

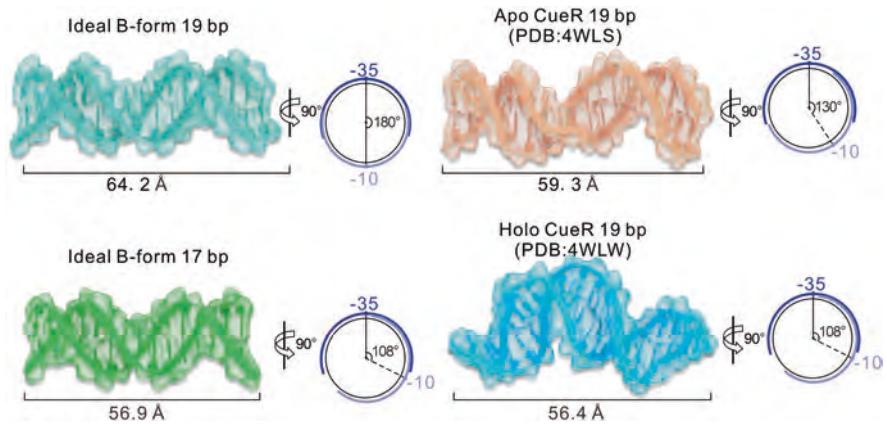


Fig. 2.5 CueR bends, twists, and kinks DNA in regulating transcription. The B-form DNA models of 19 bp and 17 bp spacer DNA are shown in cyan and green, respectively. The DNAs bound by apo CueR and the holo one are shown in orange and sky blue, respectively. The right part in each panel shows the corresponding angle between -10 and -35 elements looking down the central axis

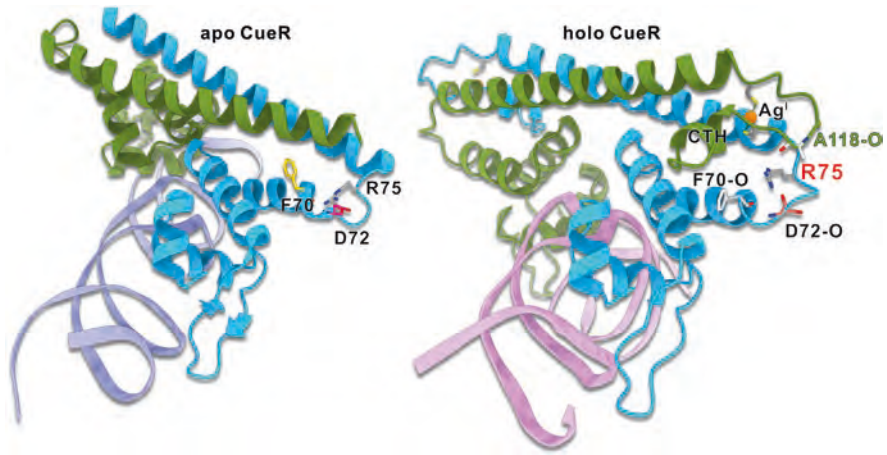


Fig. 2.6 Structural differences of the apo and holo forms of CueR. R75 in the holo CueR interacts with the backbone carbonyl oxygens of F70, D72, and A118

CueR has been shown in a dynamic nature and may undergo a range of conformational states that assist in the transcription process (Joshi et al. 2012; Sameach et al. 2017, 2019; Yakobov et al. 2022). Generally, CueR resides in three states: the apo-CueR state, the apo-CueR-DNA state (a repression state), and the holo-CueR-DNA state (an activation state). The population of holo-CueR-DNA complex increases with increasing concentrations of Cu^I in solution and is dominant at the highest

concentration of Cu^I (Sameach et al. 2017). These analyses suggest that copper is indeed the main trigger that induces the protein to shift from a repression state into an activation state, thereby initiating transcription.

Transcriptional Activation by CueR

After sensing monovalent metal ions, CueR binds to promoter DNA to regulate the transcription carried out by RNAP. The bacterial RNAP holoenzyme consists of a multi-subunit core enzyme ($\alpha 2 \beta \beta' \omega$) and a sigma factor (σ) responsible for promoter recognition (Murakami and Darst 2003; Borukhov and Nudler 2008). Generally, the primary σ^{70} -type RNAP recognizes promoter -10 and -35 elements, which are separated by a ~ 17 bp spacer region (Paget 2015). As described above, CueR-regulated promoters have a 19 or 20 bp spacer between -10 and -35 elements, which results in a longer distance for σ^{70} recognition in RNAP holoenzyme and a different phase angle between -35 and -10 elements (Fig. 2.5). Thus, σ^{70} -type RNAP cannot efficiently regulate such noncanonical promoters without CueR. Within the structure of CueR-promoter complexes, the target promoter is kinked and twisted by holo CueR (Philips et al. 2015). The distance between -35 and -10 elements is shortened to 56.4 Å, and the angle between these two elements is adjusted to 108°, matching the features of 17 bp spacer canonical promoter recognition by σ^{70} -type RNAP (Fig. 2.5). Similar changes have also been observed in other MerR regulator-associated complexes, including MtaN, SoxR, and BmrR (Heldwein and Brennan 2001; Newberry and Brennan 2004; Watanabe et al. 2008). Thus, CueR has been proposed to activate transcription by modifying promoter conformation, similar to other MerR regulators.

Cryo-EM structures of intact CueR-dependent transcription activation complexes have been reported (Fang et al. 2021; Shi et al. 2021). Proper binding of RNAP to the -10 and -35 elements of a noncanonical promoter has been detected and characterized in these structures, confirming the major activation mechanism through which CueR bends the 19 bp promoter spacer to realign the -35 and -10 elements for recognition by the σ^{70} -type RNAP holoenzyme. Although the DNA region recognized by CueR has a 3 bp overlap with the promoter -35 element, CueR and σ^{70} in the complex bind to the opposite faces of DNA to avoid possible steric clashes in promoter recognition (Fig. 2.7). In addition to CueR-DNA interaction, CueR also shows protein-protein interactions with RNAP. The DNA-binding domain from one CueR molecule directly interacts with the σ^{70} nonconserved region (σ^{70} -NCR) in the complex. Mutations of key residues involved in this interaction influence the transcription activated by CueR, suggesting that CueR- σ^{70} -NCR interaction plays an auxiliary role in CueR-dependent transcription activation. This auxiliary role of MerR- σ^{70} -NCR interaction is also seen in the structures of activation complexes involving EcmrR (another MerR family regulator) (Yang et al. 2021).

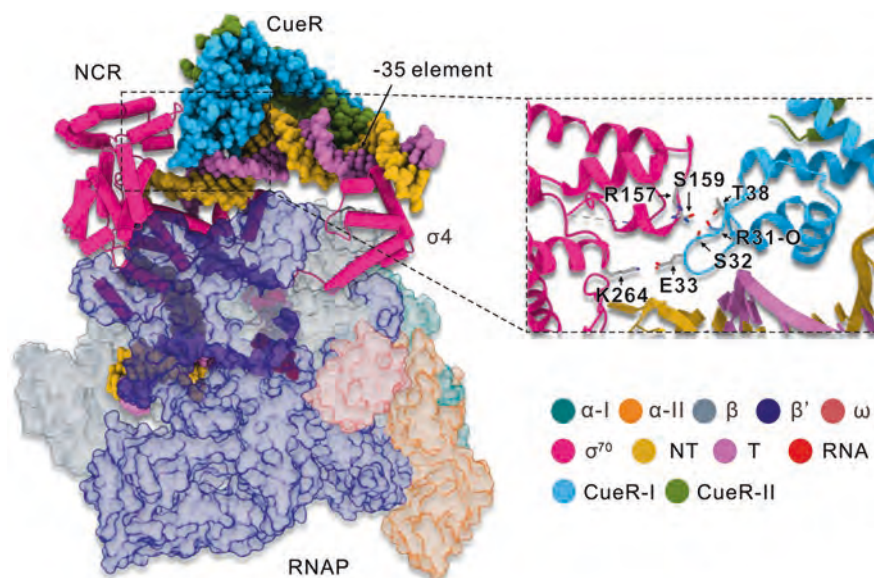


Fig. 2.7 The overall structure of *E. coli* CueR-dependent transcription activation complex with de novo-synthesized RNA transcript. The insert is the zoom-in view of the interface between σ^{70} -NCR and the CueR DBD wing. The involved residues are shown in sticks

As the DNA-binding domain and σ^{70} -NCR widely exist and are conserved in bacteria (Fig. 2.7), this mechanism might be used by other MerR family regulators.

As the CueR proteins observed in transcription activation complexes are all in a holo form, how metal ions influence the transcription process mediated by CueR remains unclear. Studies using smFRET technology revealed that RNAP may bind reversibly to a noncanonical promoter (*copA* promoter) to form two distinct non-interconverting complexes (Martell et al. 2015) (Fig. 2.8). The majority of complex, assigned as a dead-end closed-like complex, cannot initiate transcription in the presence of nucleotides, while the minority one, defined as an open promoter complex (R_{Po}), may explain the leakage transcription at this promoter. Without monovalent metal ions, the apo CueR binds the target promoter but does not bend it, therefore facilitating formation of the dead-end closed-like complex by RNAP binding promoter and then inhibiting the unbound kinetics of RNAP from this closed complex (Fig. 2.8). Once the copper ion is supplemented, holo CueR bends the target promoter to facilitate its recognition by RNAP to form an active open promoter complex. The apo-CueR-bound closed complex is unlikely to be activated by introducing copper ions. Thus, CueR functions via a “biased sampling” instead of “dynamic equilibrium shifting” mechanism in regulating transcription initiation (Fig. 2.8), which has been confirmed by electron paramagnetic resonance analyses (Sameach et al. 2019; Casto et al. 2022).

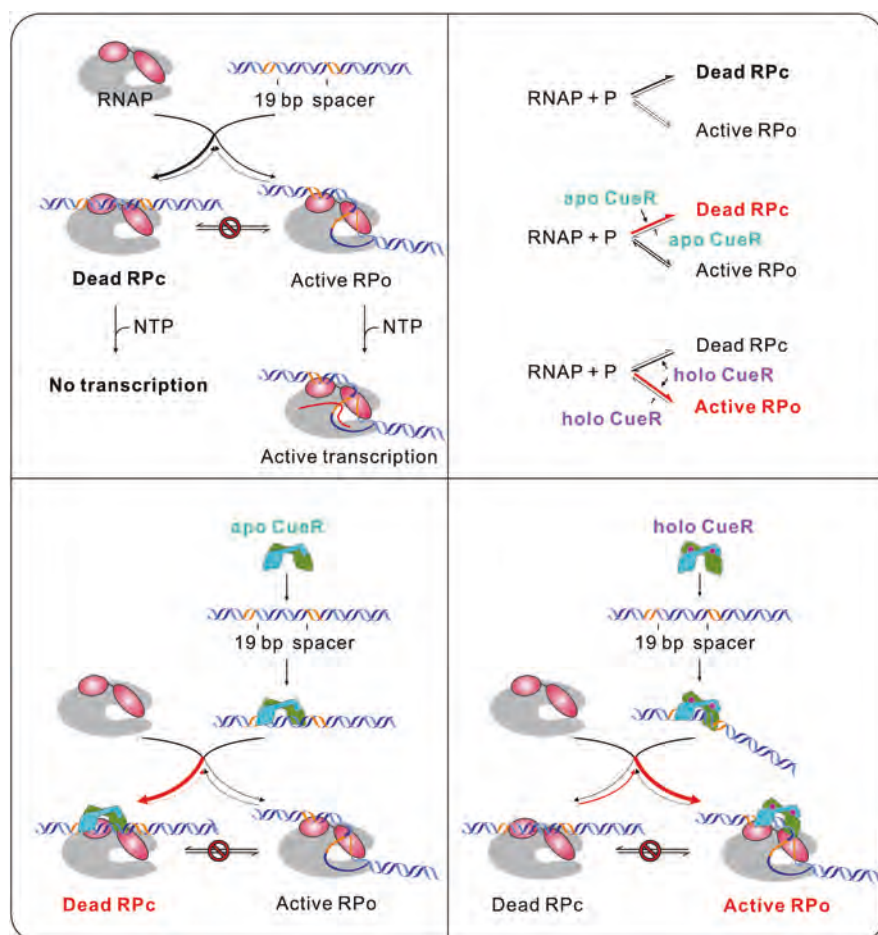


Fig. 2.8 Schematic mechanism of CueR activating transcription. RNAP binds promoter DNA to form either an active RPo or a dead RPC. The majority of complex of apo CueR-binding to promoter DNA and RNAP is the dead-end closed-like complex. Holo CueR facilitates the formation of active RNAP-promoter open complex for initiating transcription

Summary

In this chapter, we have provided a comprehensive overview of the regulatory mechanisms of the MerR family metalloregulator CueR. As a representative member of that family, CueR contains three MerR domains: an N-terminal conserved DNA-binding domain, a helix dimeric domain, and a C-terminal effector-binding domain. CueR recognizes a palindrome DNA sequence and interacts with RNAP holoenzyme on the σ^{70} -NCR. The metal-binding selectivity of CueR is a crucial step for activating the CueR protein, which is accomplished through cysteines located at its

C-terminal effector-binding domain and extended tail. In the absence of metal ions, apo-CueR binds to the promoter as a dimer, increasing the amount of dead-end closed-like complex that inhibits transcription. After sensing Cu^I (as well as other monovalent metal ions, including Ag^I and Au^I), holo CueR activates genes involved in copper homeostasis by kinking and twisting target promoters, similar to other MerR family metalloregulators. Although further investigations are still necessary for clarifying the common and specific actions between CueR and other MerR family regulators, progress has been made in characterizing how the widely existing MerR family regulators are involved in controlling gene transcription.

References

- Andoy NM, Sarkar SK, Wang Q, Panda D, Benitez JJ, Kalininskiy A, Chen P (2009) Single-molecule study of metalloregulator CueR-DNA interactions using engineered Holliday junctions. *Biophys J* 97(3):844–852. <https://doi.org/10.1016/j.bpj.2009.05.027>
- Andrei A, Ozturk Y, Khalfaoui-Hassani B, Rauch J, Marckmann D, Trasnea PI, Daldal F, Koch HG (2020) Cu homeostasis in bacteria: the ins and outs. *Membranes (Basel)* 10(9). <https://doi.org/10.3390/membranes10090242>
- Ansari AZ, Bradner JE, O'Halloran TV (1995) DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature* 374(6520):371–375. <https://doi.org/10.1038/374370a0>
- Balogh RK, Gyurcsik B, Hunyadi-Gulyas E, Schell J, Thulstrup PW, Hemmingsen L, Jancso A (2019) C-terminal cysteines of CueR act as auxiliary metal site ligands upon Hg(II) binding—a mechanism to prevent transcriptional activation by divalent metal ions? *Chemistry* 25(66):15030–15035. <https://doi.org/10.1002/chem.201902940>
- Balogh RK, Gyurcsik B, Jensen M, Thulstrup PW, Koster U, Christensen NJ, Jensen ML, Hunyadi-Gulyas E, Hemmingsen L, Jancso A (2022) Tying up a loose end: on the role of the C-terminal CCHHRAG fragment of the Metalloregulator CueR. *Chembiochem* 23(16):e202200290. <https://doi.org/10.1002/cbic.202200290>
- Borukhov S, Nudler E (2008) RNA polymerase: the vehicle of transcription. *Trends Microbiol* 16(3):126–134. <https://doi.org/10.1016/j.tim.2007.12.006>
- Brenner AJ, Harris ED (1995) A quantitative test for copper using bichinchoninic acid. *Anal Biochem* 226(1):80–84. <https://doi.org/10.1006/abio.1995.1194>
- Brown NL, Stoyanov JV, Kidd SP, Hobman JL (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* 27(2–3):145–163. [https://doi.org/10.1016/S0168-6445\(03\)00051-2](https://doi.org/10.1016/S0168-6445(03)00051-2)
- Casto J, Mandato A, Hofmann L, Yakobov I, Ghosh S, Ruthstein S, Saxena S (2022) Cu(II) -based DNA labeling identifies the structural link between transcriptional activation and termination in a metalloregulator. *Chem Sci* 13(6):1693–1697. <https://doi.org/10.1039/d1sc06563g>
- Chandrangsu P, Rensing C, Helmann JD (2017) Metal homeostasis and resistance in bacteria. *Nat Rev Microbiol* 15(6):338–350. <https://doi.org/10.1038/nrmicro.2017.15>
- Chang CC, Lin LY, Zou XW, Huang CC, Chan NL (2015) Structural basis of the mercury(II)-mediated conformational switching of the dual-function transcriptional regulator MerR. *Nucleic Acids Res* 43(15):7612–7623. <https://doi.org/10.1093/nar/gkv681>
- Changela A, Chen K, Xue Y, Holschen J, Outten CE, O'Halloran TV, Mondragon A (2003) Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* 301(5638):1383–1387. <https://doi.org/10.1126/science.1085950>
- Chen K, Yuldasheva S, Penner-Hahn JE, O'Halloran TV (2003) An atypical linear Cu(I)-S_2 center constitutes the high-affinity metal-sensing site in the CueR metalloregulatory protein. *J Am Chem Soc* 125(40):12088–12089. <https://doi.org/10.1021/ja036070y>

- Fang C, Philips SJ, Wu X, Chen K, Shi J, Shen L, Xu J, Feng Y, O'Halloran TV, Zhang Y (2021) CueR activates transcription through a DNA distortion mechanism. *Nat Chem Biol* 17(1):57–64. <https://doi.org/10.1038/s41589-020-00653-x>
- Fang C, Zhang Y (2022) Bacterial MerR family transcription regulators: activation by distortion. *Acta Biochim Biophys Sin Shanghai* 54(1):25–36. <https://doi.org/10.3724/abbs.2021003>
- Finney LA, O'Halloran TV (2003) Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* 300(5621):931–936. <https://doi.org/10.1126/science.1085049>
- Godsey MH, Baranova NN, Neyfakh AA, Brennan RG (2001) Crystal structure of MtaN, a global multidrug transporter gene activator. *J Biol Chem* 276(50):47178–47184. <https://doi.org/10.1074/jbc.M105819200>
- Goldstein S, Czapski G (1986) The role and mechanism of metal ions and their complexes in enhancing damage in biological systems or in protecting these systems from the toxicity of O₂. *J Free Radic Biol Med* 2(1):3–11. [https://doi.org/10.1016/0748-5514\(86\)90117-0](https://doi.org/10.1016/0748-5514(86)90117-0)
- Guzikevich-Guerstein G, Shakked Z (1996) A novel form of the DNA double helix imposed on the TATA-box by the TATA-binding protein. *Nat Struct Biol* 3(1):32–37. <https://doi.org/10.1038/nsb0196-32>
- Heldwein EE, Brennan RG (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* 409(6818):378–382. <https://doi.org/10.1038/35053138>
- Hobman JL, Wilkie J, Brown NL (2005) A design for life: prokaryotic metal-binding MerR family regulators. *Biometals* 18(4):429–436. <https://doi.org/10.1007/s10534-005-3717-7>
- Joshi CP, Panda D, Martell DJ, Andoy NM, Chen TY, Gaballa A, Helmann JD, Chen P (2012) Direct substitution and assisted dissociation pathways for turning off transcription by a MerR-family metalloregulator. *Proc Natl Acad Sci USA* 109(38):15121–15126. <https://doi.org/10.1073/pnas.1208508109>
- Ma Z, Jacobsen FE, Giedroc DP (2009) Coordination chemistry of bacterial metal transport and sensing. *Chem Rev* 109(10):4644–4681. <https://doi.org/10.1021/cr900077w>
- Martell DJ, Joshi CP, Gaballa A, Santiago AG, Chen TY, Jung W, Helmann JD, Chen P (2015) Metalloregulator CueR biases RNA polymerase's kinetic sampling of dead-end or open complex to repress or activate transcription. *Proc Natl Acad Sci USA* 112(44):13467–13472. <https://doi.org/10.1073/pnas.1515231112>
- Murakami KS, Darst SA (2003) Bacterial RNA polymerases: the whole story. *Curr Opin Struct Biol* 13(1):31–39. [https://doi.org/10.1016/s0959-440x\(02\)00005-2](https://doi.org/10.1016/s0959-440x(02)00005-2)
- Newberry KJ, Brennan RG (2004) The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. *J Biol Chem* 279(19):20356–20362. <https://doi.org/10.1074/jbc.M400960200>
- O'Halloran TV (1993) Transition metals in control of gene expression. *Science* 261(5122):715–725. <https://doi.org/10.1126/science.8342038>
- Outten FW, Outten CE, Hale J, O'Halloran TV (2000) Transcriptional activation of an *Escherichia coli* copper efflux regulon by the chromosomal MerR homologue, cueR. *J Biol Chem* 275(40):31024–31029. <https://doi.org/10.1074/jbc.M006508200>
- Page MS (2015) Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 5(3):1245–1265. <https://doi.org/10.3390/biom5031245>
- Perez Audero ME, Podoroska BM, Ibanez MM, Cauerhff A, Checa SK, Soncini FC (2010) Target transcription binding sites differentiate two groups of MerR-monovalent metal ion sensors. *Mol Microbiol* 78(4):853–865. <https://doi.org/10.1111/j.1365-2958.2010.07370.x>
- Petersen C, Moller LB (2000) Control of copper homeostasis in *Escherichia coli* by a P-type ATPase, CopA, and a MerR-like transcriptional activator, CopR. *Gene* 261(2):289–298. [https://doi.org/10.1016/s0378-1119\(00\)00509-6](https://doi.org/10.1016/s0378-1119(00)00509-6)
- Philips SJ, Canalizo-Hernandez M, Yildirim I, Schatz GC, Mondragon A, O'Halloran TV (2015) Transcription. Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Science* 349(6250):877–881. <https://doi.org/10.1126/science.aaa9809>

- Price EE, Boyd JM (2020) Genetic regulation of metal ion homeostasis in *Staphylococcus aureus*. *Trends Microbiol* 28(10):821–831. <https://doi.org/10.1016/j.tim.2020.04.004>
- Rao L, Cui Q, Xu X (2010) Electronic properties and desolvation penalties of metal ions plus protein electrostatics dictate the metal binding affinity and selectivity in the copper efflux regulator. *J Am Chem Soc* 132(51):18092–18102. <https://doi.org/10.1021/ja103742k>
- Rapisarda VA, Montelongo LR, Farias RN, Massa EM (1999) Characterization of an NADH-linked cupric reductase activity from the *Escherichia coli* respiratory chain. *Arch Biochem Biophys* 370(2):143–150. <https://doi.org/10.1006/abbi.1999.1398>
- Rensing C, Fan B, Sharma R, Mitra B, Rosen BP (2000) CopA: an *Escherichia coli* Cu(I)-translocating P-type ATPase. *Proc Natl Acad Sci USA* 97(2):652–656. <https://doi.org/10.1073/pnas.97.2.652>
- Robinson NJ, Winge DR (2010) Copper metallochaperones. *Annu Rev Biochem* 79:537–562. <https://doi.org/10.1146/annurev-biochem-030409-143539>
- Sameach H, Ghosh S, Gevorkyan-Airapetov L, Saxena S, Ruthstein S (2019) EPR spectroscopy detects various active state conformations of the transcriptional regulator CueR. *Angew Chem Int Ed Engl* 58(10):3053–3056. <https://doi.org/10.1002/anie.201810656>
- Sameach H, Narunsky A, Azoulay-Ginsburg S, Gevorkyan-Airapetov L, Zehavi Y, Moskovitz Y, Juven-Gershon T, Ben-Tal N, Ruthstein S (2017) Structural and dynamics characterization of the MerR family Metalloregulator CueR in its repression and activation states. *Structure* 25(7):988–996 e983. <https://doi.org/10.1016/j.str.2017.05.004>
- Shi W, Zhang B, Jiang Y, Liu C, Zhou W, Chen M, Yang Y, Hu Y, Liu B (2021) Structural basis of copper-efflux-regulator-dependent transcription activation. *iScience* 24(5):102449. <https://doi.org/10.1016/j.isci.2021.102449>
- Stoyanov JV, Brown NL (2003) The *Escherichia coli* copper-responsive copA promoter is activated by gold. *J Biol Chem* 278(3):1407–1410. <https://doi.org/10.1074/jbc.C200580200>
- Stoyanov JV, Hobman JL, Brown NL (2001) CueR (YbbI) of *Escherichia coli* is a MerR family regulator controlling expression of the copper exporter CopA. *Mol Microbiol* 39(2):502–511. <https://doi.org/10.1046/j.1365-2958.2001.02264.x>
- Summers AO (2009) Damage control: regulating defenses against toxic metals and metalloids. *Curr Opin Microbiol* 12(2):138–144. <https://doi.org/10.1016/j.mib.2009.02.003>
- Thaden JT, Lory S, Gardner TS (2010) Quorum-sensing regulation of a copper toxicity system in *Pseudomonas aeruginosa*. *J Bacteriol* 192(10):2557–2568. <https://doi.org/10.1128/JB.01528-09>
- Waldron KJ, Rutherford JC, Ford D, Robinson NJ (2009) Metalloproteins and metal sensing. *Nature* 460(7257):823–830. <https://doi.org/10.1038/nature08300>
- Watanabe S, Kita A, Kobayashi K, Miki K (2008) Crystal structure of the [2Fe-2S] oxidative-stress sensor SoxR bound to DNA. *Proc Natl Acad Sci USA* 105(11):4121–4126. <https://doi.org/10.1073/pnas.0709188105>
- Yakovov I, Mandato A, Hofmann L, Singewald K, Shenberger Y, Gevorkyan-Airapetov L, Saxena S, Ruthstein S (2022) Allosteric-driven changes in dynamics regulate the activation of bacterial copper transcription factor. *Protein Sci* 31(5):e4309. <https://doi.org/10.1002/pro.4309>
- Yamamoto K, Ishihama A (2005) Transcriptional response of *Escherichia coli* to external copper. *Mol Microbiol* 56(1):215–227. <https://doi.org/10.1111/j.1365-2958.2005.04532.x>
- Yang Y, Liu C, Zhou W, Shi W, Chen M, Zhang B, Schatz DG, Hu Y, Liu B (2021) Structural visualization of transcription activated by a multidrug-sensing MerR family regulator. *Nat Commun* 12(1):2702. <https://doi.org/10.1038/s41467-021-22990-8>

Chapter 3

The Structural Biology of Catalase Evolution



María Belén López, María Belén Oterino, and Javier M. González 

Abstract Catalases are essential enzymes for removal of hydrogen peroxide, enabling aerobic and anaerobic metabolism in an oxygenated atmosphere. Monofunctional heme catalases, catalase-peroxidases, and manganese catalases, evolved independently more than two billion years ago, constituting a classic example of convergent evolution. Herein, the diversity of catalase sequences is analyzed through sequence similarity networks, providing the context for sequence distribution of major catalase families, and showing that many divergent catalase families remain to be experimentally studied.

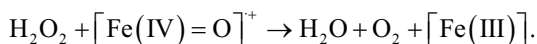
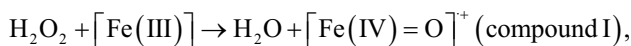
Keywords Catalase · Hydrogen peroxide · Enzyme evolution · Sequence similarity networks

Introduction

With the advent of oxygenic photosynthesis and aerobic metabolism about 2.3 billion years ago, reactive oxygen species (ROS) became an issue for all forms of life (Sánchez-Baracaldo and Cardona 2020). The umbrella term oxidative stress is often used to describe the variety of detrimental effects ROS have on biological molecules (Halliwell and Gutteridge 2015; Sies et al. 2017). Such strong selective pressure led to the evolution of efficient mechanisms for ROS removal and oxygen scavenging, including the production of catalases, metalloenzymes that dismutate the toxic hydrogen peroxide into water and molecular oxygen. Catalases have been extensively studied for over a century due to the ease of catalase activity assay, with the readily available H_2O_2 substrate and straightforward measurement of O_2 production

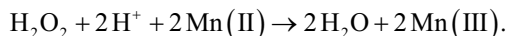
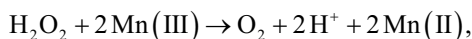
M. B. López · M. B. Oterino · J. M. González (✉)
Instituto de Bionanotecnología del NOA (INBIONATEC-CONICET),
Universidad Nacional de Santiago del Estero (UNSE), Santiago del Estero, Argentina
e-mail: javierng@conicet.gov.ar

(Chelikani et al. 2004; Díaz et al. 2012). Despite the simplicity of the H_2O_2 dismutation reaction, $2\text{H}_2\text{O}_2 \rightarrow 2\text{H}_2\text{O} + \text{O}_2$, its mechanism proved to be rather complex, and many aspects remain unknown even today. Most catalases use H_2O_2 to oxidize Fe(III) heme *b* or *d* into an Fe(IV) oxo-ferryl radical (compound I) and then reduce it back to Fe(III) heme using a second molecule of H_2O_2 as electron donor, thereby producing molecular oxygen (Alfonso-Prieto et al. 2009, 2012):



Compound I is a highly reactive species that may oxidize alternative electron donors depending on the enzyme active site environment. Although the name *catalase* was naively coined by biochemists in the early days of enzymology, by modern standards, catalases are hydroperoxidases. Indeed, the abovementioned half reactions are respectively known as catalytic and peroxidatic reactions. While catalases only allow a H_2O_2 molecule to reduce compound I, peroxidases use varied organic substrates as electron donors (Zámocký et al. 2015). Bifunctional catalase-peroxidases apparently evolved from plant-fungal peroxidases (Chelikani et al. 2004) and do not share significant sequence homology with monofunctional heme catalases. The peroxidatic reaction of catalase-peroxidases may occur through two one-electron steps or one two-electron step, with H_2O_2 or varied electron donors. In a way, these enzymes evolved to harness the power of H_2O_2 to oxidize a variety of organic and inorganic compounds.

Another type of catalase uses a radically different mechanism to catalyze the dismutation of hydrogen peroxide, using a dinuclear manganese center instead of heme iron. Such non-heme Mn-dependent catalases are typically slower than monofunctional heme catalases, with turnover rates in the order of $\sim 10^5 \text{ s}^{-1}$ versus $\sim 10^7 \text{ s}^{-1}$ of their heme-dependent counterparts (Chance et al. 1952; Shank et al. 1994). In Mn catalases, hydrogen peroxide is first oxidized to molecular oxygen, while the enzyme extracts two electrons and two protons, which are later used to reduce a second hydrogen peroxide molecule into two water molecules (Whittaker 2012):



Traditional Classification of Catalases

A number of studies have reviewed the structure, evolution, and diversity of the three classes of catalases (Kirkman and Gaetani 1984; Chelikani et al. 2004; Zámocký 2004; Smulevich et al. 2006; Passardi et al. 2007; Whittaker 2012; Zámocký et al. 2012). Monofunctional heme catalases share a characteristic fold for

heme binding and tetramerization through an intricate network of noncovalent interactions between subunits (Fig. 3.1a). They often exhibit an intense Soret band at 406 nm that parallels the protein absorption intensity at 280 nm. The so-called catalase fold is about 460 amino acid residues, with a heme-binding structure conserved throughout all kingdoms of life, highlighting its central role for coping with oxygen and ROS. The catalase fold comprises a β -barrel domain and an α -helical domain connected by a wrapping segment. The ferric protoporphyrin IX (heme *b*) present in each subunit is located on one side of the β -barrel and defines two sides: the proximal side, harboring an essential tyrosine residue from an α -helix in the wrapping segment, covalently coordinated to the heme iron, and the distal side, where the main channel for substrate entrance is located, lined with residues evolved to prevent any side reactions during H_2O_2 dismutation that could potentially generate deleterious ROS.

Early phylogenetic analyses have grouped monofunctional catalases into three clades (Klotz et al. 1997) (see Table 3.1 and Fig. 3.2): clade I enzymes comprise about 500 residues per monomer and are found mainly in plants and some bacteria; clade II enzymes are larger, due to an additional flavodoxin-like ~250-residue domain, and often belong to bacteria and fungi; and catalases from clade III are smaller and more widespread among eubacteria, fungi, cyanobacteria, and higher eukaryotes. Curiously, all eukaryotic and most bacterial clade III catalases bind a molecule of NAD(P)H; however, it is not required for catalysis given its absence in enzymes from clades 1 and 2 (Díaz et al. 2012). The presence of this NAD(P)H molecule has been linked to protection of catalases from oxidative damage by hydrogen peroxide and derivative ROS (Fita and Rossmann 1985; Kirkman et al. 1999).

At present, about 140 structures of monofunctional catalases have been experimentally determined by X-ray crystallography and deposited in the Protein Data Bank, corresponding to 4 enzymes from clade I, 9 from clade II, and 13 from clade III (Table 3.1). So far, most determined structures correspond to microbes (bacteria and fungi), there are no structures for monofunctional catalases from plants nor archaea, and the only experimental structures of catalases from higher eukaryotes are those from beef liver and human erythrocytes. From the structural viewpoint, each clade shares distinctive features that parallel the phylogenetic classification; however, the catalase folds of clade III enzymes from *S. cerevisiae* (PDB 1a4e), *P. pastoris* (PDB 2xq1), and *M. luteus* (PDB 1gwe) resemble the catalase fold of clade II enzymes (Fig. 3.2).

Catalase-peroxidases (often named KatGs) exhibit the typical fold of plant and fungal peroxidases, a homodimeric structure with two polypeptides of about 750 residues that likely arose from a duplication of an α -helix-rich heme-binding domain (Welinder 1991) (Fig. 3.1b). Only the N-terminal domain retains the ability to bind a heme *b* molecule, whose iron atom is covalently coordinated to a conserved His residue in the proximal side. A unique adduct of Met-Tyr-Trp residues bonded together is found in the distal side, along with an extensive network of solvent molecules occupying the void of a narrow channel for substrate entrance. The mechanism of reaction of catalase-peroxidases is naturally more complex than that of monofunctional catalases. Compound I in catalase-peroxidases is sequentially

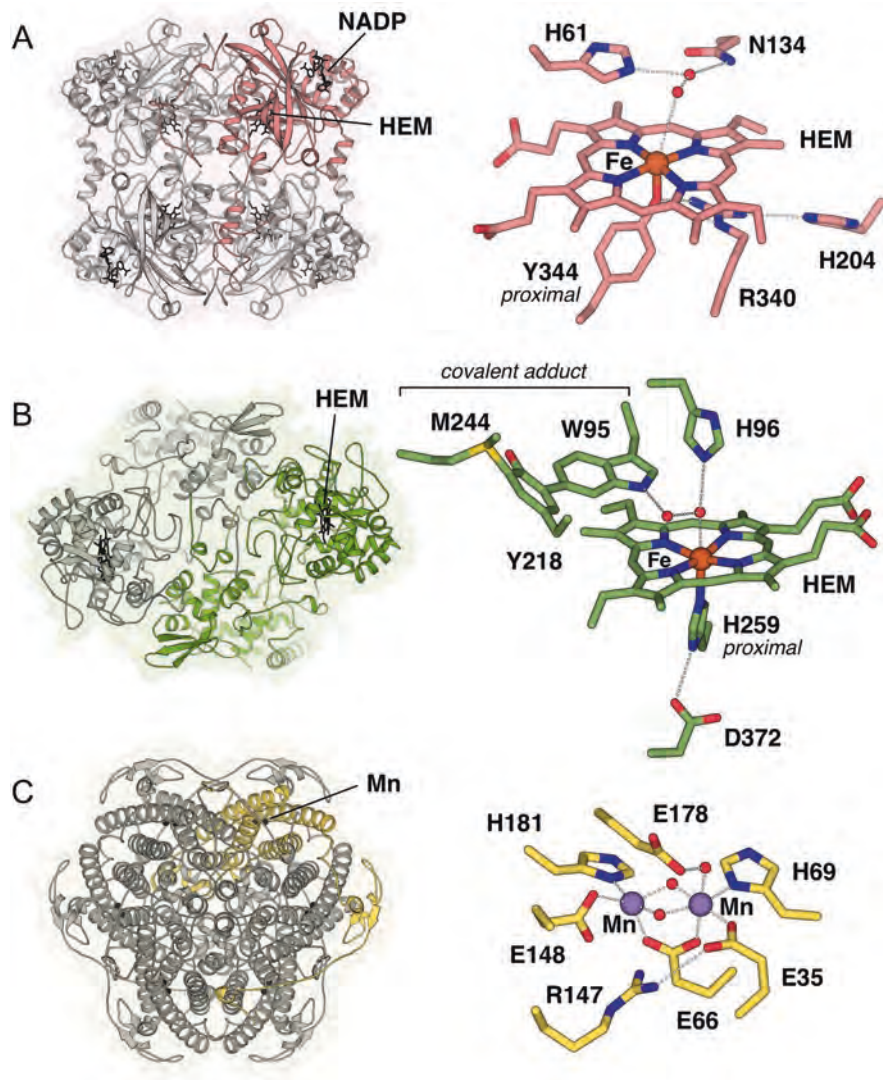


Fig. 3.1 Representative catalase structures obtained by X-ray crystallography, showing their quaternary structure and active site details. **(a)** Monofunctional catalase KatE1 (clade III, tetrameric) from *Acinetobacter* sp. (PDB 6pt7) with bound heme *b* (HEM) and nicotinamide adenine dinucleotide phosphate (NADP), **(b)** catalase-peroxidase HmCPx (dimeric) from archaeon *Haloarcula marismortui* (PDB 1ltk), and **(c)** Mn catalase LPC (hexameric) from *Lactiplantibacillus (Lactobacillus) plantarum* (PDB 1jku). Selected colored subunits highlight the intricate interaction mode between monomers, particularly in monofunctional and Mn catalases. Dotted lines indicate important noncovalent interactions in the first and second shells of each metal ion. Red spheres indicate solvent molecules that point the likely H_2O_2 binding sites

Table 3.1 Catalases with available experimentally determined three-dimensional structures

Enzyme	Clade	Name	PDB	UniProt	Organism	Phylum
CAT	1	CatF	1M7S	P46206	<i>Pseudomonas syringae</i>	<i>Pseudomonadota</i>
	1		4CAB	Q59337	<i>Deinococcus radiodurans</i>	<i>Deinococcota</i>
	1		4QOL	–	<i>Bacillus pumilus</i>	<i>Bacillota</i>
	1		2J2M	A2A136	<i>Exiguobacterium oxidotolerans</i>	<i>Bacillota</i>
	2		4CAT	–	<i>Penicillium vitale</i>	<i>Ascomycota</i>
	2		3E4W	Q73WB6	<i>Mycobacterium paratuberculosis</i>	<i>Actinomycetota</i>
	2	PVC	2IUF	D9N167	<i>Penicillium janthinellum</i>	<i>Ascomycota</i>
	2	CATPO	4B7A	M4GGR5	<i>Mycothermus thermophilus</i>	<i>Ascomycota</i>
	2	CATPO	4AUL	M4GGR6	<i>Mycothermus thermophilus</i>	<i>Ascomycota</i>
	2	Cat3	4AJ9	Q9C169	<i>Neurospora crassa</i>	<i>Ascomycota</i>
	2	Cat1	1SY7	Q9C168	<i>Neurospora crassa</i>	<i>Ascomycota</i>
	2	KatE, HPII	3P9P	P21179	<i>Escherichia coli</i>	<i>Pseudomonadota</i>
	2	eKatE	6LFK	A0A618WFM0	<i>Escherichia coli</i> M8	<i>Pseudomonadota</i>
	3	MLC	1GWE	P29422	<i>Micrococcus luteus</i>	<i>Actinomycetota</i>
	3	CatA	1A4E	P15202	<i>Saccharomyces cerevisiae</i>	<i>Ascomycota</i>
	3		2XQ1	P30263	<i>Ogataea angusta</i>	<i>Ascomycota</i>
	3	BLC	3RGP	P00432	<i>Bos taurus</i>	<i>Chordata</i>
	3	HEC	1DGF	P04040	<i>Homo sapiens</i>	<i>Chordata</i>
	3		1SI8	Q834P5	<i>Enterococcus faecalis</i>	<i>Bacillota</i>
	3	HPC	1QWL	P77872	<i>Helicobacter pylori</i>	<i>Campylobacterota</i>
	3		4E37	O52762	<i>Pseudomonas aeruginosa</i>	<i>Pseudomonadota</i>
	3	KatA	1E93	P42321	<i>Proteus mirabilis</i>	<i>Pseudomonadota</i>
	3		2ISA	Q3LSM1	<i>Alivibrio salmonicida</i>	<i>Bacteroidota</i>
	3	KatE1	6PT7	A0A031LXI5	<i>Acinetobacter</i> sp. Ver3	<i>Pseudomonadota</i>

(continued)

Table 3.1 (continued)

Enzyme	Clade	Name	PDB	UniProt	Organism	Phylum
CAT-PER	3		6RJN	C1PHG1	<i>Komagataella pastoris</i>	Ascomycota
	3		6RJR	Q6CR58	<i>Kluyveromyces lactis</i>	Ascomycota
		HmCPx	IITK	O59651	<i>Haloarcula marismortui</i>	Euryarchaeota
		KatG	ISJ2	P9WIE5	<i>Mycobacterium tuberculosis</i>	Actinomycetota
			IUB2	Q31MN3	<i>Synechococcus elongatus</i>	Cyanobacteriota
		MagKatG2	3UT2	A4QUT2	<i>Pyricularia oryzae</i>	Ascomycota
		BpKatG	5 L02	Q3JNW6	<i>Burkholderia pseudomallei</i>	Pseudomonadota
			5WHQ	Q8X182	<i>Neurospora crassa</i>	Ascomycota
		KatG	7JZ6	A0A037YMI1	<i>Escherichia coli</i>	Pseudomonadota
		LPC	IJKU	P60355	<i>Lactiplantibacillus plantarum</i>	Bacillota
MN-CAT		TTC	2V8U	Q84DB4	<i>Thermus thermophilus</i>	Thermoproteota
		KatB	4R42	Q8YSJ5	<i>Anabaena</i> sp. (<i>Nostoc</i> sp.)	Cyanobacteriota

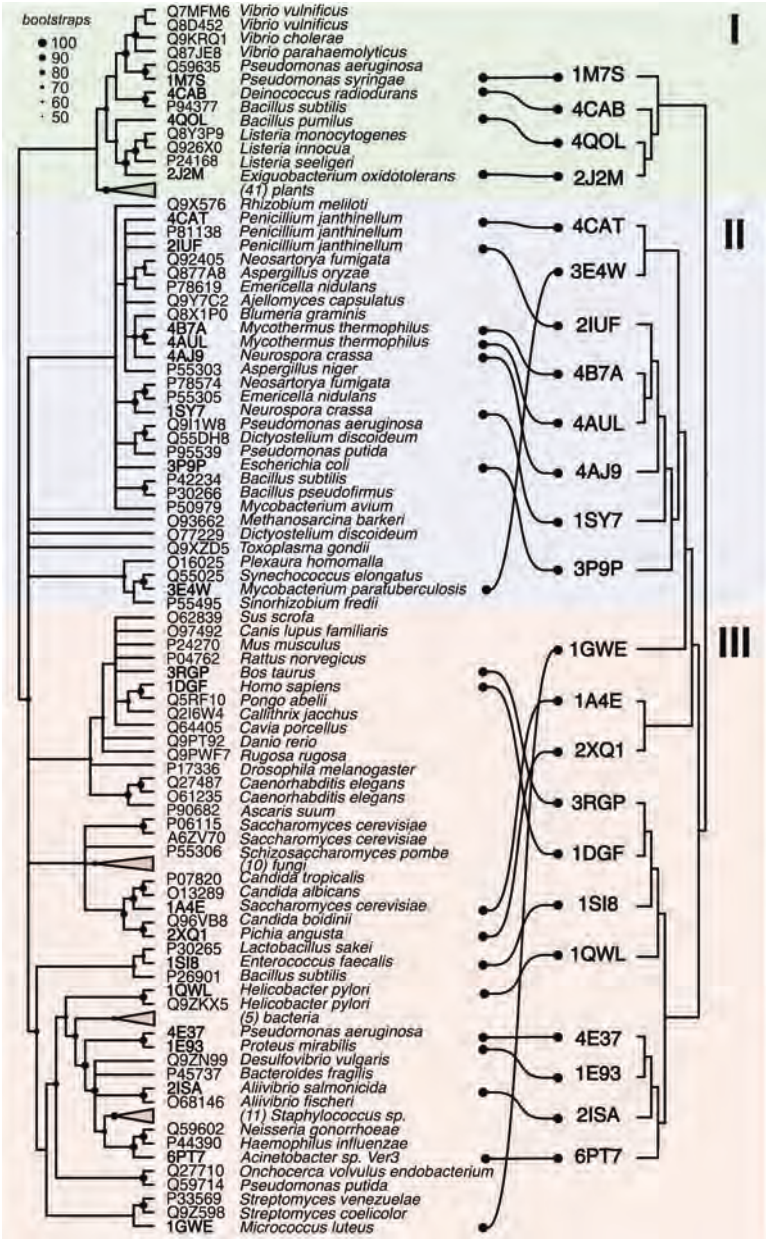


Fig. 3.2 Comparison of monofunctional catalase sequences. Tanglegram matching a bootstrapped maximum-likelihood cladogram (152 sequences, *left*) with a structural phenogram (23 structures, *right*) calculated from an all-to-all distance matrix of superposed crystallographic structures. Each polypeptide is indicated by UniProt or PDB identifiers. Members of clades I, II, and III were assigned considering both tree topologies in the context of literature information. Adapted from Sartorio et al. (2021)

reduced to hydroxy-ferryl protoporphyrin [Fe(IV)–OH] (compound II), an unstable superoxo-ferric protoporphyrin [Fe(III)–O₂^{•−}] (compound III) and back to the ferric heme ground state; the details of this peroxidatic process depend largely on the electron donor chemical features (Vlasits et al. 2010). Seven experimental catalase-peroxidase structures have been deposited in the Protein Data Bank, spanning enzymes from archaea, cyanobacteria, fungi, actinomycetes, and pseudomonads (Table 3.1).

Mn catalases, also known as KatN catalases, belong to the ferritin superfamily, sharing a four-helix bundle motif comprising two antiparallel α -helical hairpins, with a dinuclear μ -oxo bridged Mn center in the middle of the bundle, coordinated to His and Glu residues. The C-terminal segment adopts an extended structure that symmetrically interacts with another monomer, thereby stabilizing a pseudodimer that gives rise to a homohexameric quaternary structure with three pseudodimers (Fig. 3.1c). Putative Mn catalases have been found in a wide range of bacteria and archaea, but not in eukaryotes; exceptions to this rule may result from horizontal gene transfer events. Some obligate anaerobes like catalase-negative *Clostridium* sp. and *Anaerofustis* sp. and catalase-positive *Desulfobacterium* sp. express Mn catalases as a defense mechanism to survive with small concentrations of H₂O₂ under microaerophilic conditions. Mn catalases in *Lactobacillus* may come from ancient *Firmicutes* (*Bacillota*). Some pathogenic bacteria like *Salmonella enterica* sv. *typhimurium*, *Escherichia coli* O157:H7, and *Pseudomonas aeruginosa* express Mn catalases as a defense mechanism against oxidative stress and iron deficiency conditions in phagolysosomes, part of the antimicrobial weaponry in macrophages (Slauch 2011).

Updated Classification Scheme for Catalases

The InterPro database (Blum et al. 2021) contains specific entry identifiers for each class of catalases, namely, IPR002226 for monofunctional catalases (CAT), IPR000763 for catalase-peroxidases (CAT-PER), and IPR007760 for manganese catalases (MN-CAT). For InterPro version 95 (UniProt version 2023_03), the three sets total 71,713 sequences that can be mapped to 28,372 clusters in the UniRef90 database. In the following sections, this sequence set will be analyzed using sequence similarity networks (SSN) (Table 3.2). SSNs constitute a convenient alternative to the classical classification of sequences based on multiple sequence alignments and cladograms, particularly suitable for large sets of divergent sequences (Atkinson et al. 2009). SSN graphs consist of sets of nodes representing one or more sequences, connected by edges indicative of a specified sequence similarity threshold, and have been applied to describe a number of protein families (Atkinson and Babbitt 2009; Baier and Tokuriki 2014; Davidson et al. 2018; Copp et al. 2019; Malik and Kim 2019; Tararina and Allen 2020; Shi et al. 2020; Sartorio et al. 2021). Roughly, for every Mn catalase sequence known, there are two catalase-peroxidases and three monofunctional catalases. This 1:2:3 ratio turns to 1:1:2 when considering

Table 3.2 Statistics of each catalase class sequence similarity network

	CAT	CAT-PER	MN-CAT
InterPro entry	IPR002226	IPR000763	IPR007760
UniProt sequences	35,079	24,495	12,139
UniRef90 clusters	11,348	12,185	4839
RedoxiBase sequences	465	546	101
SSN rep nodes (NC >5)	2624	2728	1953
SSN edges (60% ID)	124,256	1,183,475	114,238
SSN edges (70% ID)	–	63,770	–

the UniRef90 database, where sequences sharing 90% or higher sequence similarity are grouped into a single identifier. Furthermore, this ratio is modified to 2:3:3 when considering only representative nodes in each SSN with an average number of neighbors of 5 or more, as defined by a neighborhood connectivity (NC) of 5 or higher. By considering sequences of NC > 5 only, most outliers like fragments and rare sequences that are not included in any clusters are discarded. Thus, even though there are more sequences of monofunctional catalases than catalase-peroxidases, their sequence diversity is comparable, whereas Mn catalases are somewhat more conserved, less diverse enzymes. Surprisingly, this trend is also followed by the subset of sequences from the RedoxiBase (Savelli et al. 2019). The set of 1112 sequences considered for network analysis is represented in 35 nodes, comprising 13 monofunctional catalases, 14 catalase-peroxidases, and 8 Mn catalases

Table 3.3 List of catalases from RedoxiBase included in SSN analysis

	Name	Organism	Phylum
CAT	AnKat05	<i>Aspergillus niger</i>	<i>Ascomycota</i>
	EcamKat01	<i>Eucalyptus camaldulensis</i>	<i>Eudicotyledons</i>
	EglKat[P]07	<i>Eucalyptus globulus</i>	<i>Eudicotyledons</i>
	EguKat[P]05	<i>Eucalyptus gunnii</i>	<i>Eudicotyledons</i>
	EguKat13	<i>Eucalyptus gunnii</i>	<i>Eudicotyledons</i>
	HaKat08	<i>Helianthus annuus (sunflower)</i>	<i>Eudicotyledons</i>
	KflKat03	<i>Klebsormidium flaccidum</i>	<i>Streptophyta algae</i>
	LeKat[P]03	<i>Lycopersicon esculentum (tomato)</i>	<i>Eudicotyledons</i>
	LeKat04	<i>Lycopersicon esculentum (tomato)</i>	<i>Eudicotyledons</i>
	PinvKat01	<i>Paxillus involutus</i>	<i>Basidiomycota</i>
	PinvKat04	<i>Paxillus involutus</i>	<i>Basidiomycota</i>
	SbreKat03	<i>Suillus brevipes</i>	<i>Basidiomycota</i>
CAT-PER	ThasKat02	<i>Tarenaya hassleriana</i>	<i>Eudicotyledons</i>
	AwenCP01	<i>Aspergillus wentii</i>	<i>Ascomycota</i>
	CfoCP02_PJ7	<i>Colletotrichum fioriniae (Glomerella acutata)</i>	<i>Ascomycota</i>
	CfulCP01	<i>Cladosporium fulvum</i>	<i>Ascomycota</i>
	CgloCP01_Narage5	<i>Colletotrichum gloeosporioides (Glomerella cingulata)</i>	<i>Ascomycota</i>

(continued)

Table 3.3 (continued)

	Name	Organism	Phylum
	CgloCP01_ST23	<i>Colletotrichum gloeosporioides</i> (<i>Glomerella cingulata</i>)	Ascomycota
	CheCP_C5	<i>Cochliobolus heterostrophus</i>	Ascomycota
	COsaCP	<i>Cochliobolus sativus</i>	Ascomycota
	CtoCP	<i>Chrysochromulina tobin</i> (sp. CCMP291)	Prymniales
	Fcy1CP02	<i>Fragilariopsis cylindrus</i>	Bacillariophyta (diatoms)
	LpaluCP01	<i>Lepidopterella palustris</i>	Ascomycota
	MpopCP	<i>Mycosphaerella populorum</i> (<i>Septoria musiva</i>)	Ascomycota
	OCgCP02	<i>Ochlochaete</i> sp.	Chlorophyta (green algae)
	TterCP01	<i>Thielavia terrestris</i> (<i>Acremonium alabamense</i>)	Ascomycota
	XhypCP01	<i>Xylaria hypoxylon</i>	Ascomycota
MN-CAT	ACxMnCat01	<i>Achromobacter xylosoxidans</i>	Betaproteobacteria
	BcenMnCat	<i>Burkholderia cenocepacia</i>	Betaproteobacteria
	HlaMnCat	<i>Halorubrum lacusprofundi</i>	Euryarchaeota
	KpMnCat01	<i>Klebsiella pneumoniae</i>	Gammaproteobacteria
	MsiMnCat03	<i>Methylocella silvestris</i>	Alphaproteobacteria
	PlmaMnCat	<i>Planctomyces maris</i>	Other bacteria
	SmedMnCat	<i>Sinorhizobium medicae</i>	Alphaproteobacteria
	SuMnCat02	<i>Solibacter usitatus</i> (<i>Candidatus Solibacter usitatus</i>)	Other bacteria

(Table 3.3), roughly the same 3:3:2 ratio as the general SSN. The distribution of NC values throughout SSNs clearly shows major groups or clusters of sequences of monofunctional and Mn catalases sharing at least 60% sequence similarity (Fig. 3.3a and b); however, catalase-peroxidases seem to comprise a major set of sequences sharing more than 70% sequence similarity (Fig. 3.3c).

The main group of monofunctional catalases (i.e., those with the larger NC values) comprises enzymes from phyla *Pseudomonadota* (mostly α -proteobacteria, with some β - and γ -proteobacteria), *Actinomycetota* (*Actinomycetes*, with some *Acidimicrobiia*), and *Planctomycetota* (*Planctomycetia*) of about 700–800 residues. This is also visible in the clade distribution and size distribution histogram for these enzymes (Fig. 3.3g and h), showing that most monofunctional catalases are prokaryotic, large polypeptides, from clade II. The second largest group includes diverse enzymes from *Actinomycetota* (*Actinomycetes*) of 700–800 residues, likely from clade II, although without any members experimentally characterized. Another cluster includes bacterial enzymes from *Bacillota* (*Bacilli*, *Clostridia*), *Pseudomonadota* (α -, β -, and γ -proteobacteria), *Myxococcota* (*Myxococcia*), *Deinococcota* (*Deinococci*), *Cyanobacteriota* (*Cyanophyceae*), *Chloroflexota* (*Chloroflexia*), and *Actinomycetes* and virtually all archaeal monofunctional catalases from *Euryarchaeota* (*Stenosarchaea* group, *Methanobacteria*). Next, two clusters of mostly yeast and fungal enzymes from *Ascomycota* (*Saccharomyceta*)

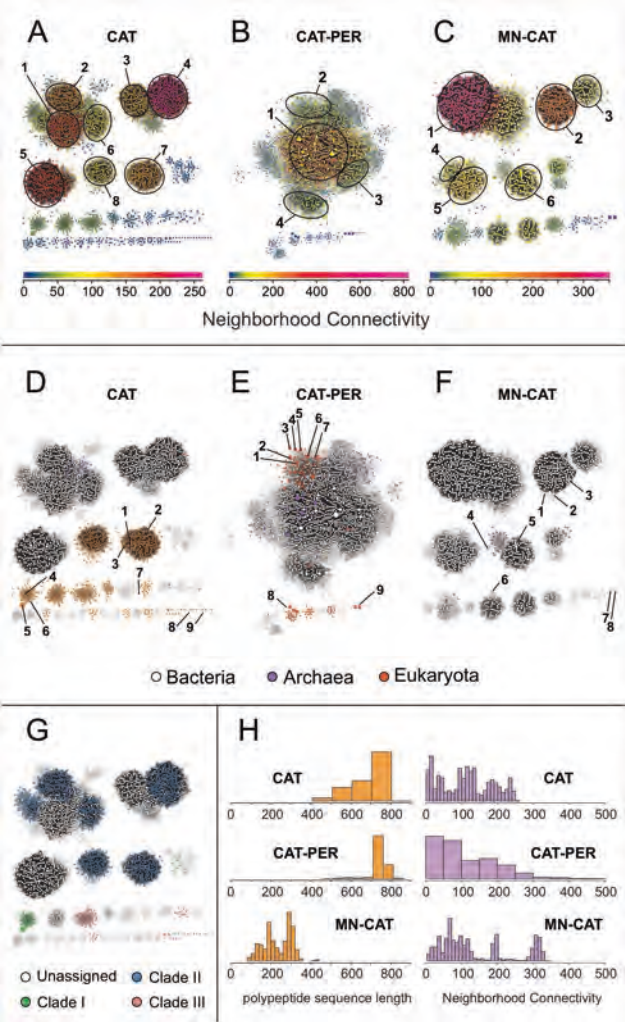
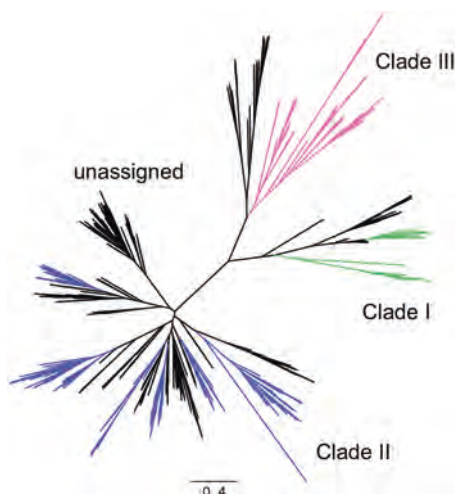


Fig. 3.3 Sequence similarity networks (SSN) of monofunctional catalases (CAT, IPR002226), catalase-peroxidases (CAT-PER, IPR002226), and Mn catalases (MN-CAT, IPR007760). (a, b, c) Phyla are indicated for major protein families, as revealed through neighborhood connectivity (NC) analysis (González 2021). (d, e, f) The three kingdoms of life mapped on each SSN, with selected catalases from RedoxiBase (Savelli et al. 2019) (<https://peroxidase.toulouse.inra.fr>) indicated as reference square nodes (cf. Table 3.3). (g) Catalase clades I, II, and III as defined in Fig. 3.2, including their first neighbors in the network. Note that several clusters remain unassigned to either clade. (h) Distribution of polypeptide sizes and neighborhood connectivity for nodes in each network. SSN calculation was conducted with the EFI-EST server (Gerlt et al. 2015). Network representative nodes indicate clusters in the UniRef90 database. Edges connecting representative nodes indicate 60% (70% for catalase-peroxidases) sequence similarity or higher, aiming to demarcate functional clusters. Cytoscape 3.9.1 (Shannon 2003) was used for SSN display and analysis. For simplicity, representative nodes whose NC value is lower than 5, with unspecified PDB or Swiss-Prot database entries, are not shown

Fig. 3.4 Maximum-likelihood bootstrapped tree of a uniform 672-sequence sample (UniRef90 identifiers) of monofunctional catalases SSN aligned with MAFFT (Kato et al. 2019) and calculated with RAxML (Stamatakis 2014). Assigned sequences from clades I, II, and III (Fig. 3.3g) are colored for reference



and *Basidiomycota* (*Agaricomycetes*), with a few members included in RedoxiBase. The remainder of the network includes enzymes from clades I and III, which are significantly less represented. Clade I enzymes appear divided in two clusters, one of bacterial *Bacillota* (*Bacilli*), *Bacteroidota* (*Cytophagia*), and *Deinococcota* (*Deinococci*) and one of plants *Streptophyta* (*Magnoliopsida*). Clade III enzymes cluster mostly with eukaryotic enzymes from *Eumetazoa* (insects, birds) and *Chordata* (reptiles, amphibians, mammals, cartilaginous fish), plus some bacterial enzymes from *Bacillota* (*Bacilli*), *Bacteroidota* (*Cytophagia*), and *Deinococcota* (*Deinococci*) (Fig. 3.3d). The fact that most monofunctional catalases are large clade II polypeptides suggests that smaller enzymes from clades I and III originate from an ancestral clade II enzyme that lost its flavodoxin-like domain. This hypothesis is in line with the observation that all archaeal (presumably ancient) monofunctional catalases belong to clade II, as well as phylogenetic analysis (Fig. 3.4).

Catalase-peroxidases comprise a large set of enzymes sharing at least 70% sequence similarity, predominantly from *Bacteria* (89%, 700–800 residues), *Eukaryota* (6.5%, 700–900 residues), and *Archaea* (4.5%, 700–750 residues). Neighborhood connectivity indicates that relatively few sequences share more than 300 neighbors on average, reflecting the overall high conservation of catalase-peroxidases, despite being widespread among all life-forms. Eukaryotic enzymes are largely from yeast and fungi, including seven RedoxiBase members (Fig. 3.3e), *Ascomycota* (*Saccharomyceta*), and *Basidiomycota* (*Agaricomycetes*, *Ustilaginomycetes*, *Tremellomycetes*); enzymes from *Archaea* are *Euryarchaeota* (*Stenosarchaea* group, *Methanobacteria*); and bacterial catalase-peroxidases are more diverse, including mainly *Actinomycetota* (*Actinomycetes*, *Acidimicrobiia*), *Acidobacteriota* (*Thermoanaerobaculia*, *Holophagae*, *Blastocatellia*, *Terriglobia*), *Bacillota* (*Bacilli*, *Clostridia*, *Tissierellia*), *Bacteroidota* (largely *Flavobacteriia*, *Cytophagia*, *Chitinophagia*, and *Saprospira*), *Balneolota* (*Balneolia*), *Bdellovibrionota* (*Bdellovibrionia*, *Bacteriovoracia*, *Oligoflexia*),

Campylobacteriota (ϵ -proteobacteria), *Chloroflexota*, *Cyanobacteriota* (*Cyanophyceae*), *Gemmatimonadota*, *Myxococcota*, *Planctomycetota*, and *Pseudomonadota* (α -, β -, and γ -proteobacteria).

Mn catalases are largely bacterial enzymes, with some archaeal members (Fig. 3.3f). The size distribution is somewhat bimodal, with subunits of ~200 and ~300 residues (Fig. 3.3h). The major groups of Mn catalases include *Bacillota* (predominantly *Clostridia*, *Bacilli*, *Erysipelotrichia*, *Mollicutes*, and *Negativicutes*), followed by a group of *Pseudomonadota* (α -, β -, and γ -proteobacteria), including RedoxiBase members ACxMnCat01, BcenMnCat, and KpMnCat01, a cluster of enzymes from *Bacillota* (*Bacilli*) and *Cyanobacteriota* (*Cyanophyceae*), and a group from *Bacteroidota* (*Cytophagia*, *Sphingobacteriia*), *Deinococcota* (*Deinococci*), *Pseudomonadota* (α -proteobacteria), including the RedoxiBase SmedMnCat and related archaeal enzymes from *Stenosarchaea* halophile group, including the RedoxiBase HlaMnCat.

Concluding Remarks

Catalases are crucial enzymes enabling aerobic and anaerobic metabolism in an oxygenated atmosphere, by readily eliminating hydrogen peroxide. Monofunctional heme catalases, catalase-peroxidases, and Mn catalases constitute a classic example of convergent evolution, since they evolved independently more than two billion years ago and are widespread over all life-forms. Heme catalase clades I and III seem to have evolved from a clade II ancestor that lost an $\alpha\beta\alpha$ extra domain, while many other heme catalases remain to be classified. Catalase-peroxidases have been surprisingly conserved throughout evolution and also harness the oxidative power of hydrogen peroxide to oxidize varied substrates. Mn catalases are particularly abundant among enterobacterial pathogens and strict anaerobes, indicative of their key role at withstanding microaerophilic conditions and oxidative stress of the host defense response. Such functional and structural diversities make these enzymes particularly suitable, not only for metabolic engineering applications, for the development of biosensors, and as targets for antimicrobial drug design but also to enlighten our understanding of protein evolution.

Acknowledgments This work was supported with funds from CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) and Agencia I+D+i (Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación), grants PIP0024 and PICT 2020-2669 to JMG. MBL and MBO are graduate fellows of CONICET. JMG is a staff member of CONICET.

References

Alfonso-Prieto M, Biarnés X, Vidossich P, Rovira C (2009) The molecular mechanism of the catalase reaction. *J Am Chem Soc* 131:11751–11761. <https://doi.org/10.1021/ja9018572>

- Alfonso-Prieto M, Vidossich P, Rovira C (2012) The reaction mechanisms of heme catalases: an atomistic view by ab initio molecular dynamics. *Arch Biochem Biophys* 525:121–130. <https://doi.org/10.1016/j.abb.2012.04.004>
- Atkinson HJ, Babbitt PC (2009) An atlas of the thioredoxin fold class reveals the complexity of function-enabling adaptations. *PLoS Comput Biol* 5:e1000541. <https://doi.org/10.1371/journal.pcbi.1000541>
- Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC (2009) Using sequence similarity networks for visualization of relationships across diverse protein Superfamilies. *PLoS One* 4:e4345. <https://doi.org/10.1371/journal.pone.0004345>
- Baier F, Tokuriki N (2014) Connectivity between catalytic landscapes of the metallo- β -lactamase superfamily. *J Mol Biol* 426:2442–2456. <https://doi.org/10.1016/j.jmb.2014.04.013>
- Blum M, Chang HY, Chuguransky S et al (2021) The inter pro protein families and domains database: 20 years on. *Nucleic Acids Res* 49:D344–D354. <https://doi.org/10.1093/nar/gkaa977>
- Chance B, Greenstein DS, Roughton FJW (1952) The mechanism of catalase action. I. Steady-state analysis. *Arch Biochem Biophys* 37:301–321. [https://doi.org/10.1016/0003-9861\(52\)90194-X](https://doi.org/10.1016/0003-9861(52)90194-X)
- Chelikani P, Fita I, Loewen PC (2004) Diversity of structures and properties among catalases. *Cell Mol Life Sci* 61:192–208. <https://doi.org/10.1007/s00018-003-3206-5>
- Copp JN, Anderson DW, Akiva E et al (2019) Exploring the sequence, function, and evolutionary space of protein superfamilies using sequence similarity networks and phylogenetic reconstructions. *Methods Enzymol* 620:315–347. <https://doi.org/10.1016/bs.mie.2019.03.015>
- Davidson R, Baas B-J, Akiva E et al (2018) A global view of structure-function relationships in the tautomerase superfamily. *J Biol Chem* 293:2342–2357. <https://doi.org/10.1074/jbc.M117.815340>
- Díaz A, Loewen PC, Fita I, Carpena X (2012) Thirty years of heme catalases structural biology. *Arch Biochem Biophys* 525:102–110. <https://doi.org/10.1016/j.abb.2011.12.011>
- Fita I, Rossmann MG (1985) The NADPH binding site on beef liver catalase. *Proc Natl Acad Sci USA* 82:1604–1608
- Gerlt JA, Bouvier JT, Davidson DB et al (2015) Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks. *Biochim Biophys Acta-Proteins Proteomics* 1854:1019–1037. <https://doi.org/10.1016/j.bbapap.2015.04.015>
- González JM (2021) Visualizing the superfamily of metallo- β -lactamases through sequence similarity network neighborhood connectivity analysis. *Heliyon* 7:e05867. <https://doi.org/10.1016/j.heliyon.2020.e05867>
- Halliwell B, Gutteridge JMC (2015) Free radicals in biology and medicine. Oxford University Press
- Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20:1160–1166. <https://doi.org/10.1093/bib/bbx108>
- Kirkman HN, Gaetani GF (1984) Catalase: a tetrameric enzyme with four tightly bound molecules of NADPH. *Proc Natl Acad Sci USA* 81:4343–4347. <https://doi.org/10.1073/pnas.81.14.4343>
- Kirkman HN, Rolfo M, Ferraris AM, Gaetani GF (1999) Mechanisms of protection of catalase by NADPH. Kinetics and stoichiometry. *J Biol Chem* 274:13908–13914. <https://doi.org/10.1074/jbc.274.20.13908>
- Klotz MG, Klassen GR, Loewen PC (1997) Phylogenetic relationships among prokaryotic and eukaryotic catalases. *Mol Biol Evol* 14:951–958. <https://doi.org/10.1093/oxfordjournals.molbev.a025838>
- Malik A, Kim SB (2019) A comprehensive in silico analysis of sortase superfamily. *J Microbiol* 57:431–443. <https://doi.org/10.1007/s12275-019-8545-5>
- Passardi F, Zamocky M, Favet J et al (2007) Phylogenetic distribution of catalase-peroxidases: are there patches of order in chaos? *Gene* 397:101–113. <https://doi.org/10.1016/j.gene.2007.04.016>
- Sánchez-Baracaldo P, Cardona T (2020) On the origin of oxygenic photosynthesis and cyanobacteria. *New Phytol* 225:1440–1446. <https://doi.org/10.1111/nph.16249>

- Sartorio MG, Cortez N, González JM (2021) Structure and functional properties of the cold-adapted catalase from *Acinetobacter* sp. Ver3 native to the Atacama plateau in northern Argentina. *Acta Crystallogr Sect D Struct Biol* 77:369–379. <https://doi.org/10.1107/S2059798321000929>
- Savelli B, Li Q, Webber M et al (2019) RedoxiBase: a database for ROS homeostasis regulated proteins. *Redox Biol* 26:101247. <https://doi.org/10.1016/j.redox.2019.101247>
- Shank M, Barynin V, Dismukes GC (1994) Protein coordination to manganese determines the high catalytic rate of dimanganese catalases. Comparison to functional catalase mimics. *Biochemistry* 33:15433–15436. <https://doi.org/10.1021/bi00255a025>
- Shannon P (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
- Shi Q, Wang H, Liu J et al (2020) Old yellow enzymes: structures and structure-guided engineering for stereocomplementary bioreduction. *Appl Microbiol Biotechnol* 104:8155–8170. <https://doi.org/10.1007/s00253-020-10845-z>
- Sies H, Berndt C, Jones DP (2017) Oxidative stress. *Annu Rev Biochem* 86:715–748. <https://doi.org/10.1146/annurev-biochem-061516-045037>
- Slauch JM (2011) How does the oxidative burst of macrophages kill bacteria? Still an open question. *Mol Microbiol* 80:580–583. <https://doi.org/10.1111/j.1365-2958.2011.07612.x>
- Smulevich G, Jakopitsch C, Droghetti E, Obinger C (2006) Probing the structure and bifunctionality of catalase-peroxidase (KatG). *J Inorg Biochem* 100:568–585. <https://doi.org/10.1016/j.jinorgbio.2006.01.033>
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Tararina MA, Allen KN (2020) Bioinformatic analysis of the Flavin-dependent amine oxidase superfamily: adaptations for substrate specificity and catalytic diversity. *J Mol Biol* 432:3269–3288. <https://doi.org/10.1016/j.jmb.2020.03.007>
- Vlasits J, Jakopitsch C, Bernroither M et al (2010) Mechanisms of catalase activity of heme peroxidases. *Arch Biochem Biophys* 500:74–81. <https://doi.org/10.1016/j.abb.2010.04.018>
- Welinder KG (1991) Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. *Biochim Biophys Acta* 1080:215–220. [https://doi.org/10.1016/0167-4838\(91\)90004-j](https://doi.org/10.1016/0167-4838(91)90004-j)
- Whittaker JW (2012) Non-heme manganese catalase—the ‘other’ catalase. *Arch Biochem Biophys* 525:111–120. <https://doi.org/10.1016/j.abb.2011.12.008>
- Zámocký M (2004) Phylogenetic relationships in class I of the superfamily of bacterial, fungal, and plant peroxidases. *Eur J Biochem* 271:3297–3309. <https://doi.org/10.1111/j.1432-1033.2004.04262.x>
- Zámocký M, Gasselhuber B, Furtmüller PG, Obinger C (2012) Molecular evolution of hydrogen peroxide degrading enzymes. *Arch Biochem Biophys* 525:131–144. <https://doi.org/10.1016/j.abb.2012.01.017>
- Zámocký M, Hofbauer S, Schaffner I et al (2015) Independent evolution of four heme peroxidase superfamilies. *Arch Biochem Biophys* 574:108–119. <https://doi.org/10.1016/j.abb.2014.12.025>

Chapter 4

Building the Bacterial Divisome at the Septum



Josiah J. Morrison and Jodi L. Camberg

Abstract Across living organisms, division is necessary for cell survival and passing heritable information to the next generation. For this reason, cell division is highly conserved among eukaryotes and prokaryotes. Among the most highly conserved cell division proteins in eukaryotes are tubulin and actin. Tubulin polymerizes to form microtubules, which assemble into cytoskeletal structures in eukaryotes, such as the mitotic spindle that pulls chromatids apart during mitosis. Actin polymerizes to form a morphological framework for the eukaryotic cell, or cytoskeleton, that undergoes reorganization during mitosis. In prokaryotes, two of the most highly conserved cell division proteins are the tubulin homolog FtsZ and the actin homolog FtsA. In this chapter, the functions of the essential bacterial cell division proteins FtsZ and FtsA and their roles in assembly of the divisome at the septum, the site of cell division, will be discussed. In most bacteria, including *Escherichia coli*, the tubulin homolog FtsZ polymerizes at midcell, and this step is crucial for recruitment of many other proteins to the division site. For this reason, both FtsZ abundance and polymerization are tightly regulated by a variety of proteins. The actin-like FtsA protein polymerizes and tethers FtsZ polymers to the cytoplasmic membrane. Additionally, FtsA interacts with later stage cell division proteins, which are essential for division and for building the new cell wall at the septum. Recent studies have investigated how actin-like polymerization of FtsA on the lipid membrane may impact division, and we will discuss this and other ways that division in bacteria is regulated through FtsZ and FtsA.

Keywords Actin · Cell wall · Cytoskeleton · Division machinery · Divisome · Peptidoglycan · Septation · Tubulin

J. J. Morrison · J. L. Camberg (✉)
Department of Cell and Molecular Biology, The University of Rhode Island,
Kingston, RI, USA
e-mail: josymorrison96@uri.edu; cambergj@uri.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2024

J. R. Harris, J. Marles-Wright (eds.), *Macromolecular Protein Complexes V*,
Subcellular Biochemistry 104, https://doi.org/10.1007/978-3-031-58843-3_4

The Essential Cell Division Machinery

In bacterial cell division, such as in the widely studied Gram-negative bacterium *Escherichia coli*, the highly conserved tubulin homolog FtsZ polymerizes at midcell to form a ringlike structure, termed the Z-ring. FtsZ may act as a scaffold for the recruitment of many proteins to the cell center to ensure division occurs. Chief among these proteins is FtsA and ZipA, which tether FtsZ polymers to the inner membrane. The actin homolog FtsA is the second most conserved cell division protein among bacteria (after FtsZ) and polymerizes with ATP, similar to actin. FtsA also interacts with other cell division proteins that stimulate peptidoglycan (PG) cell wall remodeling. In an initial assembly step, FtsZ, FtsA, ZipA, FtsE, and FtsX coalesce to recruit downstream cell division proteins (Fig. 4.1). Later stage cell division proteins FtsK, FtsB, FtsL, FtsQ, FtsW, FtsI, and finally FtsN arrive in a later step to initiate constriction and cell wall synthesis. Each step in bacterial cell division requires strict regulation to ensure that division occurs with high fidelity. Here, we will provide an overview of FtsZ-ring assembly and regulation. We will then discuss how regulation of the FtsA oligomeric state and interaction with other division proteins impact division. We will explore models of FtsA and FtsZ function in bacterial cells such as *E. coli*.

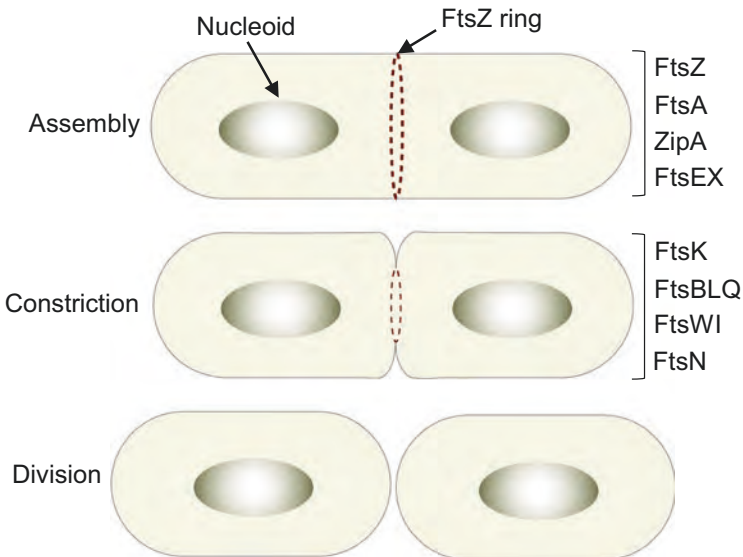


Fig. 4.1 Assembly and constriction of the cell division machinery in *E. coli*. During assembly of the divisome, FtsZ, FtsA, ZipA, FtsE, and FtsX localize first to the septum and recruit late-stage cell division proteins. Then, FtsK, FtsB, FtsL, FtsQ, FtsW, FtsI, and FtsN are recruited and lead to constriction, cell wall synthesis, and separation

Divisome Structure

In *E. coli* cell division, FtsZ polymerizes at midcell to form the Z-ring and is the initial step in the division pathway which results in recruitment of at least 30 cell division proteins, 12 of which are essential, to form a protein ultrastructure known as the divisome (Fig. 4.2) (Bi and Lutkenhaus 1991; McQuillen and Xiao 2020; Vaughan et al. 2004). FtsZ is tethered to the membrane by direct protein interactions with ZipA and FtsA (Hale and de Boer 1997; Tormo and Vicente 1984). FtsZ, FtsA, and ZipA are the earliest proteins to localize to the division site and appear to localize concurrently by microscopy (Rueda et al. 2003). The FtsEX complex also localizes in this first step and appears critical under conditions of low osmolarity. FtsE is a cytoplasmic ATPase, while FtsX is a membrane protein, and together, the FtsEX complex binds to FtsA and FtsZ during cell division and may regulate the FtsA oligomeric state (Corbin et al. 2007; Du et al. 2016; Reddy 2007; Schmidt et al. 2004). The FtsEX complex also appears to have another function in activating PG synthesis enzymes and amidases, which requires ATPase activity (Arends et al. 2009; Du et al. 2016). Although it is a predicted ABC transporter, it remains unclear whether the FtsEX complex transports anything important for division. FtsK, FtsB, FtsL, FtsQ, FtsW, FtsI, and FtsN all localize later in cell division but are critical for this process. FtsK binds to chromosomes and is embedded in the cell membrane and appears to play a role in coordinating division with DNA segregation (Liu et al. 1998; Stouf et al. 2013; Wang and Lutkenhaus 1998). FtsB, FtsL, and FtsQ form a complex that is critical for recruitment of the FtsW and FtsI to the midcell and may inhibit the PG synthesis activities of FtsW and FtsI (Liu et al. 2015; Marmont and

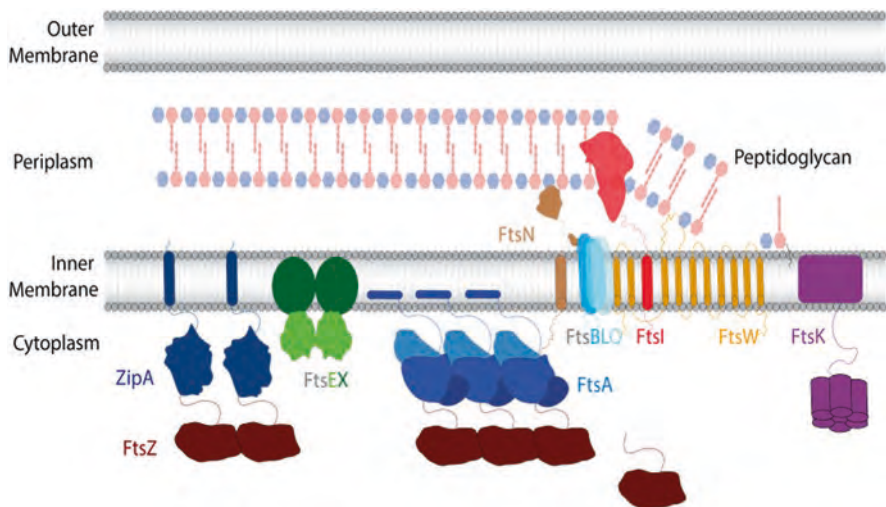


Fig. 4.2 Divisome assembly in *E. coli*. Divisome proteins in the cytoplasm, the inner membrane, and the periplasm are shown. Whenever possible, confirmed protein-protein, protein-membrane, and protein-PG interactions are reflected by their respective spatial arrangements

Bernhardt 2020). FtsW contains ten transmembrane domains that anchor it in the cytoplasmic membrane and has glycosyltransferase activity to link the *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAMA) sugar backbone of PG (Khattar et al. 1994). FtsI contains a single transmembrane domain with a large domain in the periplasm that catalyzes the cross-linking of PG stem peptides, called transpeptidation (Weiss et al. 1997). FtsN is a bitopic membrane protein that is thought to act as a trigger for septation through coordinated interaction with both FtsA and FtsWI, as well as the FtsBLQ complex to allow division to occur (Liu et al. 2015; Marmont and Bernhardt 2020; Park et al. 2020). One model for FtsN's role as a trigger for septation is of FtsN relieving repression of FtsBLQ complex on FtsWI. Evidence for this is that FtsN requires FtsQ for localization to the septum and that suppressor mutations in FtsB have been shown to suppress deletion of the FtsN essential domain, which indicates a role for FtsBLQ in FtsN-mediated stimulation of PG synthesis (Goehring et al. 2006; Liu et al. 2015). The interaction with FtsA in the cytoplasm could lead to a critical mass of FtsN accumulating at the midcell to stimulate PG synthesis.

FtsZ Conformations and Oligomerization into Polymers

While a complete understanding of the dynamics of FtsZ polymerization in vitro is still unclear, significant progress has been made thus far. FtsZ is a bacterial tubulin homolog and like tubulin, polymerizes in a GTP-dependent manner (Bramhill and Thompson 1994). However, recent work in *E. coli* has shown that polymerization of FtsZ occurs with opposite kinetic polarity and curvature to that found in microtubules, which may reflect the different roles of the proteins in prokaryotic versus eukaryotic division (Du et al. 2018; Housman et al. 2016). Additionally, while tubulin assembles into microtubules that are composed of 13 filaments assembled around a hollow core, FtsZ filaments assemble linearly and can bundle laterally to form larger complexes. Interestingly, two distinct conformations have been observed for FtsZ, known as the relaxed (R) state and tense (T) state (Fig. 4.3a) (Fujita et al. 2017; Matsui et al. 2014). The R state has been observed for monomeric FtsZ regardless of nucleotide state, while the T state refers to FtsZ incorporated into filaments (Wagstaff et al. 2017). FtsZ polymerizes into filaments in a head-to-tail arrangement when bound to GTP, which, due to being hydrolyzed to GDP, induces filament bending (Lu et al. 2000; Lu and Erickson 1999). This leads to a filament that is growing at the GTP-bound end and disassembling at the GDP-bound end (Fig. 4.3b) (Ramirez-Diaz et al. 2018). As a result, it has been observed in vitro and in vivo that FtsZ polymers exhibit directional growth, known as treadmilling, and the significance of this in division will be discussed in the section below (Bisson-Filho et al. 2017; Loose and Mitchison 2014; Ramirez-Diaz et al. 2018; Yang et al. 2017). It has been proposed that this observed directional growth is a result of a conformational switch that occurs in FtsZ when bound within a filament as opposed to when it is monomeric, which would cause it to be more energetically favorable

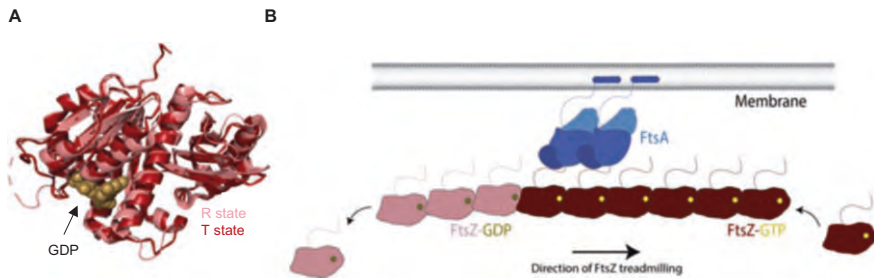


Fig. 4.3 FtsZ conformational changes and treadmilling. (a) Representative R (salmon) and T (dark red) states of *E. coli* FtsZ. *E. coli* FtsZ structure crystalized with GDP (gold) (PDB 6LL6) (Yoshizawa et al. 2020) was used as representative of the R state, and *E. coli* FtsZ sequence was modeled onto *M. jannaschii* FtsZ (PDB 1W5A) (Oliva et al. 2004) using ProMod3 by the SWISS-MODEL homology server as representative of the T state (Guex et al. 2009; Studer et al. 2021). (b) Treadmilling by FtsZ filaments. FtsZ binding to GTP induces FtsZ filament assembly which is tethered to the membrane by FtsA. GTP hydrolysis by FtsZ causes filament bending and favors release of FtsZ from the filament. Thus, FtsZ filaments exhibit directional growth (treadmilling) toward the FtsZ-GTP end

for FtsZ to stay as a polymer rather than disassemble (Wagstaff et al. 2017). In addition to GTP dependency, FtsZ polymerization requires other factors such as magnesium and is affected by pH fluctuations, ion composition, and crowding (Pacheco-Gómez et al. 2011; Rivas et al. 2000). One major implication of FtsZ polymerization is that it increases interactions with other divisome proteins through multivalency. Indeed, it has been observed that the membrane anchors FtsA, ZipA, and the nucleoid exclusion factor SlmA interact more readily with FtsZ polymers rather than unassembled FtsZ (Conti et al. 2018; Du et al. 2015).

Z-Ring Architecture and Function

Closer examination of the *in vivo* architecture of the Z-ring reveals that it is not one singular ring wrapped around midcell, but is instead a highly dynamic, rapidly exchanging structure. Super-resolution photoactivated localization microscopy (PALM) imaging of fluorescent FtsZ at midcell shows that FtsZ forms a staggered and discontinuous ring which may contain overlapping FtsZ filaments (Fu et al. 2010). The discontinuous nature of the Z-ring may allow dynamic FtsZ polymerization and protein interactions to occur at various locations about the midcell. FtsZ dynamic polymerization within the cell appears as a treadmilling effect and has been observed *in vivo* by total internal reflection fluorescence microscopy (TIRFM), where FtsZ monomers are added to one end of the filament while disassembling at the opposite end (Bisson-Filho et al. 2017). Interestingly, in *B. subtilis*, the treadmilling rate of FtsZ appears to dictate the rate of PG synthesis, while in *E. coli* the rate of treadmilling has no effect. This may reflect differences in cell wall architecture between Gram-negative and Gram-positive bacteria (Bisson-Filho et al. 2017;

Yang et al. 2017). Recent studies have indicated that there may be several treadmilling tracks with different rates in *E. coli*, including one track going at the speed of FtsZ treadmilling and another slower track by PG synthesis enzymes such as FtsW and FtsI, moving to synthesize new septal PG (Lyu et al. 2022; McCausland et al. 2021; Yang et al. 2021). In this model, FtsZ treadmilling would allow for more even distribution of PG insertion by PG synthesis enzymes around the division site (McQuillen and Xiao 2020).

Many studies have also attempted to address whether FtsZ polymers can exert the force required to divide the cell into two in vivo. Polymerization of *E. coli* FtsZ has been shown to bend and curve lipids in vitro when a membrane targeting sequence (MTS) is attached to it; however, it was unclear whether this force would be physiologically relevant in vivo (Osawa et al. 2008, 2009). Additionally, FtsZ adopts straight and curved conformations in the presence of GTP and GDP nucleotides, respectively, which prompted speculation that hydrolysis of GTP by FtsZ may exert force on the membrane to begin invagination (Lu et al. 2000; Lu and Erickson 1999). However, one study by Coltharp and colleagues showed that disruption of critical functions of FtsZ including GTPase activity disrupts treadmilling activity but does not change the overall cell constriction rate, indicating that FtsZ is likely not the main constriction force generator (Coltharp et al. 2016). This is additionally supported by FtsZ temperature-sensitive alleles of *ftsZ* in *E. coli*, such as *ftsZ84*, which confer reduced GTPase activity, due to a substitution mutation at glycine 105, and prevent growth by the strain at high temperatures; however, the strains are able to grow normally at the permissive temperature (Addinall et al. 1996). Lastly, FtsZ appears to leave the divisome before the septum has fully divided, indicating that FtsZ is certainly not the only constriction force generator throughout septation (Söderström et al. 2014, 2016).

Regulation of FtsZ Function and Z-Ring Architecture

FtsZ assembly is critical for division, and FtsZ is regulated by a multitude of proteins and at several levels. Here, we will discuss the two broad ways FtsZ is regulated: (1) through regulation of FtsZ abundance via transcriptional regulation, synthesis, and degradation and (2) through regulation of FtsZ polymerization and bundling.

Transcriptional regulation of *ftsZ*—The *ftsZ* gene is a component of the division and cell wall (*dcw*) operon, which is a highly conserved cluster of genes required for division and cell wall synthesis in bacteria (Fig. 4.4) (Ayala et al. 1994; Miguel Vicente and Errington 1996). In *E. coli*, this cluster contains 16 genes and *ftsZ* is the second to the last gene in this cluster, which is just downstream of *ftsA* and *ftsQ*. Transcriptionally, this gene cluster is highly regulated, which allows the cell to adjust to differing environmental and cellular conditions. Under many conditions, the entire *dcw* operon is transcribed as a single polycistronic mRNA and contains only a rho-independent termination sequence at the end of *lpxC* (*envA*), the last gene

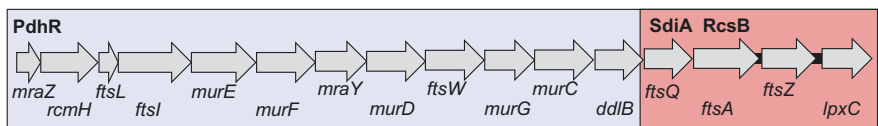


Fig. 4.4 Division and cell wall synthesis (*dcw*) operon in *E. coli*. Sixteen genes are part of the *dcw* gene cluster in *E. coli*. *ftsQAZ* and *lpxC* expressions are activated by the quorum-sensing protein SdiA and the two-component system RcsB. The entire *dcw* operon is repressed by PdhR which may allow coordination of the cell metabolic state with division. Light blue color represents regulation by PdhR. Red color represents regulation by SdiA and RcsB

in the operon (Vicente et al. 1998). There are many promoters located in the *dcw* cluster; specifically, there are several that are located upstream of the *ftsQAZ* gene region (within the *ddlB* gene) that can control expression of these essential division genes (Vicente et al. 1998). These promoters appear to be regulated through a variety of factors including SdiA, RcsB, and ppGpp (Aldea et al. 1990; Gervais et al. 1992; Yamamoto et al. 2001). However, only about 30–40% of the transcription of *ftsQAZ* is initiated at these promoter sequences, as evidenced by transposon insertion in the gene immediately upstream of this region (Flärdh et al. 1998; de la Fuente et al. 2001). This means that a large portion of *ftsZ* transcriptional regulation is coming at the expense of the entire *dcw* gene cluster. Recently, the pyruvate dehydrogenase regulator protein PdhR, which is activated by CRP-cAMP, has been shown to repress *dcw* gene cluster activity (Göhler et al. 2011). This may act as a key link between the regulation of cell division and sensing of the cell metabolic state as it has previously been shown that PdhR repression is relieved when a high abundance of pyruvate is present, indicating a high rate of glycolytic flux (Kotte et al. 2010). Additionally, the influence of PdhR-negative regulation on the *dcw* cluster is particularly noticeable in a deletion strain. The *pdhR* deletion strain showed a completely abrogated lag phase, indicative of PdhR no longer being present to sense the cell metabolic state or repress transcription of the *dcw* cluster (Quail et al. 1994; Sekar et al. 2018).

Degradation of FtsZ—FtsZ abundance is regulated proteolytically by the ATP-dependent protease ClpXP (Camberg et al. 2009; Flynn et al. 2003). ClpX is a hexameric member of the AAA+ protein family, and ClpP is a serine protease that contains two stacked heptameric rings that form a barrel structure; ClpX and ClpP are highly conserved in bacteria (Hanson and Whiteheart 2005; Maurizi et al. 1990; Wojtkowiak et al. 1993; Yu and Houry 2007). ClpX assembles with one or both ends of the ClpP barrel and acts to recognize and unfold substrates for subsequent ClpP degradation (Grimaud et al. 1998). ClpX binds to specific recognition sequences located within the substrate, called degrons. Specifically, ClpX recognizes two degrons within FtsZ for recognition and subsequent degradation (Camberg et al. 2014). Degradation of FtsZ by ClpXP has been observed both in vitro and in vivo (Camberg et al. 2014; Camberg et al. 2009; Männik et al. 2018). Overexpression of ClpXP leads to long cell filaments, presumably from excess FtsZ degradation (Camberg et al. 2009; Viola et al. 2017). During normal logarithmic growth of *E. coli*, FtsZ levels appear fairly constant; however, during periods of slow growth

or starvation conditions for the cell, ClpXP plays a role in rapid FtsZ degradation after division (Adams and Errington 2009; Männik et al. 2018; Rueda et al. 2003; Sekar et al. 2018). ClpXP severs FtsZ polymers and deletion of *clpX* or *clpP* leads to slower recovery of Gfp-FtsZ after photobleaching (FRAP), which indicates that ClpXP is critical for regulating the FtsZ polymeric state and thus FtsZ-ring dynamics (Camberg et al. 2009; Viola et al. 2017).

MinCDE regulation of Z-ring placement—Since FtsZ polymerization is critical for division to occur at midcell, regulation of its polymerization and placement over the midcell location is crucial. Since the discovery of mini cells, short cells lacking chromosomes, much has been discovered about the Min system, a critical negative regulator of FtsZ assembly (Adler et al. 1967). The Min system acts to negatively regulate FtsZ polymerization at the cell poles and ensure its correct assembly at the midcell. The Min system in *E. coli* is composed of three proteins, MinC, MinD, and MinE, and oscillates from cell pole to cell pole creating a time-averaged gradient that ensures MinC concentration is greatest at the cell pole to prevent FtsZ polar assembly (Fig. 4.5a) (de Boer et al. 1989; Raskin and de Boer 1999). The MinC dimer acts as the FtsZ polymerization inhibitor and interacts with the FtsZ C-terminal end (CTE) and another region near the polymerization interface on FtsZ to destabilize the filament (Hu et al. 1999; LaBreck et al. 2018; Park et al. 2018). MinD binds to MinC and is also capable of binding to ATP, which enables dimerization and binding to the membrane (de Boer et al. 1991). MinE forms a ringlike structure adjacent to the lagging edge of MinD, stimulating its ATPase activity to release it from the membrane, which causes both MinD and MinC to begin the oscillatory event (Hu and Lutkenhaus 1999; Hu et al. 2001; Raskin and de Boer 1997). MinC

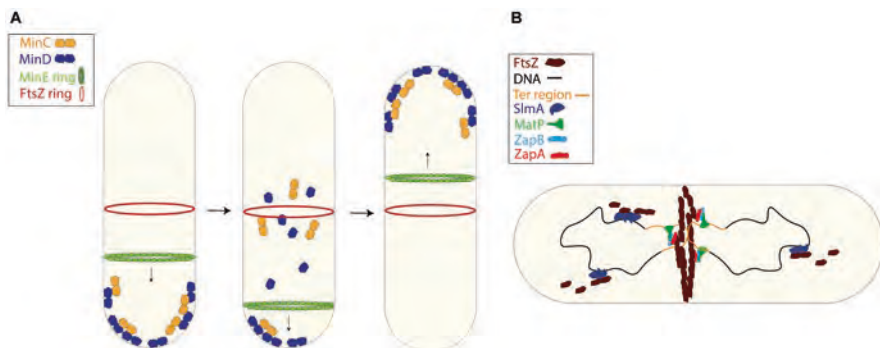


Fig. 4.5 Regulation of FtsZ-ring placement in *E. coli*. (a) The Min system in *E. coli* ensures Z-ring placement at the midcell. As described in text, MinC destabilizes FtsZ polymers, and the MinE ring stimulates MinD membrane release. This enables these components to oscillate from cell pole to cell pole to prevent Z-ring assembly at the poles. (b) SlmA-mediated nucleoid occlusion and ter linkage. SlmA binds to DNA everywhere except the ter region and destabilizes FtsZ polymers to prevent chromosome severing. MatP binds to the ter region and is linked to the FtsZ-ring through interactions with ZapB and ZapA. This makes assembly of the Z-ring over the ter region more favorable.

and MinD dimers each have two binding sites for their partners and assemble into alternating copolymers in the presence of ATP *in vitro* (Conti et al. 2015; Ghosal et al. 2014). However, mutations that prevent copolymer formation but preserve the single interactions between MinC and MinD do not impair division *in vivo*, so it remains unclear what possible role copolymers may play (Park et al. 2015). Through the oscillation process in *E. coli*, FtsZ assembly is prevented at the cell poles where MinC concentration is the highest, thus favoring FtsZ-ring assembly at midcell. It is important to note that not all organisms contain an oscillatory Min system like *E. coli*. *Bacillus subtilis* lacks *minE* and has long been thought to have a stable population of MinC and MinD at the cell poles to prevent Z-ring assembly there (Marston et al. 1998; Marston and Errington 1999). However, more recent reports have challenged this idea and suggested that MinC and MinD may still have a dynamic system in *B. subtilis* that oscillates between the midcell and the cell pole and additionally suggest that MinC does not destabilize FtsZ polymers and instead prevents FtsZ polymers from furthering division by preventing lateral interactions (Gregory et al. 2008; Helge et al. 2021). Interestingly, a recent study reported that MinD is critical for disassembly of polar Z-rings remaining from the previous division event and also inhibits polar PG synthesis (Yu et al. 2020).

SlmA and ter linkage prevents chromosome severing by the Z-ring in E. coli—Another important step in regulation of FtsZ polymerization is ensuring that FtsZ rings do not assemble over the chromosome(s), termed nucleoid occlusion (NO), which could lead to severing of the DNA. In *E. coli*, SlmA is a DNA-binding protein that prevents FtsZ polymerization over the nucleoid directly (Fig. 4.5b) (Bernhardt and de Boer 2005). SlmA binds to the CTE of FtsZ and severs FtsZ polymers, an activity which is stimulated upon SlmA binding to recognition signals on the chromosome (Cho et al. 2011; Du and Lutkenhaus 2014). These recognition sequences appear throughout the *E. coli* genome with the notable exception of the terminus (*ter*) region (Cho et al. 2011). Interestingly, SlmA has recently been shown to bind to lipids, indicating that the membrane and membrane proteins may help regulate the SlmA-FtsZ interaction (Ángel et al. 2020). In addition to SlmA-negative regulation, the *ter* region of the chromosome appears to positively regulate Z-ring placement (Bailey et al. 2014). Studies have shown that MatP, which is an organizer of the *ter* region, binds to ZapB, which indirectly associates with the Z-ring through interaction with ZapA (Fig. 4.5b) (more on Zap regulation of FtsZ below) (Espéli et al. 2012; Galli and Gerdes 2010; Mercier et al. 2008). Although ZapA, ZapB, and MatP appear to play a role in the localization and assembly of FtsZ at the *ter* region, it is still unclear precisely how this process occurs (Bailey et al. 2014).

Regulation of FtsZ polymerization during the SOS response and glucose metabolism—During the cellular stress response to DNA damage in *E. coli* (the SOS response), division is halted by SulA, which directly interacts with FtsZ (Bi and Lutkenhaus 1993; Dorina et al. 1998; Mukherjee et al. 1998). SulA inhibits FtsZ assembly through a sequestering mechanism to prevent free FtsZ subunits from polymerizing and thus increasing the critical concentration of assembly (Chen et al. 2012). A fascinating link between division of the cell and overall cell metabolism

has been discovered in OpgH. OpgH is an enzyme that is important for osmoregulated periplasmic glucan synthesis and also interacts with and inhibits FtsZ polymer formation (Bohin and Kennedy 1984; Hill et al. 2013; Weissborn and Kennedy 1984). UDP-glucose stimulates OpgH inhibition of FtsZ assembly to coordinate division with nutritional availability (Hill et al. 2013).

Zap proteins facilitate FtsZ lateral interactions—Another way of regulating FtsZ assembly is through stabilization of FtsZ polymer lateral interactions. The Zap proteins, ZapA, ZapB, ZapC, ZapD, and ZapE, appear to modulate FtsZ filament bundling, which may aid in development and maturity of the Z-ring (Marteyn et al. 2014; Durand-Heredia et al. 2012; Ebersbach et al. 2008; Gueiros-Filho and Losick 2002; Hale et al. 2011; DiBiasio et al. 2021). ZapA is a tetramer, an element critical to its FtsZ bundling capabilities (Low et al. 2004; Pacheco-Gómez et al. 2013). Moreover, ZapA helps align FtsZ filaments in clusters that are more organized and less dynamic as revealed by PALM imaging of the Z-ring in *zapA* deletion strains (Buss et al. 2013). An interesting recent study using TIRF microscopy on lipid bilayers has showed that ZapA straightens FtsZ filaments via bundling in a cooperative manner; however, this activity had no effect on the overall treadmilling rate of FtsZ (Caldas et al. 2019). ZapB does not interact directly with FtsZ and requires ZapA; however, ZapB significantly enhances the bundling activity by ZapA (Buss et al. 2013; Galli and Gerdes 2010). ZapC and ZapD both interact directly with FtsZ subunits and enhance lateral bundling although with distinct mechanisms as ZapC is a monomer and interacts with the globular FtsZ core and ZapD is a dimer that interacts with the FtsZ CTE (Schumacher et al. 2016, 2017). ZapE is an ATPase that binds to FtsZ. Although an initial report suggested it may destabilize FtsZ polymers, it has been recently demonstrated in vitro to bundle FtsZ polymers (Marteyn et al. 2014; DiBiasio et al. 2021). While not essential for division, ZapE may play a more important role during conditions of cellular stress (DiBiasio et al. 2021).

The membrane anchors ZipA and FtsA regulate FtsZ polymerization dynamics—Lastly, FtsZ assembly may be influenced by its anchoring proteins ZipA and FtsA. Several early reports showed that purified soluble ZipA (lacking its transmembrane domain) caused FtsZ bundling in vitro by transmission electron microscopy (TEM) and prevented dilution-induced FtsZ disassembly (Hale et al. 2000; Kuchibhatla et al. 2011; RayChaudhuri 1999). This was further supported by the observation that ZipA promoted static FtsZ polymer formation by TIRF microscopy on lipid monolayers, in contrast to the highly dynamic structures observed when FtsA was substituted for ZipA (Loose and Mitchison 2014). Another study showed that ZipA appears to stabilize FtsZ-GDP mini-rings formed in solution in addition to FtsZ-GTP filaments (Chen et al. 2017). A report contradicting this shows that ZipA, at physiologically relevant concentrations, may play a static role in anchoring FtsZ polymers to the membrane (Krupka et al. 2018). In contrast to ZipA, the actin homolog FtsA destabilizes FtsZ polymer formation (Conti et al. 2018; Krupka et al. 2017; Loose and Mitchison 2014). Destabilization of FtsZ polymers may allow FtsA to regulate FtsZ interaction partners by increasing FtsZ polymer turnover.

FtsA Is an Actin Homolog that Tethers FtsZ to the Membrane and Interacts with Late Divisome Proteins to Facilitate Division in *E. coli*

FtsA is a bacterial actin homolog, which, similar to actin, self-interacts and hydrolyzes ATP (Bork et al. 1992; Conti et al. 2018; Herricks et al. 2014; Szwedziak et al. 2012; Tormo and Vicente 1984). FtsA has a C-terminal amphipathic helix, referred to as the membrane targeting sequence (MTS), and acts as a membrane anchor to tether FtsZ polymers to the cytoplasmic membrane (Hale and de Boer 1997; Pichoff and Lutkenhaus 2002; Pichoff et al. 2005). Although ZipA has a similar function in tethering FtsZ polymers to the membrane, FtsA is more highly conserved than ZipA. Mutations in the sequence of FtsA can compensate for *zipA* deletion, and, unlike ZipA, FtsA interacts with a multitude of cell division proteins, including FtsZ, ZipA, FtsEX, FtsK, FtsQ, FtsWI, and FtsN. This suggests that FtsA may play a significant organizing role in *E. coli* cell division (Busiek et al. 2012; Corbin et al. 2004; Du et al. 2016; Geissler et al. 2003; Geissler and Margolin 2005; Park et al. 2021; Vega and Margolin 2019a, b). We will first discuss FtsA interaction with other proteins and then examine how FtsA functional abilities help achieve division.

FtsA interactions with FtsZ and ZipA—FtsA interacts directly with the C-terminal tail of FtsZ where it competes with ZipA for binding. FtsA also binds directly to the cytoplasmic membrane via a C-terminal amphipathic helix which is required for in vivo function (Haney et al. 2001; Mosyak et al. 2000; Pichoff et al. 2005; Szwedziak et al. 2012). Although overlapping with FtsA in basic function, ZipA contains a transmembrane domain, which anchors it in the membrane for FtsZ tethering (Hale and De Boer 1997; Pichoff and Lutkenhaus 2002). FtsA enhances FtsZ filament dynamics, and ZipA promotes stable FtsZ polymers. Thus, this competition between FtsA and ZipA for FtsZ binding may aid in regulation of dynamic treadmilling behavior of FtsZ (García-Soriano et al. 2020; Loose and Mitchison 2014).

FtsA interacts with late-stage division proteins to promote division—FtsA has been implicated in direct interactions with many other division proteins. One of these protein partners is FtsEX. FtsEX is an ABC transporter that includes an ATPase that localizes to the division ring (Schmidt et al. 2004). FtsEX binds to both FtsZ and FtsA, and it has been hypothesized that its ATPase activity may activate FtsA for recruitment of later stage cell division proteins as well as stimulate cell wall hydrolysis and cleavage (Arends et al. 2009; Corbin et al. 2007; Du et al. 2016). The membrane protein FtsK, which is a DNA translocase, also has been implicated to interact with FtsA, and facilitates chromosome segregation during cell division (Berezuk et al. 2014, 2020; Wang and Lutkenhaus 1998). Lastly, the bitopic membrane protein FtsN was originally isolated as a suppressor of the *ftsA 12* temperature-sensitive (ts) phenotype and is thought to act as a trigger for septation (Dai et al. 1993; Lutkenhaus 2009). FtsN contains three domains in its periplasmic region, a transmembrane domain, and a cytoplasmic domain which interacts with the 1C domain of FtsA (Busiek et al. 2012; Corbin et al. 2004; Dai et al. 1996). FtsN stimulates PG synthesis by FtsWI both through relieving repression of FtsWI by FtsBLQ

and through direct interaction with FtsWI to redistribute it around the circumference of the midcell (Liu et al. 2015; Lyu et al. 2022). Recently, it has been shown that FtsN may regulate the FtsA oligomeric state (Nierhaus et al. 2022). It is thought that the FtsA-FtsN interaction is crucial for linking the early division process with PG remodeling, thus enabling division to occur.

FtsA structural models—The first FtsA structural model was published by Jan Lowe’s group describing FtsA from the thermophile *Thermotoga maritima* (van den Ent and Löwe 2000). This structure confirmed the previous hypothesis that FtsA is a bacterial actin homolog with an actin fold for binding of ATP (Bork et al. 1992; Kabsch and Holmes 1995). Other notable bacterial actin homologs include the plasmid segregation factor ParM and the cell shape determining protein MreB (Shaevitz and Gitai 2010; Stoddard et al. 2017). The structure also revealed that FtsA contains a 1C domain that is switched in orientation with subdomain 2 of actin (van den Ent and Löwe 2000). However, this crystal structure was that of an FtsA monomer, and later a crystal structure of an FtsA dimer from *T. maritima* was published (Szwedziak et al. 2012) (Fig. 4.6a). This structure showed significant changes in the FtsA dimer interface when compared to yeast actin or even another bacterial actin homolog MreB, as FtsA interacts with itself through contacts with the 1A and 2B subdomains and the 1C and 2A subdomains (Fig. 4.6a), whereas actin self-interacts via repeated contacts with domains 2 and 4 with 3 (Fig. 4.6b) (Carballido-López 2006; Szwedziak et al. 2012).

The role of ATP interaction for FtsA function—The role of ATP interaction for FtsA function is not well characterized. The crystal structures of *T. maritima* FtsA and a recent *E. coli* FtsA structure show co-crystallization of ATP in the ATP-binding site indicating that FtsA binds to ATP with high affinity (van den Ent and Löwe 2000; Nierhaus et al. 2022; Szwedziak et al. 2012). Many temperature-sensitive FtsA mutations map near the ATP-binding pocket, indicating that ATP interaction may be important in vivo (Herrick et al. 2014). In one study, ATP

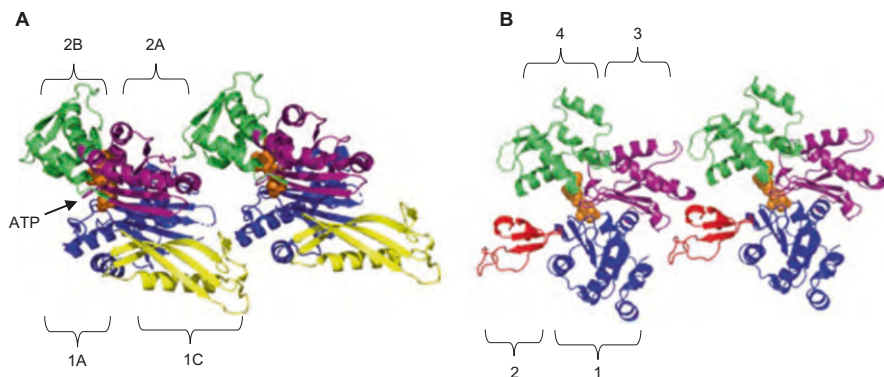


Fig. 4.6 FtsA and actin structural homology. (a) *T. maritima* FtsA dimer (PDB 4A2A) (Szwedziak et al. 2012) crystallized with ATP. (b) Yeast actin (PDB 1YAG) (Vorobiev et al. 2003) arranged as a dimer. Subdomains of homologous regions in FtsA and actin are the same color

promoted recruitment and dynamic reorganization of fluorescent FtsZ on a supported membrane via TIRF (Loose and Mitchison 2014). ATP also appears to aid *S. pneumoniae* FtsA helical polymerization and recruitment to lipids (Krupka et al. 2014). Lastly, a recent study showed that ATP enhances FtsA recruitment of FtsZ to lipid vesicles in vitro (Conti et al. 2018).

Since FtsA contains an ATPase domain similar to that of yeast actin, it was hypothesized that it can hydrolyze ATP (Bork et al. 1992). Many early studies were unable to determine significant FtsA-ATP hydrolysis activity, likely due to difficulties purifying it from inclusion bodies (Martos et al. 2012; Yim et al. 2000). Low FtsA-ATP hydrolysis rates were finally detected first in *B. subtilis* and then later in *E. coli* (Feucht et al. 2001; Herricks et al. 2014). Recently, robust ATP hydrolysis has been reported for *E. coli* FtsA by a method that copurifies phospholipid (PL) vesicles (Conti et al. 2018). This study reports that the PL interaction is critical for rapid ATP hydrolysis as removal of PLs abolishes ATPase activity (Conti et al. 2018). Additionally, this study shows that ATP stimulates reshaping and reorganization of lipid vesicles by TEM, likely as a result of polymerization on the lipid surface (Conti et al. 2018). An additional study showed that a mutation in the active site of FtsA, targeting a Glu residue (E14R) that directly contacts the Mg^{2+} in the active site, reduces ATP hydrolysis and lipid remodeling activity. This led to a model that local membrane remodeling through FtsA-ATP interaction may provide an open and accessible location on the opposing side of the membrane to direct PG synthesis (Conti et al. 2018; Morrison et al. 2022).

Actin conformational dynamics may provide insight into the role of FtsA-ATP interaction for function—Actin is known to form double-stranded filaments that are dependent on ATP and divalent cations and have a variety of functions in eukaryotes, ranging from maintenance of cell shape and cell motility, to driving muscle contraction (Kang et al. 2012; Narita 2011). The divalent cation and ATP help to stabilize the G-actin molecule, and essentially all G-actin monomers are nucleotide bound as the cellular concentration of ATP is in the millimolar range (Kabsch et al. 1990; Rosenblatt et al. 1995). Two G-actin-ATP monomers then associate, and a third is added as part of the thermodynamically unfavorable nucleation step (Sept and McCammon 2001). Once actin molecules bind to each other, a flattening of the molecule occurs to switch the conformation from G- to F-actin which facilitates interactions with subsequent actin molecules on that end (the barbed end) (Fig. 4.7a) (Chou and Pollard 2019; Merino et al. 2018). The flattening of the actin molecules repositions amino acids within the ATP-binding site to increase ATP hydrolysis of F-actin 42,000 fold compared to G-actin (Chou and Pollard 2019; Merino et al. 2018). The ADP and inorganic phosphate (P_i) are released very slowly, F-actin depolymerizes at the minus end, and F-actin-ADP is then converted back to G-actin-ADP and exchanged with ATP to reinitiate the cycle (Pollard 1986). Many other proteins within the eukaryotic cell help regulate actin nucleation, assembly, and depolymerization (Pollard 2016). Since FtsA is an actin-like protein, it can be hypothesized that it may also have several distinct conformations within the bacterial cell cycle. FtsA differs from actin in that it can bind directly to the membrane and does not require other proteins for tethering. Binding to PLs by FtsA greatly

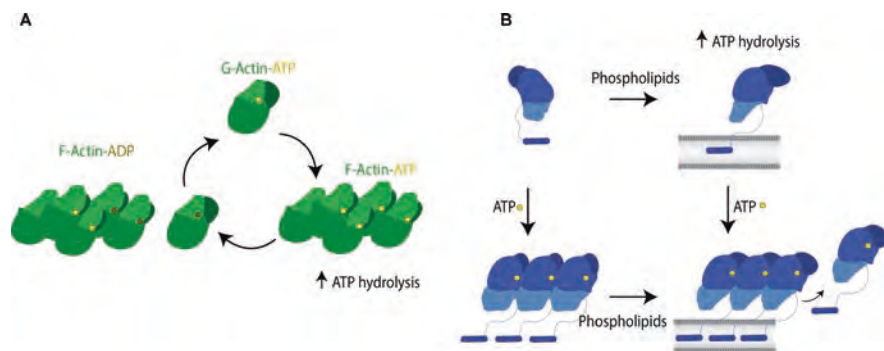


Fig. 4.7 Actin and FtsA conformational transitions. **(a)** G-actin flattens its conformation into F-actin upon being incorporated into a filament which is stimulated by ATP. This conformational change significantly increases ATP hydrolysis. As ATP becomes hydrolyzed by actin subunit, disassembly is favored and the ADP in the actin molecule becomes exchanged with ATP to change back to G-actin. **(b)** Model for FtsA conformational transitions. PL binding stimulates robust ATP hydrolysis by FtsA. ATP binding induces FtsA polymerization. FtsA bound to PLs, and ATP may allow for dynamic assembly and disassembly on the lipid membrane

stimulates ATP hydrolysis (~90% slower without PLs) suggesting that PLs stabilize a conformation associated with high catalytic activity (Conti et al. 2018). Additionally, several reports have indicated that FtsA lacking the C-terminal MTS appears to form robust, stable polymers (Krupka et al. 2014; Morrison et al. 2022; Pichoff and Lutkenhaus 2002). Therefore, binding of FtsA to the PL membrane may cause a conformational change that leads to increased ATP hydrolysis, which could regulate ATP-dependent dynamic assembly and disassembly on the lipid membrane (Fig. 4.7b). This may be similar to how ATP hydrolysis and Pi release causes actin to favor depolymerization over time. Additional proteins such as ZipA, FtsN, FtsZ, and FtsEX, among others, may regulate the FtsA oligomeric state (see later section).

FtsA self-interaction by polymerization—While much is known about how yeast actin uses cations and ATP to polymerize directionally in eukaryotes, the mechanistic details and the exact role of FtsA self-interaction in bacterial division remain unclear. It would be expected for FtsA to utilize ATP for self-interaction as actin requires ATP for polymerization (Bork et al. 1992; Pollard 2016). Difficulties in purifying active FtsA have made it problematic to study mechanistic detail of FtsA polymerization and its effects on division; however, the culmination of many studies have provided useful information on the topic. In vivo work using a Gfp-FtsA variant lacking the MTS, which tethers FtsA to the membrane, appeared to stabilize self-interaction and revealed large fluorescent rods forming in the cytoplasm (Pichoff and Lutkenhaus 2002). These fluorescent rods were hypothesized to be large polymers forming in the cytoplasm. After screening for mutations that eliminated rod formation, it was confirmed that mutations disrupting rod formation also disrupted self-interaction in a yeast two-hybrid assay (Pichoff et al. 2007). Several years after this, the FtsA dimer from *T. maritima* was observed alongside the

observation from electron microscopy that the protein forms linear polymers on lipid monolayers (Szwedziak et al. 2012). Additionally, purified FtsA from *Streptococcus pneumoniae* was shown to form ATP-dependent helical filaments in solution, which were stabilized by deletion of the MTS (Lara et al. 2005; Krupka et al. 2014). Most recently, purified FtsA from *E. coli* has been characterized for oligomerization. In several studies, FtsA was shown to form lipid-bound minirings on two dimensional lipid monolayers (Krupka et al. 2017; Nierhaus et al. 2022). These minirings appear to form spontaneously on the monolayer, without addition of exogenous ATP although some ADP/ATP is likely carried over from purification of FtsA (Krupka et al. 2017). However, miniring assembly in vivo has not been observed. Interestingly, a recent study revealed that addition of purified *E. coli* FtsN to FtsA minirings assembled on monolayers induced straight, antiparallel, double filament formation, and double filaments were observed via cross-linking experiments in vivo (Nierhaus et al. 2022). This indicates that one possible role of FtsN in the cell is to alter FtsA polymerization dynamics and increase lateral interactions, which could have implications on its interaction with FtsZ and the membrane in vivo. However, these minirings assemble without addition of endogenous ATP, so the role of ATP for FtsA function could not be determined in this study. A recent report shows that *E. coli* FtsA lacking its MTS forms stable linear polymers in solution in an ATP- and Mg^{2+} -dependent manner (Morrison et al. 2022). These polymers are antagonized by unassembled FtsZ suggesting that there may be coordination of FtsA and FtsZ oligomeric states depending on their nucleotide occupancy (Morrison et al. 2022). These results corroborate previous observations that deletion of the MTS stabilizes FtsA protofilaments and show that ATP may be critical for increasing self-interaction (Krupka et al. 2014; Pichoff and Lutkenhaus 2002).

What then is the purpose of FtsA self-interaction in division? One model that has dominated the field is that FtsA self-interaction blocks the 1C domain from interaction with other divisome proteins (Fig. 4.6) This stems from the observation that mutations in *ftsA* that prevent self-interaction via a two-hybrid assay and fluorescent rod formation in vivo also bypass the requirement for *zipA* (Geissler et al. 2003; Pichoff et al. 2013). Although FtsA, FtsEX, and ZipA have been shown to interact directly, evidence that ZipA or FtsEX directly inhibits FtsA oligomerization in vitro is lacking (Vega and Margolin 2019a, b; Du et al. 2016). In contrast, another study showed that mutations in the dimer interface prevented complementation of a temperature-sensitive strain in *B. subtilis*, indicating that polymerization by FtsA may be important in the cell (Szwedziak et al. 2012). Interestingly, FtsN has recently shown to promote FtsA switching from minirings to double filaments on a lipid monolayer when the cytoplasmic tail of FtsN is added (Nierhaus et al. 2022). Since FtsN activates PG synthesis and FtsA is reported to cause local, ATP-dependent, membrane reorganization, an attractive model is that FtsN interaction with FtsA may enhance local membrane remodeling events in addition to activating PG synthesis enzymes to backfill these regions to allow the cell to constrict (Fig. 4.8) (Conti et al. 2018; Liu et al. 2015).

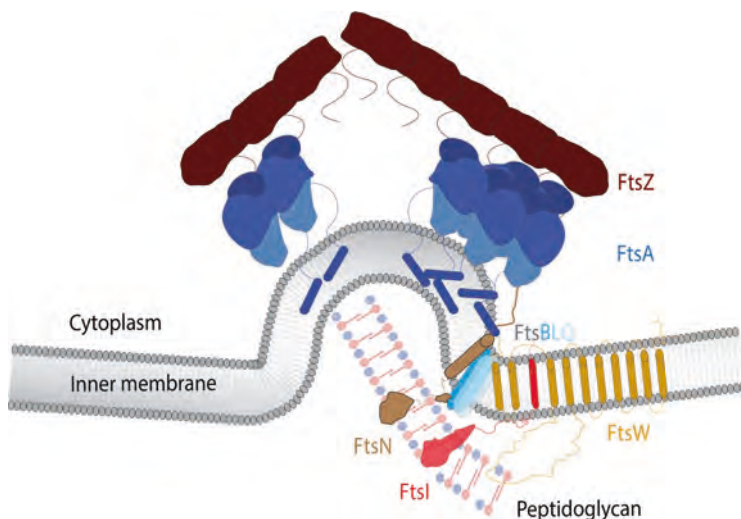


Fig. 4.8 Model for initial membrane remodeling to precede constriction. FtsA polymerization on the lipid surface induces local membrane distortions. FtsA interaction with FtsN may enhance FtsA filament formation and coordinate membrane remodeling with PG synthesis enzymes FtsWI to insert PG at these sites to promote the division event

Conclusion

In this chapter, we provided a brief overview of the function of the tubulin homolog FtsZ and the actin homolog FtsA in *E. coli*, a model Gram-negative bacterium, during cell division. FtsZ polymerizes to form the Z-ring at midcell, and its membrane tether FtsA anchors FtsZ to the membrane and recruits late-stage division proteins to the site to allow division to proceed. Polymerization and treadmilling by FtsZ may help distribute PG synthesis enzymes around the Z-ring. FtsZ levels are regulated transcriptionally by many factors including PdhR which may help link cell division and sensing of the cell metabolic state, and FtsZ is proteolytically degraded by the ClpXP complex. Polymerization by FtsZ is prevented at the cell poles by the Min system, over the nucleoid by SlmA, and during the SOS response by Sula. FtsZ polymers are laterally bundled and modulated by the Zap proteins, and FtsZ polymerization dynamics are also regulated by the membrane anchors ZipA and FtsA. FtsA tethers FtsZ polymers to the membrane of the cell and communicates with late-stage cell division proteins including FtsN to initiate PG hydrolysis and synthesis. FtsA polymerization is regulated by other division proteins such as ZipA and FtsN. Although the precise role of FtsA polymerization is unknown, a proposed model is that FtsA remodels lipid architecture directly through polymerization, which is regulated by other division proteins. Significant work, particularly biochemically, remains to fully elucidate how FtsA and FtsZ polymers coordinate division assembly and facilitate cell separation.

Acknowledgments Work in the Camberg lab is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM118927 (JLC) and the National Science Foundation Division of Molecular and Cellular Biosciences under Award Number 2131099 (JLC).

References

- Adams DW, Errington J (2009) Bacterial cell division: assembly, maintenance and disassembly of the Z ring. *Nat Rev Microbiol* 7(9):642–653
- Addinall SG, Erfei B, Lutkenhaus J (1996) FtsZ ring formation in Fts mutants. *J Bacteriol* 178(13):3877–3884
- Adler HI, Fisher WD, Cohen A, Hardigree AA (1967) Miniature *Escherichia coli* cells deficient in DNA. *Proc Natl Acad Sci* 57(2):321–326
- Aldea M, Garrido T, Pla J, Vicente M (1990) Division genes in *Escherichia coli* are expressed coordinately to cell septum requirements by gearbox promoters. *EMBO J* 9(11):3787–3794
- Ángel R-RM et al (2020) The nucleoid occlusion protein SlmA binds to lipid membranes. *MBio* 11(5):e02094–e02020
- Arends SJ, Kustusch RJ, Weiss DS (2009) ATP-binding site lesions in FtsE impair cell division. *J Bacteriol* 191(12):3772–3784
- Ayala JA, Garrido T, de Pedro MA, Vicente M (1994) Chapter 5 Molecular biology of bacterial septation. In: Ghuyssen J-M, Hakenbeck R (eds) *Bacterial cell wall*. Elsevier, Netherlands, pp 73–101
- Bailey MW et al (2014) Evidence for Divisome localization mechanisms independent of the min system and SlmA in *Escherichia coli*. *PLoS Genet* 10(8):e1004504
- Berezuk AM, Goodyear M, Khursigara CM (2014) Site-directed fluorescence labeling reveals a revised N-terminal membrane topology and functional periplasmic residues in the *Escherichia coli* cell division protein FtsK. *J Biol Chem* 289(34):23287–23301
- Berezuk AM et al (2020) FtsA G50E mutant suppresses the essential requirement for FtsK during bacterial cell division in *Escherichia coli*. *Can J Microbiol* 66(4):313–327
- Bernhardt TG, de Boer PAJ (2005) SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over chromosomes in *E. coli*. *Mol Cell* 18(5):555–564
- Bi E, Lutkenhaus J (1991) FtsZ ring structure associated with division in *Escherichia coli*. *Nature* 354(6349):161–164
- Bi E, Lutkenhaus J (1993) Cell division inhibitors SulA and MinCD prevent formation of the FtsZ ring. *J Bacteriol* 175(4):1118–1125
- Bisson-Filho AW et al (2017) Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. *Science* 355(6326):739–743
- Bohin J-p, Kennedy EP (1984) Mapping of a Locus (MdoA) that affects the biosynthesis of membrane-derived oligosaccharides in *Escherichia coli*. *J Bacteriol* 157(3):956–957
- Bork P, Sander C, Valencia A (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and Hsp70 heat shock proteins. *Proc Natl Acad Sci USA* 89(16):7290–7294. <https://pubmed.ncbi.nlm.nih.gov/1323828>
- Bramhill D, Thompson CM (1994) GTP-dependent polymerization of *Escherichia coli* FtsZ protein to form tubules. *Proc Natl Acad Sci USA* 91(13):5813–5817
- Busiek KK, Eraso JM, Wang Y, Margolin W (2012) The early divisome protein FtsA interacts directly through its 1c subdomain with the cytoplasmic domain of the late divisome protein FtsN. *J Bacteriol* 194(8):1989–2000. <https://pubmed.ncbi.nlm.nih.gov/22328664>
- Buss J et al (2013) In vivo organization of the FtsZ-ring by ZapA and ZapB revealed by quantitative super-resolution microscopy. *Mol Microbiol* 89(6):1099–1120
- Caldas P et al (2019) Cooperative ordering of treadmilling filaments in cytoskeletal networks of FtsZ and its crosslinker ZapA. *Nat Comm* 10(1)

- Camberg JL, Hoskins JR, Wickner S (2009) ClpXP protease degrades the cytoskeletal protein, FtsZ, and modulates FtsZ polymer dynamics. *Proc Natl Acad Sci* 106(26):10614 LP–10610619
- Camberg JL et al (2014) Location of dual sites in *E. coli* FtsZ important for degradation by ClpXP; one at the C-terminus and one in the disordered linker. *PLoS One* 9(4):e94964
- Carballido-López R (2006) The bacterial actin-like cytoskeleton. *Microbiol Mol Biol Rev*: MMBR 70(4):888–909
- Chen Y, Huang H, Osawa M, Erickson HP (2017) ZipA and FtsA* stabilize FtsZ-GDP minirings structures. *Sci Rep* 7(1):3650
- Chen Y, Milam SL, Erickson HP (2012) Sula inhibits assembly of FtsZ by a simple sequestration mechanism. *Biochemistry* 51(14):3100–3109
- Cho H, McManus HR, Dove SL, Bernhardt TG (2011) Nucleoid occlusion factor SlmA is a DNA-activated FtsZ polymerization antagonist. *Proc Natl Acad Sci USA* 108(9):3773–3778
- Chou SZ, Pollard TD (2019) Mechanism of actin polymerization revealed by Cryo-EM structures of actin filaments with three different bound nucleotides. *Proc Natl Acad Sci U S A* 116(10):4265–4274
- Coltharp C, Buss J, Plumer TM, Xiao J (2016) Defining the rate-limiting processes of bacterial cytokinesis. *Proc Natl Acad Sci USA* 113(8):E1044–E1053
- Conti J, Viola MG, Camberg JL (2015) The bacterial cell division regulators MinD and MinC form polymers in the presence of nucleotide. *FEBS Lett* 589(2):201–206
- Conti J, Viola MG, Camberg JL (2018) FtsA reshapes membrane architecture and remodels the Z-ring in *Escherichia coli*. *Mol Microbiol* 107(4):558–576
- Corbin BD, Geissler B, Sadasivam M, Margolin W (2004) Z-ring-independent interaction between a subdomain of FtsA and late septation proteins as revealed by a polar recruitment assay. *J Bacteriol* 186(22):7736–7744
- Corbin BD, Wang Y, Beuria TK, Margolin W (2007) Interaction between cell division proteins FtsE and FtsZ. *J Bacteriol* 189(8):3026–3035
- Dai K, Xu Y, Lutkenhaus J (1993) Cloning and characterization of FtsN, an essential cell division gene in *Escherichia coli* isolated as a multicopy suppressor of FtsA12(Ts). *J Bacteriol* 175(12):3790–3797
- Dai K, Yifan X, Lutkenhaus J (1996) Topological characterization of the essential *Escherichia coli* cell division protein FtsN. *J Bacteriol* 178(5):1328–1334
- de Boer PA, Crossley RE, Hand AR, Rothfield LI (1991) The MinD protein is a membrane ATPase required for the correct placement of the *Escherichia coli* division site. *EMBO J* 10(13):4371–4380
- de Boer PAJ, Crossley RE, Rothfield LI (1989) A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in *E. coli*. *Cell* 56(4):641–649
- de la Fuente A, Palacios P, Vicente M (2001) Transcription of the *Escherichia coli* Dcw cluster: evidence for distal upstream transcripts being involved in the expression of the downstream FtsZ gene. *Biochimie* 83(1):109–115
- DiBisio EC et al (2021) The stress-active cell division protein ZapE alters FtsZ filament architecture to facilitate division in *Escherichia coli*. *Front Microbiol* 12:2697
- Dorina T, Solomon S, Chris T, David B (1998) Bacterial SOS checkpoint protein Sula inhibits polymerization of purified FtsZ cell division protein. *J Bacteriol* 180(15):3946–3953
- Du S, Lutkenhaus J (2014) SlmA antagonism of FtsZ assembly employs a two-pronged mechanism like MinCD. *PLoS Genet* 10(7):e1004460
- Du S, Park K-T, Lutkenhaus J (2015) Oligomerization of FtsZ converts the FtsZ tail motif (conserved carboxy-terminal peptide) into a multivalent ligand with high avidity for partners ZipA and SlmA. *Mol Microbiol* 95(2):173–188
- Du S, Pichoff S, Kruse K, Lutkenhaus J (2018) FtsZ filaments have the opposite kinetic polarity of microtubules. *Proc Natl Acad Sci USA* 115(42):10768–10773
- Du S, Pichoff S, Lutkenhaus J (2016) FtsEX acts on FtsA to regulate divisome assembly and activity. *Proc Natl Acad Sci USA* 113(34):E5052–E5061

- Durand-Heredia J et al (2012) Identification of ZapD as a cell division factor that promotes the assembly of FtsZ in *Escherichia coli*. *J Bacteriol* 194(12):3189–3198
- Ebersbach G et al (2008) Novel coiled-coil cell division factor ZapB stimulates Z ring assembly and cell division. *Mol Microbiol* 68(3):720–735
- Espéll O et al (2012) A MatP-divisome interaction coordinates chromosome segregation with cell division in *E. coli*. *EMBO J* 31(14):3198–3211
- Feucht A, Lucet I, Yudkin MD, Errington J (2001) Cytological and biochemical characterization of the FtsA cell division protein of *Bacillus subtilis*. *Mol Microbiol* 40(1):115–125
- Flärdh K, Palacios P, Vicente M (1998) Cell division genes FtsQAZ in *Escherichia coli* require distant cis-acting signals upstream of DdIB for full expression. *Mol Microbiol* 30(2):305–315
- Flynn JM et al (2003) Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol Cell* 11(3):671–683
- Fu G et al (2010) In vivo structure of the *E. coli* FtsZ-ring revealed by photoactivated localization microscopy (PALM). *PLoS One* 5(9):e12682
- Fujita J et al (2017) Identification of the key interactions in structural transition pathway of FtsZ from *Staphylococcus aureus*. *J Struct Biol* 198(2):65–73
- Galli E, Gerdes K (2010) Spatial resolution of two bacterial cell division proteins: ZapA recruits ZapB to the inner face of the Z-ring. *Mol Microbiol* 76(6):1514–1526
- García-Soriano DA et al (2020) The speed of FtsZ treadmilling is tightly regulated by membrane binding. *Sci Rep* 10(1):10447
- Geissler B, Elraheb D, Margolin W (2003) A gain-of-function mutation in FtsA bypasses the requirement for the essential cell division gene ZipA in *Escherichia coli*. *Proc Natl Acad Sci USA* 113(34):E5052–E5061
- Geissler B, Margolin W (2005) Evidence for functional overlap among multiple bacterial cell division proteins: compensating for the loss of FtsK. *Mol Microbiol* 58(2):596–612
- Gervais FG, Phoenix P, Drapeau GR (1992) The RcsB gene, a positive regulator of colanic acid biosynthesis in *Escherichia coli*, is also an activator of FtsZ expression. *J Bacteriol* 174(12):3964–3971
- Ghosal D, Trambaiolo D, Amos LA, Löwe J (2014) MinCD cell division proteins form alternating copolymeric cytomotive filaments. *Nat Commun* 5:5341
- Goehring NW, Gonzalez MD, Beckwith J (2006) Premature targeting of cell division proteins to midcell reveals hierarchies of protein interactions involved in divisome assembly. *Mol Microbiol* 61(1):33–45
- Göhler A-K et al (2011) More than just a metabolic regulator—elucidation and validation of new targets of PdhR in *Escherichia coli*. *BMC Syst Biol* 5:197
- Gregory JA, Becker EC, Pogliano K (2008) *Bacillus subtilis* MinC destabilizes FtsZ-rings at new cell poles and contributes to the timing of cell division. *Genes Dev* 22(24):3475–3488
- Grimaud R et al (1998) Enzymatic and structural similarities between the *Escherichia coli* ATP-dependent proteases, ClpXP and ClpAP. *J Biol Chem* 273(20):12476–12481
- Gueiros-Filho FJ, Losick R (2002) A widely conserved bacterial cell division protein that promotes assembly of the tubulin-like protein FtsZ. *Genes Dev* 16(19):2544–2556
- Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* 30(Suppl 1):S162–S173
- Hale CA, De Boer PAJ (1997) Direct binding of FtsZ to ZipA, an essential component of the septal ring structure that mediates cell division in *E. coli*. *Cell* 88(2):175–185
- Hale CA, Rhee AC, de Boer PA (2000) ZipA-induced bundling of FtsZ polymers mediated by an interaction between C-terminal domains. *J Bacteriol* 182(18):5153–5166
- Hale CA et al (2011) Identification of *Escherichia coli* ZapC (YcbW) as a component of the division apparatus that binds and bundles FtsZ polymers. *J Bacteriol* 193(6):1393–1404
- Haney SA et al (2001) Genetic analysis of the *Escherichia coli* FtsZ-ZipA interaction in the yeast two-hybrid system: characterization of FtsZ residues essential for the interactions with zipA and with FtsA*210. *J Biol Chem* 276(15):11980–11987

- Hanson PI, Whiteheart SW (2005) AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* 6(7):519–529
- Helge F, Laeschkir W, Erwin F, Marc B (2021) Dynamics of the *Bacillus subtilis* Min system. *MBio* 12(2):e00296–e00221
- Herricks JR, Nguyen D, Margolin W (2014) A thermosensitive defect in the ATP binding pocket of FtsA can be suppressed by allosteric changes in the dimer Interface. *Mol Microbiol* 94(3):713–727
- Hill NS, Buske PJ, Shi Y, Levin PA (2013) A moonlighting enzyme links *Escherichia coli* cell size with central metabolism. *PLoS Genet* 9(7):e1003663
- Housman M et al (2016) FtsZ protofilament curvature is the opposite of tubulin rings. *Biochemistry* 55(29):4085–4091
- Hu Z, Lutkenhaus J (1999) Topological regulation of cell division in *Escherichia coli* involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE. *Mol Microbiol* 34(1):82–90
- Hu Z, Mukherjee A, Pichoff S, Lutkenhaus J (1999) The MinC component of the division site selection system in *Escherichia coli* interacts with FtsZ to prevent polymerization. *Proc Natl Acad Sci USA* 96(26):14819–14824
- Hu Z et al (2001) Topological regulation of cell division in *E. coli*: spatiotemporal oscillation of MinD requires stimulation of its ATPase by MinE and phospholipid. *Mol Cell* 7(6):1337–1343
- Kabsch W, Holmes KC (1995) The actin fold. *FASEB J* 9(2):167–174
- Kabsch W et al (1990) Atomic structure of the actin:DNase I complex. *Nature* 347(6288):37–44
- Kang H et al (2012) Identification of cation-binding sites on actin that drive polymerization and modulate bending stiffness. *Proc Natl Acad Sci USA* 109(42):16923–16927
- Khattar MM, Begg KJ, Donachie WD (1994) Identification of FtsW and characterization of a new FtsW division mutant of *Escherichia coli*. *J Bacteriol* 176(23):7140–7147
- Kotte O, Zaugg JB, Heinemann M (2010) Bacterial adaptation through distributed sensing of metabolic fluxes. *Mol Syst Biol* 6:355
- Krupka M, Cabré EJ, Jiménez M, Rivas G, Rico AI, Vicente M (2014) Role of the FtsA C terminus as a switch for polymerization and membrane association. *MBio* 5(6):e02221
- Krupka M et al (2017) *Escherichia coli* FtsA forms lipid-bound minirings that antagonize lateral interactions between FtsZ protofilaments. *Nat Commun* 8:15957
- Krupka M et al (2018) *Escherichia coli* ZipA organizes FtsZ polymers into dynamic ring-like protofilament structures. *MBio* 9(3):e01008–e01018
- Kuchibhatla A, Bhattacharya A, Panda D (2011) ZipA binds to FtsZ with high affinity and enhances the stability of FtsZ protofilaments. *PLoS One* 6(12):e28262–e28262
- LaBreck CJ, Conti J, Viola MG, Camberg JL (2018) MinC N- and C-domain interactions modulate FtsZ assembly, division site selection and MinD-dependent oscillation in *Escherichia coli*. *J Bacteriol* 201:00374–00318
- Lara B et al (2005) Cell division in cocci: localization and properties of the streptococcus pneumoniae FtsA protein. *Mol Microbiol* 55(3):699–711
- Liu G, Craig Draper G, Donachie WD (1998) FtsK is a bifunctional protein involved in cell division and chromosome localization in *Escherichia coli*. *Mol Microbiol* 29(3):893–903
- Liu B, Persons L, Lee L, de Boer PAJ (2015) Roles for both FtsA and the FtsBLQ subcomplex in FtsN-stimulated cell constriction in *Escherichia coli*. *Mol Microbiol* 95(6):945–970
- Loose M, Mitchison TJ (2014) The bacterial cell division proteins FtsA and FtsZ self-organize into dynamic cytoskeletal patterns. *Nat Cell Biol* 16(1):38–46
- Low HH, Moncrieffe MC, Löwe J (2004) The crystal structure of ZapA and its modulation of FtsZ polymerisation. *J Mol Biol* 341(3):839–852
- Lu C, Erickson HP (1999) The straight and curved conformation of FtsZ protofilaments-evidence for rapid exchange of GTP into the curved protofilament. *Cell Struct Funct* 24(5):285–290
- Lu C, Reedy M, Erickson HP (2000) Straight and curved conformations of FtsZ are regulated by GTP hydrolysis. *J Bacteriol* 182(1):164–170
- Lutkenhaus J (2009) FtsN—trigger for septation. *J Bacteriol* 191(24):7381–7382

- Lyu Z et al (2022) FtsN maintains active septal cell wall synthesis by forming a processive complex with the septum-specific peptidoglycan synthases in *E. coli*. *Nat Commun* 13(1):5751
- Männik J, Walker BE, Männik J (2018) Cell cycle-dependent regulation of FtsZ in *Escherichia coli* in slow growth conditions. *Mol Microbiol* 110(6):1030–1044
- Marmont LS, Bernhardt TG (2020) A conserved subcomplex within the bacterial cytokinetic ring activates cell wall synthesis by the FtsW-FtsI synthase. *Proc Natl Acad Sci USA* 117(38):23879–23885
- Marston AL, Errington J (1999) Selection of the midcell division site in *Bacillus subtilis* through MinD-dependent polar localization and activation of MinC. *Mol Microbiol* 33(1):84–96
- Marston AL et al (1998) Polar localization of the MinD protein of *Bacillus subtilis* and its role in selection of the mid-cell division site. *Genes Dev* 12(21):3419–3430
- Marteyn B et al (2014) ZapE is a novel cell division protein interacting with FtsZ and modulating the Z-ring dynamics. *MBio* 5(2):e00022–e00014
- Martos A et al (2012) Isolation, characterization and lipid-binding properties of the recalcitrant FtsA division protein from *Escherichia coli*. *PLoS One* 7(6):e39829–e39829
- Matsui T et al (2014) Structural change in FtsZ induced by intermolecular interactions between bound GTP and the T7 loop. *J Biol Chem* 289(6):3501–3509
- Maurizi MR et al (1990) Sequence and structure of Clp P, the proteolytic component of the ATP-dependent Clp protease of *Escherichia coli*. *J Biol Chem* 265(21):12536–12545
- McCausland JW et al (2021) Treadmilling FtsZ polymers drive the directional movement of SPG-synthesis enzymes via a Brownian ratchet mechanism. *Nat Commun* 12(1):609
- McQuillen R, Xiao J (2020) Insights into the structure, function, and dynamics of the bacterial cytokinetic FtsZ-ring. *Annu Rev Biophys* 49:309–341
- Mercier R et al (2008) The MatP/MatS site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* 135(3):475–485
- Merino F et al (2018) Structural transitions of F-actin upon ATP hydrolysis at near-atomic resolution revealed by cryo-EM. *Nat Struct Mol Biol* 25(6):528–537
- Morrison JJ, Conti J, Camberg JL (2022) Assembly and architecture of *Escherichia coli* divisome proteins FtsA and FtsZ. *J Biol Chem* 298(3):101663
- Mosyak L et al (2000) The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *EMBO J* 19(13):3179–3191
- Mukherjee A, Cao C, Lutkenhaus J (1998) Inhibition of FtsZ polymerization by Sula, an inhibitor of septation in *Escherichia coli*. *Proc Natl Acad Sci USA* 95(6):2885–2890
- Narita A (2011) Merits of the double-stranded form of the actin filament revealed by structures of the filament ends. *Commun Integr Biol* 4(6):692–695
- Nierhaus T et al (2022) Bacterial divisome protein FtsA forms curved antiparallel double filaments when binding to FtsN. *Nat Microbiol* 7(10):1686–1701
- Oliva MA, Cordell SC, Löwe J (2004) Structural insights into FtsZ protofilament formation. *Nat Struct Mol Biol* 11(12):1243–1250
- Osawa M, Anderson DE, Erickson HP (2008) Reconstitution of contractile FtsZ rings in liposomes. *Science* 320(5877):792–794. <https://doi.org/10.1126/science.1154520>
- Osawa M et al (2009) Curved FtsZ protofilaments generate bending forces on liposome membranes. *EMBO J* 28(22):3476–3484
- Pacheco-Gómez R, Roper DI, Dafforn TR, Rodger A (2011) The PH dependence of polymerization and bundling by the essential bacterial cytoskeletal protein FtsZ. *PLoS One* 6(6):e19369
- Pacheco-Gómez R et al (2013) Tetramerization of ZapA is required for FtsZ bundling. *Biochem J* 449(3):795–802
- Park K-T, Pichoff S, Shishen D, Lutkenhaus J (2021) FtsA acts through FtsW to promote cell wall synthesis during cell division in *Escherichia coli*. *Proc Natl Acad Sci USA* 118(35)
- Park K-T, Shishen D, Lutkenhaus J (2015) MinC/MinD copolymers are not required for min function. *Mol Microbiol* 98(5):895–909
- Park K-T et al (2018) MinC and FtsZ mutant analysis provides insight into MinC/MinD-mediated Z ring disassembly. *J Biol Chem* 293(16):5834–5846

- Park K-T et al (2020) Essential role for FtsL in activation of septal peptidoglycan synthesis. *MBio* 11(6)
- Pichoff S, Lutkenhaus J (2002) Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. *EMBO J* 21(4):685–693
- Pichoff S, Shen B, Sullivan B, Lutkenhaus J (2013) FtsA mutants impaired for self-interaction bypass ZipA suggesting a model in which FtsA's self-interaction competes with its ability to recruit downstream division proteins. *Mol Microbiol* 83(1):151–167
- Pichoff S et al (2005) Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA. *Mol Microbiol* 55(6):1722–1734
- Pichoff S et al (2007) Identification of a region of FtsA required for interaction with FtsZ. *Mol Microbiol* 64(4):1129–1138
- Pollard TD (1986) Rate constants for the reactions of ATP- and ADP-actin with the ends of actin filaments. *J Cell Biol* 103(6 Pt 2):2747–2754
- Pollard TD (2016) Actin and actin-binding proteins. *Cold Spring Harb Perspect Biol* 8(8)
- Quail MA, Haydon DJ, Guest JR (1994) The PdhR-AceEF-Lpd operon of *Escherichia coli* expresses the pyruvate dehydrogenase complex. *Mol Microbiol* 12(1):95–104
- Ramírez-Díaz DA et al (2018) Treadmilling analysis reveals new insights into dynamic FtsZ ring architecture. *PLoS Biol* 16(5):e2004845–e2004845
- Raskin DM, de Boer PAJ (1997) The MinE ring: an FtsZ-independent cell structure required for selection of the correct division site in *E. coli*. *Cell* 91(5):685–694
- Raskin DM, de Boer PA (1999) Rapid pole-to-pole oscillation of a protein required for directing division to the middle of *Escherichia coli*. *Proc Natl Acad Sci USA* 96(9):4971–4976
- RayChaudhuri D (1999) ZipA is a MAP-tau homolog and is essential for structural integrity of the cytokinetic FtsZ ring during bacterial cell division. *EMBO J* 18(9):2372–2383
- Reddy M (2007) Role of FtsEX in cell division of *Escherichia coli*: viability of FtsEX mutants is dependent on functional SufI or high osmotic strength. *J Bacteriol* 189(1):98–108
- Rivas G et al (2000) Magnesium-induced linear self-association of the FtsZ bacterial cell division protein monomer: the primary steps for FtsZ assembly. *J Biol Chem* 275(16):11740–11749
- Rosenblatt J, Peluso P, Mitchison TJ (1995) The bulk of unpolymerized actin in *Xenopus* egg extracts is ATP-bound. *Mol Biol Cell* 6(2):227–236
- Rueda S, Vicente M, Mingorance J (2003) Concentration and assembly of the division ring proteins FtsZ, FtsA, and ZipA during the *Escherichia coli* cell cycle. *J Bacteriol* 185(11):3344–3351
- Schmidt KL et al (2004) A predicted ABC transporter, FtsEX, is needed for cell division in *Escherichia coli*. *J Bacteriol* 186(3):785–793
- Schumacher MA, Huang K-H, Zeng W, Janakiraman A (2017) Structure of the Z ring-associated protein, ZapD, bound to the C-terminal domain of the tubulin-like protein, FtsZ, suggests mechanism of Z ring stabilization through FtsZ cross-linking. *J Biol Chem* 292(9):3740–3750
- Schumacher MA et al (2016) Structural and functional analyses reveal insights into the molecular properties of the *Escherichia coli* Z ring stabilizing protein, ZapC*. *J Biol Chem* 291(5):2485–2498
- Sekar K et al (2018) Synthesis and degradation of FtsZ quantitatively predict the first cell division in starved bacteria. *Mol Syst Biol* 14(11):e8623
- Sept D, McCammon JA (2001) Thermodynamics and kinetics of actin filament nucleation. *Biophys J* 81(2):667–674
- Shaevitz JW, Gitai Z (2010) The structure and function of bacterial actin homologs. *Cold Spring Harb Perspect Biol* 2(9):a000364–a000364
- Söderström B et al (2014) Disassembly of the divisome in *Escherichia coli*: evidence that FtsZ dissociates before compartmentalization. *Mol Microbiol* 92(1):1–9
- Söderström B et al (2016) Coordinated disassembly of the divisome complex in *Escherichia coli*. *Mol Microbiol* 101(3):425–438
- Stoddard PR et al (2017) Evolution of polymer formation within the actin superfamily. *Mol Biol Cell* 28(19):2461–2469
- Stouf M, Meile J-C, Cornet F (2013) FtsK actively segregates sister chromosomes in *Escherichia coli*. *Proc Natl Acad Sci USA* 110(27):11157–11162

- Studer G et al (2021) ProMod3—A versatile homology modelling toolbox. *PLoS Comput Biol* 17(1):e1008667
- Szwedziak P, Wang Q, Freund SMV, Löwe J (2012) FtsA forms actin-like protofilaments. *EMBO J* 31(10):2249–2260
- Tormo A, Vicente M (1984) The FtsA gene product participates in formation of the *Escherichia coli* septum structure. *J Bacteriol* 157(3):779–784
- van den Ent F, Löwe J (2000) Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. *EMBO J* 19(20):5300–5307. <https://pubmed.ncbi.nlm.nih.gov/11032797>
- Vaughan S, Wickstead B, Gull K, Addinall SG (2004) Molecular evolution of FtsZ protein sequences encoded within the genomes of archaea, bacteria, and eukaryota. *J Mol Evol* 58(1):19–29
- Vega DE, Margolin W (2019a) Direct interaction between the two Z ring membrane. *J Bacteriol* 201(4):e00579–e00518
- Vega DE, Margolin W (2019b) Direct interaction between the two Z ring membrane anchors FtsA and ZipA. *J Bacteriol* 201(4)
- Vicente M, Errington J (1996) Structure, function and controls in microbial division. *Mol Microbiol* 20(1):1–7
- Vicente M, Gomez MJ, Ayala JA (1998) Regulation of transcription of cell division genes in the *Escherichia coli* Dcw cluster. *Cell Mol Life Sci* 54(4):317–324
- Viola MG, LaBreck CJ, Conti J, Camberg JL (2017) Proteolysis-dependent remodeling of the tubulin homolog FtsZ at the division septum in *Escherichia coli*. *PLoS One* 12(1):e0170505–e0170505
- Vorobiev S et al (2003) The structure of nonvertebrate actin: implications for the ATP hydrolytic mechanism. *Proc Natl Acad Sci* 100(10):5760–5765
- Wagstaff JM et al (2017) A polymerization-associated structural switch in FtsZ that enables treadmilling of model filaments. *MBio* 8(3):e00254–e00217
- Wang L, Lutkenhaus J (1998) FtsK is an essential cell division protein that is localized to the septum and induced as part of the SOS response. *Mol Microbiol* 29(3):731–740
- Weiss DS et al (1997) Localization of the *Escherichia coli* cell division protein FtsI (PBP3) to the division site and cell pole. *Mol Microbiol* 25(4):671–681
- Weissborn AC, Kennedy P (1984) Biosynthesis of membrane-derived oligosaccharides. *J Biol Chem* 259(20):12644–12651
- Wojtkowiak D, Georgopoulos C, Zylicz M (1993) Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of *Escherichia coli*. *J Biol Chem* 268(30):22609–22617
- Yamamoto K, Yata K, Fujita N, Ishihama A (2001) Novel mode of transcription regulation by SdiA, an *Escherichia coli* homologue of the quorum-sensing regulator. *Mol Microbiol* 41(5):1187–1198. <https://doi.org/10.1046/j.1365-2958.2001.02585.x>
- Yang X et al (2017) GTPase activity-coupled treadmilling of the bacterial tubulin FtsZ organizes septal cell wall synthesis. *Science* 355(6326):744–747
- Yang X et al (2021) A two-track model for the spatiotemporal coordination of bacterial septal cell wall synthesis revealed by single-molecule imaging of FtsW. *Nat Microbiol* 6(5):584–593
- Yim A, Vandenbussche GUY, Casanova M, Ruyschaert J-m (2000) Role of the carboxy terminus of *Escherichia coli* FtsA in self-interaction and cell division. *J Bacteriol* 182(22):6366–6373
- Yoshizawa T et al (2020) Crystal structures of the cell-division protein FtsZ from *Klebsiella pneumoniae* and *Escherichia coli*. *Acta Crystallogr F Struct Biol Commun* 76(Pt 2):86–93
- Yu AYH, Houry WA (2007) ClpP: a distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett* 581(19):3749–3757
- Yu Y et al (2020) The min system disassembles FtsZ foci and inhibits polar peptidoglycan remodeling in *Bacillus subtilis*. *MBio* 11(2):e03197–e03119

Chapter 5

Telomere C-Strand Fill-In Machinery: New Insights into the Human CST-DNA Polymerase Alpha-Primase Structures and Functions



Ci Ji Lim

Abstract Telomeres at the end of eukaryotic chromosomes are extended by a specialized set of enzymes and telomere-associated proteins, collectively termed here the telomere “replisome.” The telomere replisome acts on a unique replicon at each chromosomal end of the telomeres, the 3' DNA overhang. This telomere replication process is distinct from the replisome mechanism deployed to duplicate the human genome. The G-rich overhang is first extended before the complementary C-strand is filled in. This overhang is extended by telomerase, a specialized ribonucleoprotein and reverse transcriptase. The overhang extension process is terminated when telomerase is displaced by CTC1-STN1-TEN1 (CST), a single-stranded DNA-binding protein complex. CST then recruits DNA polymerase α -primase to complete the telomere replication process by filling in the complementary C-strand. In this chapter, the recent structure-function insights into the human telomere C-strand fill-in machinery (DNA polymerase α -primase and CST) will be discussed.

Keywords Human telomere synthesis · C-strand fill-in · DNA polymerase α -primase · CST

Abbreviations

ACD/TPP1	Adrenocortical dysplasia homolog
Cryo-EM	Cryogenic-electron microscopy
CTC1	Conserved telomere maintenance component 1
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded deoxyribonucleic acid

C. J. Lim (✉)

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, WI, USA

e-mail: ciji.lim@wisc.edu

E	Enzyme
GQ	G-quadruplexes
NTP	Nucleotide triphosphate
OB	Oligonucleotide/oligosaccharide-binding
PIC	Preinitiation complex
POLA1	DNA polymerase alpha catalytic subunit
POLA2	DNA polymerase alpha subunit B
POT1	Protection of telomere 1
PRIM1	DNA primase small subunit
PRIM2	DNA primase large subunit
RAP1	Repressor/activator site-binding protein
RC	Recruitment complex
RNA	Ribonucleic acid
RPA	Replication protein A
ssDNA	Single-stranded deoxyribonucleic acid
STN1/OBFC1	Suppressor of cdc thirteen homolog
TEN1	Telomere length regulation protein
TIN2	TERF1-interacting nuclear factor 2
TRF1	Telomere repeat-binding factor 1
TRF2	Telomere repeat-binding factor 2

Introduction

Telomeres are specialized protein-DNA structures located at the termini of eukaryotic chromosomes (Moyzis et al. 1988; Riethman 2008). They play a vital role in safeguarding the genetic information's integrity (O'Sullivan and Karlseder 2010). In the absence of telomeric protection, there is a risk of losing genetic material, and the linear chromosome ends might be erroneously identified as broken DNA segments. One defining characteristic of telomeres is their physical length. Comprising repetitive DNA sequences spanning several kilobases, telomeres terminate in a 3'-DNA overhang (Makarov et al. 1997; McElligott and Wellinger 1997; Wright et al. 1997). The length of this telomeric DNA is influenced by two opposing mechanisms: the natural attrition of telomeric DNA, a consequence of the end-replication problem inherent to linear DNA molecules (Olovnikov 1996; Harley 1991), and the extension of telomeric DNA, facilitated by a dedicated ensemble of replicative enzymes and telomere-associated proteins (Lim and Cech 2021; Bonnell et al. 2021).

Telomeres undergo progressive shortening with each cell division, and when they reach a critically short length, there is cellular senescence (Harley et al. 1990; Smogorzewska and de Lange 2002). This phenomenon is responsible for limiting the replicative capacity of mammalian cells and is referred to as Hayflick's limit (Shay and Wright 2000; Hayflick and Moorhead 1961; Hayflick 1965). In contrast, cells with excessively long telomeres have an extended replicative lifespan. Notably, cancer cells often manipulate the telomere extension process to support their

aggressive proliferation (Maciejowski and de Lange 2017; Jafri et al. 2016; Roake and Artandi 2020). Consequently, the regulation of telomere length has emerged as a pivotal area of investigation in both fundamental biology and oncology.

The intricate process of telomeric DNA synthesis is pivotal to regulating telomere length. This synthesis is broadly classified into two distinct phases (Lim and Cech 2021; Cai and de Lange 2023; He and Lim 2023): the extension of the telomere overhang and the complementary C-strand fill-in. In this book chapter, we aim to shed light on the latest insights and breakthroughs pertaining to the human telomere C-strand fill-in mechanism. Our journey begins with a concise overview of telomeric DNA synthesis, laying down the foundational knowledge to understand the integral function of the human telomere C-strand fill-in apparatus. We will then delve into the latest structure-function studies of the individual components of this machinery, namely, DNA polymerase α -primase and CTC1-STN1-TEN1 (CST), before discussing how these two protein complexes collaborate to perform telomere C-strand fill-in.

An Overview of Human Telomeric DNA Synthesis

Telomere extension (telomeric DNA synthesis) encompasses two distinct, sequential processes (Fig. 5.1). First, the 3'-DNA overhang of the telomere is extended by telomerase, a reverse transcriptase (Roake and Artandi 2020; Nguyen 2021; He and Feigon 2022; Morin 1989); telomerase uses its noncoding RNA template to sequentially add telomeric repeats to the overhang, a characteristic known as repeat addition processivity (RAP). In mammals, the highly conserved telomeric DNA repeat sequence is TTAGGG (Moyzis et al. 1988). As a result of guanine repeats, the telomere overhang is often referred to as the G-overhang. At the late S-phase, telomerase engagement to the overhang is challenged by the upregulated abundance of CTC1-STN1-TEN1 (CST) (Chen et al. 2012), a heterotrimeric protein complex that has binding specificity to G-rich single-strand DNA sequences (Miyake et al. 2009; Surovtseva et al. 2009; Barbour and Wuttke 2023).

The second half of the telomere synthesis process commences at the late S/G2-phase when DNA polymerase α -primase (pol α -primase) starts synthesizing the complementary C-strand, using the G-overhang newly produced by telomerase as its DNA template (Zhao et al. 2009; Chow et al. 2012; Takai et al. 2024). Because there is no preceding DNA strand for pol α -primase's DNA polymerase to use and perform the C-strand fill-in, pol α -primase uses its primase to de novo synthesize a short RNA primer and then uses its DNA polymerase to extend the RNA primer with DNA, completing telomeric DNA synthesis. Thus, contrary to what is commonly depicted, telomeric DNA synthesis is not exclusively carried out by telomerase. Instead, it is a collective effort between telomerase and pol α -primase (He and Lim 2023; Cai and de Lange 2023).

Telomere-associated proteins play a pivotal role in telomeric DNA synthesis (Lim and Cech 2021; Smith et al. 2020). Notable among these are the shelterin (de

Lange 2005) and CST (Lim and Cech 2021) protein complexes. Both complexes are instrumental in the recruitment, assembly, and regulatory functions of the aforementioned replicative enzymes (Schmidt and Cech 2015; Lim and Cech 2021; He and Lim 2023; Cai and de Lange 2023). The shelterin complex comprises six distinct subunits: TRF1, TRF2, RAP1, TPP1, TIN2, and POT1 (de Lange 2005; Lim and Cech 2021; Smith et al. 2020). Half of these proteins exhibit strong binding affinity for telomeric DNA. Specifically, the homodimers TRF1 and TRF2 associate with telomeric double-stranded DNA (dsDNA) (Bianchi et al. 1997; Chong et al. 1995; Broccoli et al. 1997), while POT1 interacts with the single-stranded telomeric G-overhang (Baumann and Cech 2001; Loayza and De Lange 2003; Lei et al. 2004). These specific protein-DNA interactions underpin the assembly and functionality of the shelterin complex at telomeres.

A primary function of shelterin is to facilitate the recruitment of telomerase to telomeres, enabling the extension of the telomeric G-overhang. Central to this function is the interaction between TPP1 and telomerase (Zhong et al. 2012; Sexton et al. 2012; Nandakumar et al. 2012; Schmidt et al. 2016; Schmidt et al. 2014), a conserved association also observed in ciliates (Chan et al. 2017; Jiang et al. 2015; Wu et al. 2017). When acting independently, POT1 caps the 3'-DNA overhang (Lei et al. 2004), inherently inhibiting telomerase from lengthening the G-overhang (Lei et al. 2005). It was recently discovered that POT1 also binds to the recessed 5'-phosphate group at the telomeric ss-dsDNA junction, an action important for POT1 chromosome end protection function (Tesmer et al. 2023). In conjunction with TPP1, the inhibitory role of POT1 reverses, the resulting TPP1-POT1 binary complex instead becomes an enhancer of telomerase repeat addition processivity (Wang et al. 2007). While the RAP stimulatory mechanism of TPP1-POT1 has been thoroughly delineated in biochemical studies (Latrick and Cech 2010; Dalby et al. 2015; Jansson et al. 2019), the physiological significance of the RAP mechanism remains untested. In stark contrast, the role of TPP1 in human telomerase recruitment in cells is firmly established (Schmidt and Cech 2015; Schmidt et al. 2018; Schmidt et al. 2016; Laprade et al. 2020).

The CST protein complex plays two important roles in telomeric DNA synthesis: it not only terminates telomerase extension of the G-overhang but also aids pol α -primase in the C-strand fill-in of the newly formed overhang. Delving into its first role, CST terminates telomerase extension of the G-overhang by competing with the telomerase for the telomeric G-overhang, a process that happens at the late S-phase (Chen et al. 2012; Zhao et al. 2009; Takai et al. 2024). CST's ability to displace telomerase is attributed to its affinity for G-rich ssDNA (Chen et al. 2012; Zaug et al. 2021). However, the dynamics of CST and pol α -primase converging on G-overhangs remain ambiguous (He and Lim 2023). Do they operate sequentially, or do they arrive as a pre-assembled unit? Cell-based imaging investigations have shed some light on this, revealing that CST's nuclear import is cell cycle-regulated and hinges on its cytoplasmic assembly with pol α -primase (Kelich et al. 2021). This suggests that CST might rival telomerase for the G-overhang while being part of a CST-pol α -primase ensemble. But CST's role is not limited to mere recruitment; as

mentioned above, it also stimulates the enzymatic activity of pol α -primase, promoting telomeric C-strand fill-in (Zaug et al. 2022; He et al. 2022a).

Hence, to initiate and successfully complete telomeric DNA synthesis, mammalian cells require four elemental components: telomerase, pol α -primase, shelterin, and CST. In tandem, these protein complexes form the telomere “replisome,” orchestrating the biphasic synthesis of telomeric DNA, first the G-strand and subsequently the C-strand. Intriguingly, the *in vitro* reconstitution of a minimally active human telomere replisome can be achieved without the inclusion of shelterin proteins (Zaug et al. 2022). This is attributed to telomerase’s innate ability to locate telomeric DNA primers *in vitro*, independent of shelterin assistance.

While telomerase’s extension of the telomeric G-overhang is well-understood, the intricacies surrounding the C-strand fill-in “replicon” and the assembly and activation of CST-pol α -primase at the telomeric G-overhang remain elusive. Breakthroughs in cryogenic-electron microscopy technology have helped unveil the molecular architecture of both CST and CST-pol α -primase (Lim et al. 2020; Cai et al. 2022; He et al. 2022a, b; Cai et al. 2023). These revelations are bolstered by concurrent biochemical (Zaug et al. 2021; Zaug et al. 2022) and cellular studies (Kelich et al. 2021; Wang et al. 2023), painting a more comprehensive picture of the coordination between CST and pol α -primase during the telomere C-strand fill-in process. To better grasp these recent CST-pol α -primase findings, we will first review the current understanding of pol α -primase as a central replicative enzyme and then delve into how CST, as an accessory factor, help the enzyme perform its telomere-C-strand fill-in.

DNA Polymerase α -Primase: A Dynamic Primer Synthesis Molecular Machine

DNA polymerase alpha (pol α) collaborates with primase to form a unified protein complex known as pol α -primase. This complex plays a pivotal role in initiating DNA synthesis during various replication processes, including those at the telomeres (Pellegrini 2012; Baranovskiy and Tahirov 2017; He and Lim 2023). The pol α -primase complex comprises four subunits: DNA polymerase alpha (also known as p180 or POLA1), DNA polymerase alpha subunit B (referred to as p70 or POLA2), DNA primase small subunit (p49 or PRIM1), and DNA primase large subunit (p58 or PRIM2) (Fig. 5.2a). This heterotetrameric assembly guarantees the coordinated and efficient production of chimeric RNA-DNA primers (Pellegrini 2012; Baranovskiy and Tahirov 2017), which are essential for initiating DNA replication.

PRIM1 is known to catalyze the formation of a phosphodiester bond between ribonucleotides (Nasheuer and Grosse 1988). However, when functioning independently, it lacks the capability to initiate *de novo* RNA synthesis (Copeland and Wang 1993). It needs to partner with PRIM2 to form the heterodimeric primase unit to synthesize RNA *de novo* (Copeland and Wang 1993; Santocanale et al. 1993). The primase mechanism is biochemically well-characterized. It first binds a single-stranded DNA template (E•DNA complex) and then slides along the template to

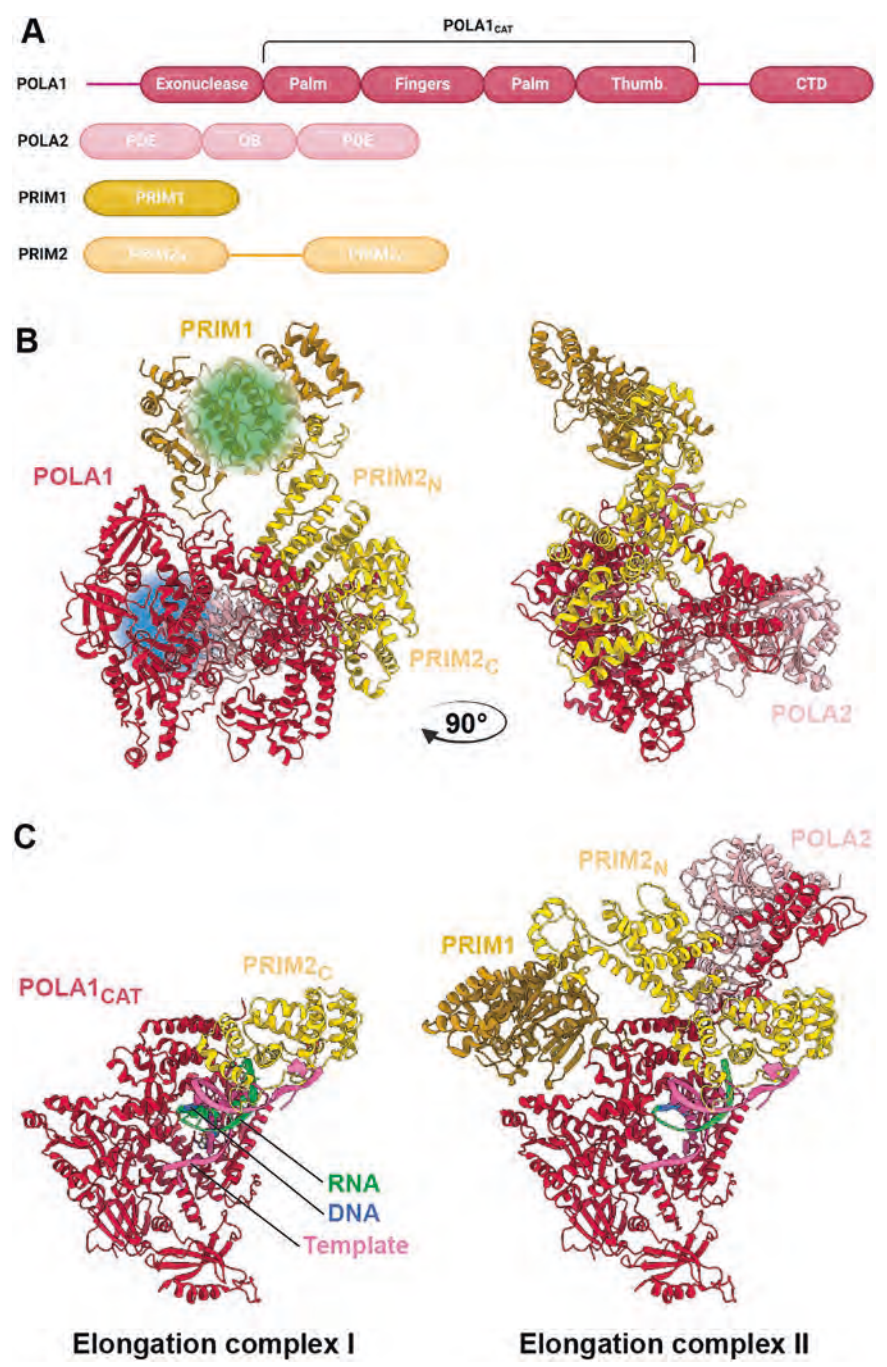


Fig. 5.2 Domain organization and structural overview of the human pol α -primase complex. (a) This panel depicts the domain organization of the human pol α -primase complex, highlighting the catalytic subunit of DNA polymerase alpha (POLA1), the accessory subunit (POLA2), and the two

find a suitable start site to initiate de novo RNA primer synthesis (Sheaff and Kuchta 1993). Notably, mammalian primase exhibits a preference for initiating RNA primer synthesis with purines (Sheaff and Kuchta 1993), indicating a certain level of specificity in its start site selection.

The primase possesses two distinct nucleotide-binding sites: the initiation and elongation sites (Kilkenny et al. 2013; Frick and Richardson 2001). For the formation of the initial dinucleotide, a crucial rate-limiting step in de novo RNA primer synthesis, nucleotides must occupy both sites in a specific sequence (Sheaff and Kuchta 1993). The process begins with the E•DNA complex binding to the second nucleotide first, resulting in the E•DNA•NTP complex. Following this, it binds to the 5′-terminal nucleotide, leading to the formation of the E•DNA complex•NTP•NTP complex (Sheaff and Kuchta 1993).

Numerous structures of the NTP-bound human PRIM1 have been determined (Kilkenny et al. 2013; Vaithiyalingam et al. 2014; Holzer et al. 2019). These structures indicate that the nucleotide-binding pocket observed in PRIM1 likely serves as the elongation site, where the catalysis of phosphodiester bonds takes place (Kilkenny et al. 2013). The role and structure of the initiation nucleotide-binding site remain more ambiguous. However, it is understood that this site involves the C-terminal domain of PRIM2, known as PRIM2_C (Zerbe and Kuchta 2002; Baranovskiy et al. 2016b). This site does not partake in catalysis; instead, its function is to stabilize the binding of the 5′-terminal nucleotide and DNA template once the E•DNA•NTP•NTP complex has formed (Baranovskiy et al. 2016b; Zerbe and Kuchta 2002; Kilkenny et al. 2013). There is a possibility that the initiation site is solely situated at PRIM2_C (Copeland and Wang 1993; Baranovskiy et al. 2016a). Alternatively, it might emerge as a composite-binding pocket formed when PRIM1 and PRIM2 assemble with the DNA template to form a preinitiation complex (E•DNA complex) (He et al. 2022a).

Following the formation of the dinucleotide, the primase undergoes a rapid RNA elongation phase and terminates promptly when the RNA primer is approximately

Fig. 5.2 (continued) primase subunits (PRIM1 and PRIM2). POLA1 is represented with distinct domains, including the exonuclease domain (catalytic dead), the polymerase “palm,” “fingers,” “thumb,” and the C-terminal domain (CTD). POLA2 contains a phosphodiesterase (PDE) domain and an oligonucleotide/oligosaccharide-binding (OB) fold. PRIM1 is illustrated with its two primase domains, and PRIM2 is shown with the N-terminal (PRIM2_N) and C-terminal (PRIM2_C) domains. **(b)** This panel shows the three-dimensional structures of the polα-primase apo-state complex (Baranovskiy et al. 2016a) (PDB: 5EXR) from two different perspectives, rotated 90 degrees relative to each other. The structural representation uses ribbon diagrams to distinguish the various subunits: POLA1 (red), POLA2 (light pink), PRIM1 (dark gold), and PRIM2 (gold). The green and blue overlays highlight the RNA and DNA catalytic centers of the PRIM1 and POLA1, respectively. **(c)** The two structures illustrate the elongation complexes I and II of the human polα-primase complex engaging with nucleic acids (He et al. 2023) (PDB: 8D96, 8D9D). In elongation complex I (EC-I), the RNA/DNA hybrid primer/template molecule is depicted with the RNA primer in green, the DNA primer in blue, and the DNA template in pink. Elongation complex II shows a larger complex with the other subunits bound to the EC-I complex, providing insight into how polα-primase can terminate primer DNA elongation using intramolecular interactions. Part of this figure is created with www.BioRender.com

7–10 nucleotides long (Sheaff and Kuchta 1993; Kuchta et al. 1990; Copeland and Wang 1993). The consistent and limited range of RNA products has led to the hypothesis of a potential “primer counting” mechanism (Frick and Richardson 2001; Kuchta and Stengel 2010; Baranovskiy and Tahirov 2017). Multiple models elucidating this mechanism have been put forth (Baranovskiy and Tahirov 2017; He and Lim 2023). One notable model suggests a physical caliper mechanism that capitalizes on the flexible linker present between the PRIM2 terminal domains (Yan et al. 2018; Baranovskiy et al. 2015, 2022; Xie 2011). Another more recent model suggests a physical interference by the pol α domain (Yuan et al. 2023; Yin et al. 2024) (see detailed discussion in the next section). In all cases, there seems to be a timely termination mechanism once the RNA primer reaches a “matured” length.

Once the 3'-end of the “matured” RNA primer detaches from the PRIM1 catalytic site, the pol α domain is primed to continue the elongation process, extending the primer with DNA. This culminates in the formation of the chimeric RNA-DNA primer. This intramolecular shift in catalytic activity takes place within a single pol α -primase complex, transitioning between the pol α and primase domains. The pol α domain consists of two subunits: POLA1 and POLA2. POLA1 acts as the DNA catalytic subunit, providing the polymerase activity responsible for synthesizing the DNA segment of the chimeric primer. On the other hand, POLA2 functions as an accessory to POLA1 and is integral to the initiation of DNA replication (Collins et al. 1993) as well as the DNA repair response (Dang and Morales 2020). While POLA2 is known to provide structural stability to the pol α -primase complex, the precise mechanisms through which it executes its essential cellular functions remain largely elusive.

Structural Biology of DNA Polymerase α -Primase

Over the past decade, the structural biology of eukaryotic pol α -primase has witnessed significant breakthroughs. A pivotal milestone was achieved with the X-ray crystal structure of the human apo (unbound)-state pol α -primase (Baranovskiy et al. 2016a) (Fig. 5.2b). This compact architecture was unexpected, especially in light of the bi-lobal structure previously observed in a negative-stain electron microscopy study of the yeast pol α -primase (Nunez-Ramirez et al. 2011). Crucially, with the POLA1 active site concealed and the PRIM1 site exposed, the apo-state structure offers insights into how the primase outcompetes the polymerase to take control of a DNA template (Baranovskiy and Tahirov 2017; Baranovskiy et al. 2016a, b). This “autoinhibitory” conformation of the apo state appears to be well-conserved, as evidenced by subsequent cryogenic-electron microscopy (cryo-EM) single-particle studies that revealed similar structures in the frog (*Xenopus laevis*) (Mullins et al. 2023) and yeast (*Saccharomyces cerevisiae*) (Yuan et al. 2023) enzymes.

Additional insights into the conformational stability of these apo-state structures are provided by the resolving power of cryo-EM heterogeneity analysis. The greater stability observed in the human pol α -primase cryo-EM apo-state structure (Yin et al. 2024), as compared to that seen in the frog (Mullins et al. 2023), ciliates (He

et al. 2022b), and yeast (Yuan et al. 2023) studies, suggests that these orthologs might transition more readily from the compact apo state to extended conformations that are poised to initiate de novo RNA primer synthesis. However, a caveat to these cryo-EM heterogeneity analyses is they were built on the assumptions that the cryo-EM samples (also prepared differently in these studies) retained their biochemical properties after the potentially harsh cryo-EM sample grid preparation process (Glaeser 2018).

So, what triggers the transition of the compact apo-state pol α -primase to its active form? This transition is essential for pol α -primase to synthesize RNA-DNA primers. The addition of a nucleotide, ssDNA template, or both did not induce this conformational change in the human pol α -primase (Yin et al. 2024). Intriguingly, the study on yeast pol α -primase suggests that template binding by Pri1 (the yeast homolog of PRIM1) facilitates this change, leading to the release of Pol1core (the yeast equivalent of POLA1) (Yuan et al. 2023). For the frog pol α -primase, this structural transition appears to occur spontaneously (Mullins et al. 2023). The shift in the human pol α -primase from its apo state to a preinitiation state has been captured using cryo-EM single-particle analysis (Yin et al. 2024). Here, the PRIM2_C domain relocates from its original position in the apo state to position itself in front of PRIM1 and DNA template, while the other subunits remain in their initial apo-state positions (Yin et al. 2024).

During this process, the POLA1 catalytic domain (POLA1_{CAT}) becomes dislodged from the rest of the enzyme, rendering it mobile and poised to bind a matured RNA/DNA primer/template for primer DNA elongation (Yin et al. 2024). A similar process is observed in the yeast pol α -primase study (Yuan et al. 2023). Biochemical and structural studies have demonstrated that the now mobile POLA1_{CAT} plays a crucial role in promptly terminating RNA primer elongation by PRIM1 (Yin et al. 2024; Yuan et al. 2023). However, whether POLA1_{CAT} performs this function through a steric clash mechanism or by engaging the RNA/DNA primer/template remains to be tested.

PRIM2_C plays a pivotal role in enabling pol α -primase to carry out several key processes, including the initiation of de novo RNA primer synthesis, the elongation of the RNA primer in a processive manner, and the intramolecular switch from RNA to DNA synthesis (Copeland and Wang 1993; Sheaff et al. 1994; Sheaff and Kuchta 1993; Arezi and Kuchta 2000; Baranovskiy et al. 2016a). The molecular foundation of these capabilities is rooted in the domain's capacity to interact with the 5'-triphosphate end of the RNA primer and the 3'-overhang of the DNA template (Baranovskiy et al. 2016b). An intriguing characteristic of PRIM2_C is its flexible attachment to the main body of the primase through an 18-amino acid linker (Baranovskiy and Tahirov 2017; Baranovskiy et al. 2016a; Arezi and Kuchta 2000). The flexibility of the linker, coupled with the abovementioned PRIM2_C interactions with the primer/template, provides a simple explanation to how PRIM2_C can perform its multiple functions (He and Lim 2023). However, alterations to the linker's length did not notably influence the length of the mature RNA primer (Baranovskiy et al. 2015). Instead, a truncated linker resulted in diminished primase activity (Baranovskiy et al. 2015). Similar results are seen in a recent study using *Sulfolobus* primase (Bergsch et al. 2021).

Another recent study has shed light on the role of the PRIM2 linker in pol α -primase function, revealing its significant involvement in terminating DNA elongation rather than RNA elongation (Baranovskiy et al. 2022). This implies that the PRIM2_C domain might persist in interacting with the primer/template even after it has been transferred to the POLA1_{CAT} domain for primer DNA elongation. Supporting this hypothesis, the first catalytic state structures of human pol α -primase in the DNA elongation state have been reported (He et al. 2023) (Fig. 5.2c). This cryo-EM study demonstrated that PRIM2_C remains bound to an RNA-DNA/DNA primer/template, even when the primer/template is engaged catalytically by POLA1_{CAT}. Similar interactions are also seen in the cryo-EM structures of the frog and yeast DNA elongation state (Mullins et al. 2023; Yuan et al. 2023), pointing toward a potentially conserved mechanism where PRIM2_C plays a pivotal role in terminating POLA1_{CAT}'s DNA elongation activity.

Notably, two distinct DNA elongation state structures were derived from a single primer/template substrate in the human DNA elongation complexes study (He et al. 2023). The first structure showcases the PRIM2_C-primer/template-POLA1_{CAT} interaction, termed the elongation complex I (EC-I), while the second captures EC-I in conjunction with the remaining pol α -primase subunits, denoted as elongation complex II (EC-II).

Intriguingly, in the EC-II structure, there are discernible interactions between PRIM1, PRIM2_N and, the POLA1_{CAT} domain. Disrupting these interactions through mutagenesis led to significantly longer RNA-DNA primers produced by the enzyme (He et al. 2023). This outcome mirrors the effects observed when the interactions between PRIM2_C and the primer/template are abolished (Baranovskiy et al. 2022). These findings suggest that pol α -primase employs two distinct mechanisms, both orchestrated by the primase, to regulate the length of the RNA-DNA primers. Considering the limited proofreading capability of pol α -primase, it seems evolutionarily prudent for the enzyme to establish a fail-safe mechanism. This mechanism ensures the primers are concise, offering just enough stability to kickstart the DNA replication process.

CTC1-STN1-TEN1 (CST): A Mammalian RPA-Like Factor for DNA Polymerase α -Primase

Replication protein A (RPA) is a highly conserved ssDNA-binding protein complex essential for a variety of DNA-based transactions in cells (Wold 1997; Fanning et al. 2006; Dueva and Iliakis 2020). Comprised of three subunits: RPA70, RPA32, and RPA14 (Wold and Kelly 1988; Fairman and Stillman 1988; Wold et al. 1989) (Fig. 5.3a), it was first discovered as an accessory factor for simian virus 40 (SV40) DNA replication in vitro (Wold and Kelly 1988; Fairman and Stillman 1988). Key roles of RPA include stabilizing unwound DNA at the replication fork and stimulating pol α -primase's activity (Challberg and Kelly 1989; Stillman 1989; Hurwitz

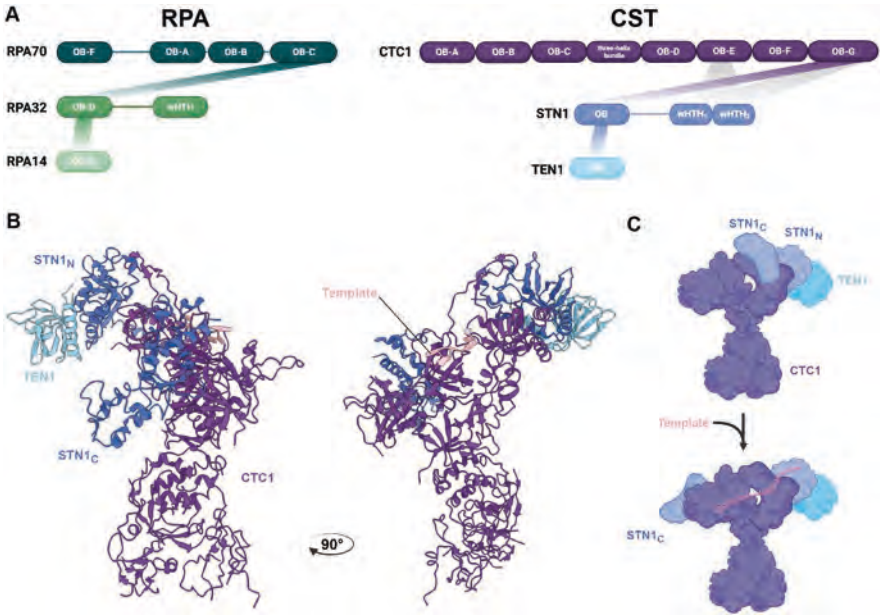


Fig. 5.3 Domain organization and structure of the RPA-like human CST. (a) The left diagram presents the domain architecture of the RPA complex, comprised of RPA70, RPA32, and RPA14 subunits, each characterized by oligonucleotide/oligosaccharide-binding (OB) folds. RPA32 features a winged-helix-turn-helix (WHTH) motif. Adjacent to it, the CST complex is depicted, consisting of CTC1, STN1, and TEN1 subunits. CTC1 includes several OB folds and a distinct three-helix bundle structure, STN1 is equipped with OB folds and a tandem WHTH motifs, and TEN1 is a single OB fold. Inter-subunit interactions are depicted. (b) Cryo-EM structure of the human CST complex bound to a telomeric ssDNA (Lim et al. 2020) (PDB: 6W6W). The CTC1 subunit (purple) acts as an assembly scaffold for the STN1 (blue) and TEN1 (light blue) subunits. STN1's N-terminal and C-terminal domains (STN1_N and STN1_C, respectively) engage CTC1 at two different positions. (c) Alternate docking positions of STN1_C, regulated by CST binding of ssDNA. When no ssDNA is bound, the STN1_C domain docks on top of CTC1 OB-G domain. CST binding to ssDNA excludes this docking position of the STN1_C domain, resulting the domain to dock on CTC1 OB-E domain. Part of this figure is created with www.BioRender.com

et al. 1990). Central to these functions is RPA's DNA-binding capability. It has a strong affinity for ssDNA and a much lower one for dsDNA (Wold 1997).

Yeast telomere research led to the discovery of an RPA-like CST complex (Cdc13-STN1-TEN1) that plays a key role in telomere maintenance (Gao et al. 2007; Lin and Zakian 1996; Nugent et al. 1996; Grandin et al. 2001; Grandin et al. 1997; Pennock et al. 2001). This finding initiated a search for a mammalian homolog, resulting in the identification of the mammalian CST complex (CTC1-STN1-TEN1) (Fig. 5.3a). Interestingly, this mammalian complex was found to have a greater resemblance to RPA than its yeast counterpart in terms of protein organization and ssDNA-binding behavior (Miyake et al. 2009; Surovtseva et al. 2009). The CTC1 and STN1 subunits were first discovered as accessory proteins for pol α -primase, initially termed as AAF-132 for CTC1 and AAF-44 for STN1 (Goulian

et al. 1990; Casteel et al. 2009). Their discovery arose from the observed elevated activity of pol α -primase during the enzyme purification (Goulian et al. 1990). Subsequent biochemical characterizations showed these accessory factors enhance pol α -primase-binding affinity to DNA templates, in addition to stimulating both the pol α and primase's activity (Goulian and Heard 1990).

While the human CST and RPA complexes share several common features, they differ in their sequence preference for DNA binding (Barbour and Wuttke 2023). RPA has a preference for binding polypyrimidines over mixed sequences or polypurines ssDNA (Kim et al. 1992), while CST favors binding to G-rich ssDNA sequences (Chen et al. 2012; Hom and Wuttke 2017). CST's affinity for G-rich ssDNA is length-dependent, diminishing when the ssDNA sequence exceeds 36 nucleotides (Miyake et al. 2009; Chen et al. 2012; Bhattacharjee et al. 2016). This behavior suggests that CST possesses at least two binding modes contingent on length, akin to RPA (Blackwell and Borowiec 1994; Arunkumar et al. 2003; Fanning et al. 2006; Pokhrel et al. 2019). Beyond ssDNA binding, CST also associates with various DNA structures, including the ss-dsDNA junction, for which it seemingly has no overhang sequence preference (Bhattacharjee et al. 2017). However, in contrast to RPA, which can unwind an ss-dsDNA junction (Lao et al. 1999; Stewart et al. 2008), CST lacks this capability (Bhattacharjee et al. 2017). A shared important function between RPA and CST is their ability to bind and unwind telomeric DNA G-quadruplexes (GQ) (Williamson et al. 1989; Sundquist and Klug 1989; Williamson 1994; Fan et al. 2009; Bhattacharjee et al. 2017), highly stable DNA structures formed from exposed G-rich ssDNA when the duplex unwinds. These GQs block DNA template translocation of processive enzymes (Stewart et al. 2012a; Spiegel et al. 2020). RPA and CST ability to resolve GQs underscores their roles in aiding replication forks' progression at genomic sequences that are prone to form GQs (Grudic et al. 2007; Zhang et al. 2019). A recent biochemical study revealed that RPA is more adept than CST at unwinding telomeric GQs (Olson et al. 2023), suggesting that the more abundant RPA might predominantly handle the resolution of telomeric GQs.

Numerous studies revealed that mammalian CST plays multiple roles in a variety of DNA-based transactions, even beyond that at telomeres. CST is pivotal in the restart of genome-wide stalled replication forks (Stewart et al. 2012b; Chastain et al. 2016). Studies indicate that CST aids replication restart by promoting origin firing, likely due to CST engaging dormant replication origins rather than primary origins (Stewart et al. 2012b; Chastain et al. 2016). Subsequent studies have provided insights into how CST can facilitate stalled replication fork recovery through multiple pathways; (1) CST recruits Rad51 to stalled replication forks to protect exposed ssDNA at the forks from nucleolytic attacks (Chastain et al. 2016), (2) CST resolves GQ-based replication blockages (Zhang et al. 2019), and (3) CST blocks MRE11 from binding to ss-dsDNA junctions at stalled replication forks to avert MRE11-mediated DNA degradation (Lyu et al. 2021). An additional mechanism might involve CST delivering pol α -primase to these stalled replication forks (Wang et al. 2012). Furthermore, CST also plays a downstream role in double-strand breaks (DSB) repair via the 53BP1-shieldin complex pathway, where the shieldin complex

recruits CST-pol α -primase to fill-in resected DSBs (Barazas et al. 2018; Mirman et al. 2018).

CST is critical in aiding the replisome to traverse the repetitive G-rich DNA sequences at telomeres (Wang et al. 2012; Stewart et al. 2012b; Kasbek et al. 2013; Gu et al. 2018). The telomeric DNA's tendency to form GQs present challenges for replisomes to progress at telomeres. CST helps the replisomes navigate these GQ roadblocks (Zhang et al. 2019) by binding to and unwinding the GQs (Bhattacharjee et al. 2017). Telomere extension commences once the genome has been duplicated by the replisomes. CST plays two essential roles in the telomeric DNA synthesis process: firstly, in terminating the telomerase extension of the G-overhang (Chen et al. 2012; Zaug et al. 2021), and secondly, in initiating the C-strand fill-in by pol α -primase (Wang et al. 2012; Wu et al. 2012; Gu et al. 2012; Feng et al. 2017; Zaug et al. 2022; He et al. 2022a; He and Lim 2023). CST terminates telomerase activity by competing with the telomerase for the 3'-termini of the G-overhang, effectively preventing telomerase from (re)binding to the DNA (Chen et al. 2012; Zaug et al. 2021). However, once telomerase engages the overhang and starts extending it, CST cannot disrupt this process (Zaug et al. 2021). This implies that, within a cellular context, CST acts as a gatekeeper, preventing telomerase from successively extending the telomeric overhangs after an initial extension event. This perspective aligns with cellular studies which posits that each telomeric overhang undergoes extension just once during a single cell division (Zhao et al. 2011).

The mammalian CST's ability to stimulate pol α -primase activity (Goulian and Heard 1990; Nakaoka et al. 2012) suggests CST plays a direct role in the telomere C-strand fill-in process by pol α -primase. Indeed, biochemical studies using reconstituted human protein complexes showed CST can stimulate pol α -primase activity on telomeric templates (Zaug et al. 2022; He et al. 2022a). Interestingly, pol α -primase initiates de novo strand synthesis at specific periodic positions along the repetitive telomeric templates—namely, at the second thymidine within the hexameric TTAGGG repeat sequence—regardless of CST presence (Nozawa et al. 2000; Zaug et al. 2022; He et al. 2022a). This primase preference for starting RNA primer synthesis at the double thymidine block of the telomeric DNA repeat sequence corroborates with previous findings that DNA pyrimidine clusters are preferred primase initiation sites (Suzuki et al. 1993). The primase's unique sequence preference may be an evolutionary adaptation, given that ATP is the most abundant ribonucleotide in cells. In terms of telomeric DNA synthesis, this implies a specific replicon for the telomere C-strand fill-in synthesis (Zaug et al. 2022); located at the second thymidine of the highly conserved mammalian telomeric DNA repeats, TTAGGG (Meyne et al. 1989).

The cell cycle-dependent cellular trafficking of CST relies on its interaction with pol α -primase (Kelich et al. 2021). This finding is consistent with earlier cellular immunoprecipitation experiments which showed CST remains tightly associated with pol α -primase across multiple purification steps (Goulian et al. 1990). Thus, it is plausible that CST also plays a role in recruiting pol α -primase to newly

synthesized telomere G-overhangs. Efforts at the cellular level to delineate the individual roles of each CST subunit indicate that TEN1 is not vital for CST termination of telomerase activity, but it is essential for CST to fulfill its role in C-strand fill-in (Feng et al. 2018). The cryo-EM structure of the human CST-pol α -primase preinitiation complex (PIC) sheds light on TEN1's role in this process (He et al. 2022a).

Structural Biology of Human CST Complex

Cryo-EM single-particle analysis has unveiled the first complete structure of the human CST complex (Lim et al. 2020) (Fig. 5.3b). The structure reveals that the telomeric ssDNA-bound CST assembles as a decamer. This assembly is facilitated by CST's association with telomeric ssDNA (Lim et al. 2020). Within this structure, CTC1 possesses seven oligonucleotide-/oligosaccharide-binding fold (OB-fold) domains, labeled OB-A through OB-G. These OB-fold domains are characterized by their compact beta-barrel structures and are frequently observed in nucleic acid-binding proteins or protein complexes that bind single-stranded DNA, like RPA (Barbour and Wuttke 2023). Notably, the core ssDNA-binding site of CST is shared across the OB-F and OB-G domains of CST (Lim et al. 2020). The OB-F domain interacts with the telomeric ssDNA molecule using nonspecific interactions, while the OB-G domain contacts the guanine bases of the telomeric ssDNA, indicating the OB-G domain utilizes a more specific set of interactions. Disruptive mutagenesis of the DNA-interacting residues within these two domains results in a significant decrease in CST's ssDNA-binding capacity (Lim et al. 2020).

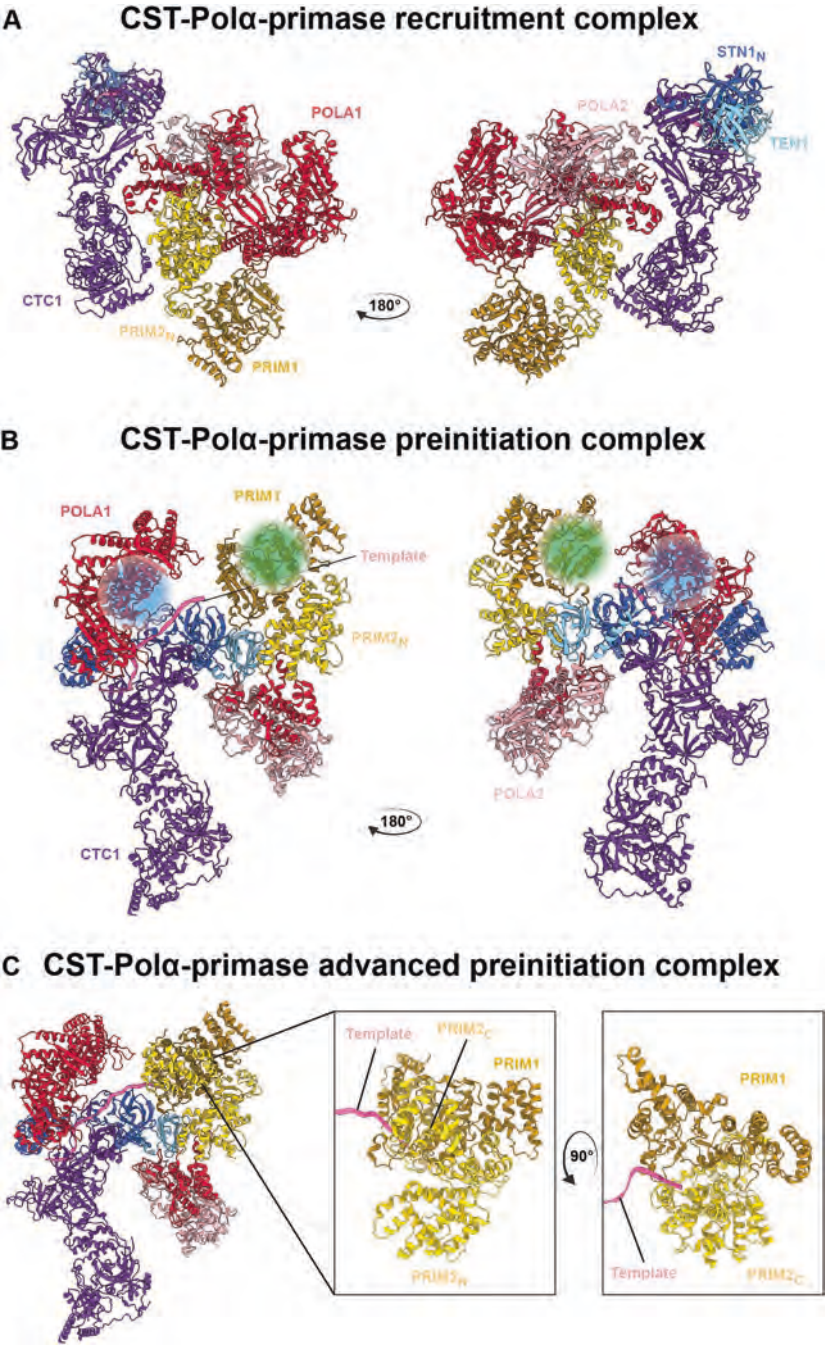
The STN1 subunit possesses two structured domains: the N-terminal domain (STN1_N) and the C-terminal domain (STN1_C) (Bryan et al. 2013; Rice and Skordalakes 2016). Both of these domains interact with CTC1 at two distinct sites (Fig. 5.3b). Specifically, the STN1_N domain, a single OB-fold domain (Bryan et al. 2013), binds to the CTC1 OB-G domain. Meanwhile, the STN1_C domain, a tandem winged helix-turn-helix domain (Bryan et al. 2013), engages the CTC1 OB-E domain. However, the latter interaction is transient in nature (Lim et al. 2020). This transient quality allows the STN1_C domain to alternate its positions on CTC1 (Lim et al. 2020) (Fig. 5.3c). In the absence of ssDNA binding, the STN1_C domain is positioned atop the CTC1 OB-G domain. When ssDNA is bound, the STN1_C domain shifts to bind the CTC1 OB-E domain. Hence, the docking position of STN1_C is influenced by the CST's ssDNA binding, which in turn dictates whether CST assembles as a decamer (Lim et al. 2020). The specific role of the CST-ssDNA decamer remains ambiguous and necessitates further investigation (Lim and Cech 2021; He and Lim 2023). As for the TEN1 subunit, which is an OB-fold domain (Bryan et al. 2013), it does not directly contact CTC1 but exclusively associates with the STN1_N domain (Lim et al. 2020) (Fig. 5.3b). This indicates that the incorporation of TEN1 into CST is dependent on STN1.

Structural Biology of Human CST-pol α -Primase Complex

Studies have indicated that the human shelterin subunit TPP1 plays a primary role in recruiting CST to telomeres (Wan et al. 2009; Chen et al. 2012; Wang et al. 2023). A recent cryo-EM study, however, challenges this view by suggesting that it is the shelterin POT1 subunit—not TPP1—that directly interacts with CST and that the POT1-CST interaction is regulated by phosphorylation (Cai et al. 2023). The authors of the study detected no interaction between human CST and POT1 under normal conditions. But when they engineered human POT1 to include a four-amino acid sequence (ESDL) from mouse POT1b, which is known to interact with mouse CST (Wu et al. 2012), they observed a more stable interaction between the chimeric human/mouse POT1 and human CST (Cai et al. 2023). This enhanced interaction was attributed to increased phosphorylation of several serine residues in the POT1 hinge region, influenced by the insertion of the mouse sequence (Cai et al. 2023). The kinase responsible for this phosphorylation in the human POT1 model has not been identified and remains a subject for future research. Notably, the cryo-EM structure of the human CST-TPP1-POT1(ESDL) complex aligns well with the structure of the human CST-pol α -primase recruitment complex (RC) (Cai et al. 2022). This structural compatibility suggests that a phosphorylated POT1 within the shelterin complex could play a role in either attracting a preassembled CST-pol α -primase complex or facilitating the recruitment of pol α -primase by CST to the telomeric overhang.

The cryo-EM structure of the CST-pol α -primase RC (Cai et al. 2022) (Fig. 5.4a) shows that upon binding to CST, pol α -primase retains the compact conformation observed in its apo-state structure (Baranovskiy et al. 2016a), indicating that the pol α -primase in the CST-pol α -primase RC is not activated for de novo primer synthesis. The activation mechanism for pol α -primase is further elucidated in the cryo-EM structure of the CST-pol α -primase preinitiation complex (PIC) (He et al. 2022a) (Fig. 5.4b). In the PIC structure, the DNA polymerase and primase domains of pol α -primase are spatially segregated by CST. CST achieves this by providing distinct binding sites for these pol α -primase domains while also engaging the telomeric ssDNA template. The binding surface for the POLA1_{CAT} domain is provided by the template-bound CTC1, whereas the TEN1 subunit is responsible for binding the primase subunits, POLA2, and POLA1_{CTD}—pol α -primase’s components collectively known as the “platform” (Baranovskiy and Tahirov 2017).

The catalyst for pol α -primase transition from a compact state in the CST-pol α -primase RC to an extended and separated conformation in the CST-pol α -primase PIC remains unclear. CST’s binding to telomeric ssDNA is unlikely the trigger—both the RC and PIC are template-bound (Cai et al. 2022; He et al. 2022a). Interestingly, the STN1_C domain exhibits flexibility in the CST-pol α -primase RC structure (Cai et al. 2022) but in the PIC structure, it forms a stable interaction with both the CTC1 and POLA1_{CAT} domains (He et al. 2022a). This observation suggests that the STN1_C domain might have a pivotal role in facilitating the conformational change of pol α -primase, beyond its previously described function in CST-ssDNA oligomerization (Lim et al. 2020). Notably, the CST-pol α -primase PIC



configuration precludes POT1 from binding to CST (Cai et al. 2023), implying that for the CST-pol α -primase RC to transit into a PIC for telomeric C-strand synthesis, POT1 must detach from CST, a process that would require POT1 dephosphorylation (Cai et al. 2023).

The human CST-pol α -primase PIC structure reveals how CST stimulates pol α -primase activity for de novo RNA primer synthesis (He et al. 2022a). In this complex, CST's binding affinity and sequence specificity come into play, stabilizing the telomeric ssDNA template to facilitate the initiation of RNA primer synthesis by pol α -primase (He et al. 2022a). The PIC structure elucidated more of the CST-bound telomeric ssDNA compared to the previously reported CST-ssDNA structure (Lim et al. 2020), with an additional 11 nucleotides of telomeric ssDNA resolved, extending from the previously determined 4 nucleotide fragment at the CST core ssDNA-binding site. This suggests that the binding of the POLA1_{CAT} domain to the template-bound CST may reinforce the more transient CST-ssDNA interactions utilized by CST. These interactions are mediated by two further sites within the CTC1 OB-G and STN1_N domains (He et al. 2022a).

CST positions the 3'-end of the template DNA toward the RNA catalytic center facet of PRIM1, where the PRIM2C domain can engage the DNA and assemble with PRIM1 to form the primase-template complex that is poised to start de novo RNA primer synthesis (He et al. 2022a) (Fig. 5.4c). The assembly of the primase-template complex, as observed in the "advanced" PIC cryo-EM structure (He et al. 2022a), resolved an additional 5 nucleotides of the template 3'-end and revealed that the template is stably positioned in front of the PRIM1 RNA catalytic center and the elongation nucleotide-binding site. This advanced PIC conformation is presumably ready to accept NTPs and initiate de novo RNA primer synthesis. However, it remains unclear how the primase elongation nucleotide-binding site is occupied before the initiation site (Sheaff and Kuchta 1993). There seems to be a pathway for free NTPs to reach these binding sites, as evidenced by the crevice formed between PRIM1 and PRIM2_C (He et al. 2022a), although the sequence of nucleotide binding and the mechanism of initial dinucleotide formation require further investigation.

The advanced PIC structure has provided structural insights into why a certain length of telomeric template is necessary for CST to enhance pol α -primase activity (He et al. 2022a). For a CST-bound template to reach the PRIM1 RNA catalytic center and participate in the assembly of the primase-template complex, it must be

Fig. 5.4 (continued) (POLA1 in red, POLA2 in light pink, PRIM1 in dark gold, and PRIM2 in gold). This arrangement signifies the initial recruitment stage, where the CST complex interacts with the pol α -primase to form an assembly at telomeric overhangs. **(b)** Cryo-EM structure of the CST-pol α -primase complex (He et al. 2022a) (PDB: 8D0B) at the preinitiation stage, with the DNA template strand (pink) shown in position for the primase to initiate primer synthesis. The RNA and DNA catalytic centers, highlighted in green and blue, respectively, are now separated (by CST) as compared to the compact form seen in the enzyme apo state (see Fig. 5.2). **(c)** Cryo-EM structure of the advanced preinitiation stage (He et al. 2022a) (PDB: 8D0K) where the PRIM2_C domain (gold) docks in front of PRIM1 (dark gold) and the DNA template (hot pink) (insets). Together, they assemble the primase unit that is poised to incorporate NTPs to start de novo RNA primer synthesis

at least 20 nucleotides in length. This requirement elucidates why CST did not stimulate pol α -primase activity with an 18 nucleotide telomeric ssDNA template, while it did with a 36 nucleotide template (He et al. 2022a). However, a template that is long enough to reach the primase catalytic center is likely not sufficient to initiate de novo RNA primer synthesis. The binding of PRIM2_C to the template is crucial for this process (Baranovskiy et al. 2016b), suggesting that a longer template, which can also accommodate PRIM2_C binding, is likely necessary to support CST's stimulatory effect.

The human CST-pol α -primase advanced PIC structure suggests that the assembly of the primase-template complex does not involve CST (He et al. 2022a). This implies that CST likely does not influence the rate-limiting step in the initial dinucleotide formation (Sheaff and Kuchta 1993) and that CST's role in enhancing pol α -primase activity is confined to template binding and the reshaping of pol α -primase (He et al. 2022a; Zaug et al. 2022). Whether CST has additional functions in other steps of pol α -primase catalysis remains to be investigated. Notably, certain conspicuous features of the human CST-pol α -primase PIC structure could participate in the processes of primase RNA elongation and termination, as covered in a related review (He and Lim 2023). One salient feature is the seemingly restrictive enclosure for the ssDNA template created by CST and the POLA1_{CAT} domain (Fig. 5.4b). This feature, also present in the *Tetrahymena* CST-pol α -primase cryo-EM structure (He et al. 2022b), suggests a conserved structural characteristic, the function of which is not yet understood. It has been proposed to be related to the template translocation mechanism (He et al. 2022a; He and Lim 2023). Additionally, the flexibly tethered PRIM2_C domain is strategically positioned between the RNA and DNA catalytic centers of pol α -primase (He et al. 2022a) (Fig. 5.4b), facilitating its involvement in RNA primer synthesis and termination as well as the transition to DNA elongation without necessitating additional conformational changes in the CST-pol α -primase PIC. This is corroborated by the cryo-EM structure of the human DNA elongation complex, which is structurally compatible with the CST-pol α -primase PIC structure (He et al. 2023).

Conclusion

Recent advances in cryo-EM technology have significantly enhanced our structural understanding of telomerase holoenzyme biology over the past decade (Sauerwald et al. 2013; Jiang et al. 2013, 2015, 2018; Nguyen et al. 2018; Ghanim et al. 2021; Wan et al. 2021; Sekne et al. 2022; Liu et al. 2022; He et al. 2022b). They have yielded valuable insights into the structural-functional relationship between shelterin, telomerase, and their role in extending the telomeric G-overhang. The pioneering work conducted in the ciliate system, followed by the human system, has established a foundational framework for telomerase complex assembly and function. This research has been instrumental in informing and enriching our understanding across various organisms.

Recently, our comprehension of the second phase of telomeric DNA synthesis involving C-strand fill-in by CST and DNA polymerase alpha-primase has also expanded. This chapter delves into the latest structures and functions of human CST and pol α -primase, shedding light on the complex orchestration of telomeric C-strand synthesis. Although the focus of this chapter is on the human system, it is noteworthy that our structural understanding of the *Tetrahymena* C-strand fill-in process has advanced in parallel, providing comparative insights that highlight the universality and diversity of telomere biology (He et al. 2022b). Looking ahead, the challenge will be to harness these insights to further unravel the molecular mechanics of CST-pol α -primase interactions and the critical conformational transitions required for RNA-DNA primer synthesis. A deeper understanding in this domain is crucial for elucidating broader biological processes, such as cellular aging and cancer proliferation.

References

- Arezi B, Kuchta RD (2000) Eukaryotic DNA primase. *Trends Biochem Sci* 25(11):572–576. [https://doi.org/10.1016/s0968-0004\(00\)01680-7](https://doi.org/10.1016/s0968-0004(00)01680-7)
- Arunkumar AI, Stauffer ME, Bochkareva E, Bochkarev A, Chazin WJ (2003) Independent and coordinated functions of replication protein a tandem high affinity single-stranded DNA binding domains. *J Biol Chem* 278(42):41077–41082. <https://doi.org/10.1074/jbc.M305871200>
- Baranovskiy AG, Babayeva ND, Zhang Y, Gu J, Suwa Y, Pavlov YI, Tahirov TH (2016a) Mechanism of concerted RNA-DNA primer synthesis by the human primosome. *J Biol Chem* 291(19):10006–10020. <https://doi.org/10.1074/jbc.M116.717405>
- Baranovskiy AG, Lisova AE, Morstadt LM, Babayeva ND, Tahirov TH (2022) Insight into RNA-DNA primer length counting by human primosome. *Nucleic Acids Res* 50(11):6264–6270. <https://doi.org/10.1093/nar/gkac492>
- Baranovskiy AG, Tahirov TH (2017) Elaborated action of the human primosome. *Genes (Basel)* 8(2). <https://doi.org/10.3390/genes8020062>
- Baranovskiy AG, Zhang Y, Suwa Y, Babayeva ND, Gu J, Pavlov YI, Tahirov TH (2015) Crystal structure of the human primase. *J Biol Chem* 290(9):5635–5646. <https://doi.org/10.1074/jbc.M114.624742>
- Baranovskiy AG, Zhang Y, Suwa Y, Gu J, Babayeva ND, Pavlov YI, Tahirov TH (2016b) Insight into the human DNA primase interaction with template-primer. *J Biol Chem* 291(9):4793–4802. <https://doi.org/10.1074/jbc.M115.704064>
- Barazas M, Annunziato S, Pettitt SJ, de Krijger I, Ghezraoui H, Roobol SJ, Lutz C, Frankum J, Song FF, Brough R, Evers B, Gogola E, Bhin J, van de Ven M, van Gent DC, Jacobs JJJ, Chapman R, Lord CJ, Jonkers J, Rottenberg S (2018) The CST complex mediates end protection at double-strand breaks and promotes PARP inhibitor sensitivity in BRCA1-deficient cells. *Cell Rep* 23(7):2107–2118. <https://doi.org/10.1016/j.celrep.2018.04.046>
- Barbour AT, Wuttke DS (2023) RPA-like single-stranded DNA-binding protein complexes including CST serve as specialized processivity factors for polymerases. *Curr Opin Struct Biol* 81:102611. <https://doi.org/10.1016/j.sbi.2023.102611>
- Baumann P, Cech TR (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* 292(5519):1171–1175. <https://doi.org/10.1126/science.1060036>
- Bergsch J, Devillier JC, Jeschke G, Lipps G (2021) Stringent primer termination by an archaeo-eukaryotic DNA primase. *Front Microbiol* 12:652928. <https://doi.org/10.3389/fmicb.2021.652928>

- Bhattacharjee A, Stewart J, Chaiken M, Price CM (2016) STN1 OB fold mutation alters DNA binding and affects selective aspects of CST function. *PLoS Genet* 12(9):e1006342. <https://doi.org/10.1371/journal.pgen.1006342>
- Bhattacharjee A, Wang Y, Diao J, Price CM (2017) Dynamic DNA binding, junction recognition and G4 melting activity underlie the telomeric and genome-wide roles of human CST. *Nucleic Acids Res* 45(21):12311–12324. <https://doi.org/10.1093/nar/gkx878>
- Bianchi A, Smith S, Chong L, Elias P, de Lange T (1997) TRF1 is a dimer and bends telomeric DNA. *EMBO J* 16(7):1785–1794. <https://doi.org/10.1093/emboj/16.7.1785>
- Blackwell LJ, Borowiec JA (1994) Human replication protein a binds single-stranded DNA in two distinct complexes. *Mol Cell Biol* 14(6):3993–4001. <https://doi.org/10.1128/mcb.14.6.3993-4001.1994>
- Bonnell E, Pasquier E, Wellinger RJ (2021) Telomere replication: solving multiple end replication problems. *Front Cell Dev Biol* 9:668171. <https://doi.org/10.3389/fcell.2021.668171>
- Broccoli D, Smogorzewska A, Chong L, de Lange T (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat Genet* 17(2):231–235. <https://doi.org/10.1038/ng1097-231>
- Bryan C, Rice C, Harkisheimer M, Schultz DC, Skordalakes E (2013) Structure of the human telomeric Stn1-Ten1 capping complex. *PLoS One* 8(6):e66756. <https://doi.org/10.1371/journal.pone.0066756>
- Cai SW, de Lange T (2023) CST-polalpha/primase: the second telomere maintenance machine. *Genes Dev* 37(13–14):555–569. <https://doi.org/10.1101/gad.350479.123>
- Cai SW, Takai H, Walz T, Lange Td (2023) POT1 recruits and regulates CST–Pol α /Primase at human telomeres. *bioRxiv*:2023.2005.2008.539880. <https://doi.org/10.1101/2023.05.08.539880>
- Cai SW, Zinder JC, Svetlov V, Bush MW, Nudler E, Walz T, de Lange T (2022) Cryo-EM structure of the human CST-polalpha/primase complex in a recruitment state. *Nat Struct Mol Biol* 29:813. <https://doi.org/10.1038/s41594-022-00766-y>
- Casteel DE, Zhuang S, Zeng Y, Perrino FW, Boss GR, Goulian M, Pilz RB (2009) A DNA polymerase- α -primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J Biol Chem* 284(9):5807–5818. <https://doi.org/10.1074/jbc.M807593200>
- Challberg MD, Kelly TJ (1989) Animal virus DNA replication. *Annu Rev Biochem* 58:671–717. <https://doi.org/10.1146/annurev.bi.58.070189.003323>
- Chan H, Wang Y, Feigon J (2017) Progress in human and Tetrahymena telomerase structure determination. *Annu Rev Biophys* 46:199–225. <https://doi.org/10.1146/annurev-biophys-062215-011140>
- Chastain M, Zhou Q, Shiva O, Fadri-Moskwik M, Whitmore L, Jia P, Dai X, Huang C, Ye P, Chai W (2016) Human CST facilitates genome-wide RAD51 recruitment to GC-rich repetitive sequences in response to replication stress. *Cell Rep* 16(5):1300–1314. <https://doi.org/10.1016/j.celrep.2016.06.077>
- Chen LY, Redon S, Lingner J (2012) The human CST complex is a terminator of telomerase activity. *Nature* 488(7412):540–544. <https://doi.org/10.1038/nature11269>
- Chong L, van Steensel B, Broccoli D, Erdjument-Bromage H, Hanish J, Tempst P, de Lange T (1995) A human telomeric protein. *Science* 270(5242):1663–1667. <https://doi.org/10.1126/science.270.5242.1663>
- Chow TT, Zhao Y, Mak SS, Shay JW, Wright WE (2012) Early and late steps in telomere overhang processing in normal human cells: the position of the final RNA primer drives telomere shortening. *Genes Dev* 26(11):1167–1178. <https://doi.org/10.1101/gad.187211.112>
- Collins KL, Russo AA, Tseng BY, Kelly TJ (1993) The role of the 70 kDa subunit of human DNA polymerase alpha in DNA replication. *EMBO J* 12(12):4555–4566. <https://doi.org/10.1002/j.1460-2075.1993.tb06144.x>
- Copeland WC, Wang TS (1993) Enzymatic characterization of the individual mammalian primase subunits reveals a biphasic mechanism for initiation of DNA replication. *J Biol Chem* 268(35):26179–26189

- Dalby AB, Hofr C, Cech TR (2015) Contributions of the TEL-patch amino acid cluster on TPP1 to telomeric DNA synthesis by human telomerase. *J Mol Biol* 427(6 Pt B):1291–1303. <https://doi.org/10.1016/j.jmb.2015.01.008>
- Dang TT, Morales JC (2020) Involvement of POLA2 in double Strand break repair and genotoxic stress. *Int J Mol Sci* 21(12). <https://doi.org/10.3390/ijms21124245>
- de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19(18):2100–2110. <https://doi.org/10.1101/gad.1346005>
- Dueva R, Iliakis G (2020) Replication protein a: a multifunctional protein with roles in DNA replication, repair and beyond. *NAR. Cancer* 2(3):zcaa022. <https://doi.org/10.1093/narcan/zcaa022>
- Fairman MP, Stillman B (1988) Cellular factors required for multiple stages of SV40 DNA replication in vitro. *EMBO J* 7(4):1211–1218. <https://doi.org/10.1002/j.1460-2075.1988.tb02933.x>
- Fan JH, Bochkareva E, Bochkarev A, Gray DM (2009) Circular dichroism spectra and electrophoretic mobility shift assays show that human replication protein a binds and melts intramolecular G-quadruplex structures. *Biochemistry* 48(5):1099–1111. <https://doi.org/10.1021/bi801538h>
- Fanning E, Klimovich V, Nager AR (2006) A dynamic model for replication protein a (RPA) function in DNA processing pathways. *Nucleic Acids Res* 34(15):4126–4137. <https://doi.org/10.1093/nar/gkl550>
- Feng X, Hsu SJ, Bhattacharjee A, Wang Y, Diao J, Price CM (2018) CTC1-STN1 terminates telomerase while STN1-TEN1 enables C-strand synthesis during telomere replication in colon cancer cells. *Nat Commun* 9(1):2827. <https://doi.org/10.1038/s41467-018-05154-z>
- Feng X, Hsu SJ, Kasbek C, Chaiken M, Price CM (2017) CTC1-mediated C-strand fill-in is an essential step in telomere length maintenance. *Nucleic Acids Res* 45(8):4281–4293. <https://doi.org/10.1093/nar/gkx125>
- Frick DN, Richardson CC (2001) DNA primases. *Annu Rev Biochem* 70:39–80. <https://doi.org/10.1146/annurev.biochem.70.1.39>
- Gao H, Cervantes RB, Mandell EK, Otero JH, Lundblad V (2007) RPA-like proteins mediate yeast telomere function. *Nat Struct Mol Biol* 14(3):208–214. <https://doi.org/10.1038/nsmb1205>
- Ghanim GE, Fountain AJ, van Roon AM, Rangan R, Das R, Collins K, Nguyen THD (2021) Structure of human telomerase holoenzyme with bound telomeric DNA. *Nature* 593(7859):449–453. <https://doi.org/10.1038/s41586-021-03415-4>
- Glaeser RM (2018) Proteins, interfaces, and cryo-Em grids. *Curr Opin Colloid Interface Sci* 34:1–8. <https://doi.org/10.1016/j.cocis.2017.12.009>
- Goulian M, Heard CJ (1990) The mechanism of action of an accessory protein for DNA polymerase alpha/primase. *J Biol Chem* 265(22):13231–13239
- Goulian M, Heard CJ, Grimm SL (1990) Purification and properties of an accessory protein for DNA polymerase alpha/primase. *J Biol Chem* 265(22):13221–13230
- Grandin N, Damon C, Charbonneau M (2001) Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *EMBO J* 20(5):1173–1183. <https://doi.org/10.1093/emboj/20.5.1173>
- Grandin N, Reed SI, Charbonneau M (1997) Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes Dev* 11(4):512–527. <https://doi.org/10.1101/gad.11.4.512>
- Grudic A, Jul-Larsen A, Haring SJ, Wold MS, Lonning PE, Bjerkvig R, Boe SO (2007) Replication protein a prevents accumulation of single-stranded telomeric DNA in cells that use alternative lengthening of telomeres. *Nucleic Acids Res* 35(21):7267–7278. <https://doi.org/10.1093/nar/gkm738>
- Gu P, Jia S, Takasugi T, Smith E, Nandakumar J, Hendrickson E, Chang S (2018) CTC1-STN1 coordinates G- and C-strand synthesis to regulate telomere length. *Aging Cell* 17(4):e12783. <https://doi.org/10.1111/acer.12783>
- Gu P, Min JN, Wang Y, Huang C, Peng T, Chai W, Chang S (2012) CTC1 deletion results in defective telomere replication, leading to catastrophic telomere loss and stem cell exhaustion. *EMBO J* 31(10):2309–2321. <https://doi.org/10.1038/emboj.2012.96>

- Harley CB (1991) Telomere loss: mitotic clock or genetic time bomb? *Mutat Res* 256(2–6):271–282. [https://doi.org/10.1016/0921-8734\(91\)90018-7](https://doi.org/10.1016/0921-8734(91)90018-7)
- Harley CB, Futcher AB, Greider CW (1990) Telomeres shorten during ageing of human fibroblasts. *Nature* 345(6274):458–460. <https://doi.org/10.1038/345458a0>
- Hayflick L (1965) The limited in vitro lifetime of human diploid cell strains. *Exp Cell Res* 37:614–636. [https://doi.org/10.1016/0014-4827\(65\)90211-9](https://doi.org/10.1016/0014-4827(65)90211-9)
- Hayflick L, Moorhead PS (1961) The serial cultivation of human diploid cell strains. *Exp Cell Res* 25:585–621. [https://doi.org/10.1016/0014-4827\(61\)90192-6](https://doi.org/10.1016/0014-4827(61)90192-6)
- He Q, Baranovskiy AG, Morstadt LM, Lisova AE, Babayeva ND, Lusk BL, Lim CJ, Tahirov TH (2023) Structures of human primosome elongation complexes. *Nat Struct Mol Biol* 30(5):579–583. <https://doi.org/10.1038/s41594-023-00971-3>
- He Q, Lim CJ (2023) Models for human telomere C-strand fill-in by CST-polalpha-primase. *Trends Biochem Sci* 48:860. <https://doi.org/10.1016/j.tibs.2023.07.008>
- He Q, Lin X, Chavez BL, Agrawal S, Lusk BL, Lim CJ (2022a) Structures of the human CST-polalpha-primase complex bound to telomere templates. *Nature* 608(7924):826–832. <https://doi.org/10.1038/s41586-022-05040-1>
- He Y, Feigon J (2022) Telomerase structural biology comes of age. *Curr Opin Struct Biol* 76:102446. <https://doi.org/10.1016/j.sbi.2022.102446>
- He Y, Song H, Chan H, Liu B, Wang Y, Susac L, Zhou ZH, Feigon J (2022b) Structure of Tetrahymena telomerase-bound CST with polymerase alpha-primase. *Nature* 608(7924):813–818. <https://doi.org/10.1038/s41586-022-04931-7>
- Holzer S, Rzechorzek NJ, Short IR, Jenkyn-Bedford M, Pellegrini L, Kilkenny ML (2019) Structural basis for inhibition of human primase by arabinofuranosyl nucleoside analogues fludarabine and vidarabine. *ACS Chem Biol* 14(9):1904–1912. <https://doi.org/10.1021/acscchembio.9b00367>
- Hom RA, Wuttke DS (2017) Human CST prefers G-rich but not necessarily telomeric sequences. *Biochemistry* 56(32):4210–4218. <https://doi.org/10.1021/acs.biochem.7b00584>
- Hurwitz J, Dean FB, Kwong AD, Lee SH (1990) The in vitro replication of DNA containing the SV40 origin. *J Biol Chem* 265(30):18043–18046
- Jafri MA, Ansari SA, Alqahtani MH, Shay JW (2016) Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med* 8(1):69. <https://doi.org/10.1186/s13073-016-0324-x>
- Jansson LI, Hentschel J, Parks JW, Chang TR, Lu C, Baral R, Bagshaw CR, Stone MD (2019) Telomere DNA G-quadruplex folding within actively extending human telomerase. *Proc Natl Acad Sci USA* 116(19):9350–9359. <https://doi.org/10.1073/pnas.1814777116>
- Jiang J, Chan H, Cash DD, Miracco EJ, Ogorzalek Loo RR, Upton HE, Cascio D, O'Brien Johnson R, Collins K, Loo JA, Zhou ZH, Feigon J (2015) Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science* 350(6260):aab4070. <https://doi.org/10.1126/science.aab4070>
- Jiang J, Miracco EJ, Hong K, Eckert B, Chan H, Cash DD, Min B, Zhou ZH, Collins K, Feigon J (2013) The architecture of Tetrahymena telomerase holoenzyme. *Nature* 496(7444):187–192. <https://doi.org/10.1038/nature12062>
- Jiang J, Wang Y, Susac L, Chan H, Basu R, Zhou ZH, Feigon J (2018) Structure of telomerase with Telomeric DNA. *Cell* 173(5):1179–1190. e1113. <https://doi.org/10.1016/j.cell.2018.04.038>
- Kasbek C, Wang F, Price CM (2013) Human TEN1 maintains telomere integrity and functions in genome-wide replication restart. *J Biol Chem* 288(42):30139–30150. <https://doi.org/10.1074/jbc.M113.493478>
- Kelich JM, Papaioannou H, Skordalakes E (2021) Pol alpha-primase dependent nuclear localization of the mammalian CST complex. *Commun Biol* 4(1):349. <https://doi.org/10.1038/s42003-021-01845-4>
- Kilkenny ML, Longo MA, Perera RL, Pellegrini L (2013) Structures of human primase reveal design of nucleotide elongation site and mode of pol alpha tethering. *Proc Natl Acad Sci USA* 110(40):15961–15966. <https://doi.org/10.1073/pnas.1311185110>

- Kim C, Snyder RO, Wold MS (1992) Binding properties of replication protein a from human and yeast cells. *Mol Cell Biol* 12(7):3050–3059. <https://doi.org/10.1128/mcb.12.7.3050-3059.1992>
- Kuchta RD, Reid B, Chang LM (1990) DNA primase. Processivity and the primase to polymerase alpha activity switch. *J Biol Chem* 265(27):16158–16165
- Kuchta RD, Stengel G (2010) Mechanism and evolution of DNA primases. *Biochim Biophys Acta* 1804(5):1180–1189. <https://doi.org/10.1016/j.bbapap.2009.06.011>
- Lao Y, Lee CG, Wold MS (1999) Replication protein a interactions with DNA. 2. Characterization of double-stranded DNA-binding/helix-destabilization activities and the role of the zinc-finger domain in DNA interactions. *Biochemistry* 38(13):3974–3984. <https://doi.org/10.1021/bi982371m>
- Laprade H, Querido E, Smith MJ, Guerit D, Crimmins H, Conomos D, Pourret E, Chartrand P, Sfeir A (2020) Single-molecule imaging of telomerase RNA reveals a recruitment-retention model for telomere elongation. *Mol Cell* 79:115. <https://doi.org/10.1016/j.molcel.2020.05.005>
- Lattrick CM, Cech TR (2010) POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. *EMBO J* 29(5):924–933. <https://doi.org/10.1038/emboj.2009.409>
- Lei M, Podell ER, Cech TR (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol* 11(12):1223–1229. <https://doi.org/10.1038/nsmb867>
- Lei M, Zaug AJ, Podell ER, Cech TR (2005) Switching human telomerase on and off with hPOT1 protein in vitro. *J Biol Chem* 280(21):20449–20456. <https://doi.org/10.1074/jbc.M502212200>
- Lim C, Barbour AT, Zaug AJ, Goodrich KJ, McKay AE, Wuttke DS, Cech TR (2020) The structure of human CST reveals a decameric assembly bound to telomeric DNA. *Science* 368(6495):1081–1085. <https://doi.org/10.1126/science.aaz9649>
- Lim C, Cech TR (2021) Shaping human telomeres: from shelterin and CST complexes to telomeric chromatin organization. *Nat Rev Mol Cell Biol* 22(4):283–298. <https://doi.org/10.1038/s41580-021-00328-y>
- Lin JJ, Zakian VA (1996) The *Saccharomyces* CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior in vivo. *Proc Natl Acad Sci USA* 93(24):13760–13765. <https://doi.org/10.1073/pnas.93.24.13760>
- Liu B, He Y, Wang Y, Song H, Zhou ZH, Feigon J (2022) Structure of active human telomerase with telomere shelterin protein TPP1. *Nature* 604(7906):578–583. <https://doi.org/10.1038/s41586-022-04582-8>
- Loayza D, De Lange T (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* 423(6943):1013–1018. <https://doi.org/10.1038/nature01688>
- Lyu X, Lei KH, Biak Sang P, Shiva O, Chastain M, Chi P, Chai W (2021) Human CST complex protects stalled replication forks by directly blocking MRE11 degradation of nascent-strand DNA. *EMBO J* 40(2):e103654. <https://doi.org/10.15252/embj.2019103654>
- Maciejowski J, de Lange T (2017) Telomeres in cancer: tumour suppression and genome instability. *Nat Rev Mol Cell Bio* 18(3):175–186. <https://doi.org/10.1038/nrm.2016.171>
- Makarov VL, Hirose Y, Langmore JP (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88(5):657–666. [https://doi.org/10.1016/s0092-8674\(00\)81908-x](https://doi.org/10.1016/s0092-8674(00)81908-x)
- McElligott R, Wellinger RJ (1997) The terminal DNA structure of mammalian chromosomes. *EMBO J* 16(12):3705–3714. <https://doi.org/10.1093/emboj/16.12.3705>
- Meyne J, Ratliff RL, Moyzis RK (1989) Conservation of the human telomere sequence (TTAGGG) n among vertebrates. *Proc Natl Acad Sci USA* 86(18):7049–7053. <https://doi.org/10.1073/pnas.86.18.7049>
- Mirman Z, Lottersberger F, Takai H, Kibe T, Gong Y, Takai K, Bianchi A, Zimmermann M, Durocher D, de Lange T (2018) 53BP1-RIF1-shieldin counteracts DSB resection through CST- and polalpha-dependent fill-in. *Nature* 560(7716):112–116. <https://doi.org/10.1038/s41586-018-0324-7>

- Miyake Y, Nakamura M, Nabetani A, Shimamura S, Tamura M, Yonehara S, Saito M, Ishikawa F (2009) RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. *Mol Cell* 36(2):193–206. <https://doi.org/10.1016/j.molcel.2009.08.009>
- Morin GB (1989) The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* 59(3):521–529. [https://doi.org/10.1016/0092-8674\(89\)90035-4](https://doi.org/10.1016/0092-8674(89)90035-4)
- Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL, Wu JR (1988) A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci USA* 85(18):6622–6626. <https://doi.org/10.1073/pnas.85.18.6622>
- Mullins EA, Salay LE, Durie CL, Bradley NP, Jackman JE, Ohi MD, Chazin WJ, Eichman BF (2023) A mechanistic model of primer synthesis from catalytic structures of DNA polymerase α -primase. *Nat Struct Mol Biol* (2024). <https://doi.org/10.1038/s41594-024-01227-4>
- Nakaoka H, Nishiyama A, Saito M, Ishikawa F (2012) *Xenopus laevis* Ctc1-Stn1-Ten1 (xCST) protein complex is involved in priming DNA synthesis on single-stranded DNA template in *Xenopus* egg extract. *J Biol Chem* 287(1):619–627. <https://doi.org/10.1074/jbc.M111.263723>
- Nandakumar J, Bell CF, Weidenfeld I, Zaug AJ, Leinwand LA, Cech TR (2012) The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity. *Nature* 492(7428):285–289. <https://doi.org/10.1038/nature11648>
- Nasheuer HP, Grosse F (1988) DNA polymerase α -primase from calf thymus. Determination of the polypeptide responsible for primase activity. *J Biol Chem* 263(18):8981–8988
- Nguyen THD (2021) Structural biology of human telomerase: progress and prospects. *Biochem Soc Trans* 49(5):1927–1939. <https://doi.org/10.1042/BST20200042>
- Nguyen THD, Tam J, Wu RA, Greber BJ, Toso D, Nogales E, Collins K (2018) Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* 557(7704):190–195. <https://doi.org/10.1038/s41586-018-0062-x>
- Nozawa K, Suzuki M, Takemura M, Yoshida S (2000) In vitro expansion of mammalian telomere repeats by DNA polymerase α -primase. *Nucleic Acids Res* 28(16):3117–3124. <https://doi.org/10.1093/nar/28.16.3117>
- Nugent CI, Hughes TR, Lue NF, Lundblad V (1996) Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science* 274(5285):249–252. <https://doi.org/10.1126/science.274.5285.249>
- Nunez-Ramirez R, Klinge S, Sauguet L, Melero R, Recuero-Checa MA, Kilkenny M, Perera RL, Garcia-Alvarez B, Hall RJ, Nogales E, Pellegrini L, Llorca O (2011) Flexible tethering of primase and DNA pol α in the eukaryotic primosome. *Nucleic Acids Res* 39(18):8187–8199. <https://doi.org/10.1093/nar/gkr534>
- O'Sullivan RJ, Karlseder J (2010) Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol* 11(3):171–181. <https://doi.org/10.1038/nrm2848>
- Olovnikov AM (1996) Telomeres, telomerase, and aging: origin of the theory. *Exp Gerontol* 31(4):443–448. [https://doi.org/10.1016/0531-5565\(96\)00005-8](https://doi.org/10.1016/0531-5565(96)00005-8)
- Olson CL, Barbour AT, Wieser TA, Wuttke DS (2023) RPA engages telomeric G-quadruplexes more effectively than CST. *Nucleic Acids Res* 51(10):5073–5086. <https://doi.org/10.1093/nar/gkad315>
- Pellegrini L (2012) The pol α -primase complex. *Subcell Biochem* 62:157–169. https://doi.org/10.1007/978-94-007-4572-8_9
- Pennock E, Buckley K, Lundblad V (2001) Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell* 104(3):387–396. [https://doi.org/10.1016/S0092-8674\(01\)00226-4](https://doi.org/10.1016/S0092-8674(01)00226-4)
- Pokhrel N, Caldwell CC, Corless EI, Tillison EA, Tibbs J, Jovic N, Tabei SMA, Wold MS, Spies M, Antony E (2019) Dynamics and selective remodeling of the DNA-binding domains of RPA. *Nat Struct Mol Biol* 26(2):129–136. <https://doi.org/10.1038/s41594-018-0181-y>
- Rice C, Skordalakes E (2016) Structure and function of the telomeric CST complex. *Comput Struct Biotechnol J* 14:161–167. <https://doi.org/10.1016/j.csbj.2016.04.002>

- Riethman H (2008) Human telomere structure and biology. *Annu Rev Genomics Hum Genet* 9:1–19. <https://doi.org/10.1146/annurev.genom.8.021506.172017>
- Roake CM, Artandi SE (2020) Regulation of human telomerase in homeostasis and disease. *Nat Rev Mol Cell Biol* 21:384. <https://doi.org/10.1038/s41580-020-0234-z>
- Santocanale C, Foiani M, Lucchini G, Plevani P (1993) The isolated 48,000-Dalton subunit of yeast DNA primase is sufficient for RNA primer synthesis. *J Biol Chem* 268(2):1343–1348
- Sauerwald A, Sandin S, Cristofari G, Scheres SH, Lingner J, Rhodes D (2013) Structure of active dimeric human telomerase. *Nat Struct Mol Biol* 20(4):454–460. <https://doi.org/10.1038/nsmb.2530>
- Schmidt JC, Cech TR (2015) Human telomerase: biogenesis, trafficking, recruitment, and activation. *Genes Dev* 29(11):1095–1105. <https://doi.org/10.1101/gad.263863.115>
- Schmidt JC, Dalby AB, Cech TR (2014) Identification of human TERT elements necessary for telomerase recruitment to telomeres. *elife* 3. <https://doi.org/10.7554/eLife.03563>
- Schmidt JC, Zaugg AJ, Cech TR (2016) Live cell imaging reveals the dynamics of telomerase recruitment to telomeres. *Cell* 166(5):1188–1197. e1189. <https://doi.org/10.1016/j.cell.2016.07.033>
- Schmidt JC, Zaugg AJ, Kufer R, Cech TR (2018) Dynamics of human telomerase recruitment depend on template- telomere base pairing. *Mol Biol Cell* 29:869. <https://doi.org/10.1091/mbc.E17-11-0637>
- Sekne Z, Ghanim GE, van Roon AM, Nguyen THD (2022) Structural basis of human telomerase recruitment by TPP1-POT1. *Science* 375(6585):1173–1176. <https://doi.org/10.1126/science.abn6840>
- Sexton AN, Youmans DT, Collins K (2012) Specificity requirements for human telomere protein interaction with telomerase holoenzyme. *J Biol Chem* 287(41):34455–34464. <https://doi.org/10.1074/jbc.M112.394767>
- Shay JW, Wright WE (2000) Hayflick, his limit, and cellular ageing. *Nat Rev Mol Cell Biol* 1(1):72–76. <https://doi.org/10.1038/35036093>
- Sheaff RJ, Kuchta RD (1993) Mechanism of calf thymus DNA primase: slow initiation, rapid polymerization, and intelligent termination. *Biochemistry* 32(12):3027–3037. <https://doi.org/10.1021/bi00063a014>
- Sheaff RJ, Kuchta RD, Ilsley D (1994) Calf thymus DNA polymerase alpha-primase: “communication” and primer-template movement between the two active sites. *Biochemistry* 33(8):2247–2254. <https://doi.org/10.1021/bi00174a035>
- Smith EM, Pendlebury DF, Nandakumar J (2020) Structural biology of telomeres and telomerase. *Cell Mol Life Sci* 77(1):61–79. <https://doi.org/10.1007/s00018-019-03369-x>
- Smogorzewska A, de Lange T (2002) Different telomere damage signaling pathways in human and mouse cells. *EMBO J* 21(16):4338–4348. <https://doi.org/10.1093/emboj/cdf433>
- Spiegel J, Adhikari S, Balasubramanian S (2020) The structure and function of DNA G-Quadruplexes. *Trends Chem* 2(2):123–136. <https://doi.org/10.1016/j.trechm.2019.07.002>
- Stewart JA, Chaiken MF, Wang F, Price CM (2012a) Maintaining the end: roles of telomere proteins in end-protection, telomere replication and length regulation. *Mutat Res* 730(1–2):12–19. <https://doi.org/10.1016/j.mrfmmm.2011.08.011>
- Stewart JA, Miller AS, Campbell JL, Bambara RA (2008) Dynamic removal of replication protein a by Dna2 facilitates primer cleavage during Okazaki fragment processing in *Saccharomyces cerevisiae*. *J Biol Chem* 283(46):31356–31365. <https://doi.org/10.1074/jbc.M805965200>
- Stewart JA, Wang F, Chaiken MF, Kasbek C, Chastain PD 2nd, Wright WE, Price CM (2012b) Human CST promotes telomere duplex replication and general replication restart after fork stalling. *EMBO J* 31(17):3537–3549. <https://doi.org/10.1038/emboj.2012.215>
- Stillman B (1989) Initiation of eukaryotic DNA replication in vitro. *Annu Rev Cell Biol* 5:197–245. <https://doi.org/10.1146/annurev.cb.05.110189.001213>
- Sundquist WI, Klug A (1989) Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* 342(6251):825–829. <https://doi.org/10.1038/342825a0>
- Surovtseva YV, Churikov D, Boltz KA, Song X, Lamb JC, Warrington R, Leehy K, Heacock M, Price CM, Shippen DE (2009) Conserved telomere maintenance component 1 interacts with

- STN1 and maintains chromosome ends in higher eukaryotes. *Mol Cell* 36(2):207–218. <https://doi.org/10.1016/j.molcel.2009.09.017>
- Suzuki M, Izuta S, Savoyesky E, Sakurai T, Simbulan C, Tatebe M, Kojima K, Yoshida S (1993) Deoxypyrimidine cluster mediates the priming by calf thymus DNA primase subunit. *Biochem Mol Biol Int* 29(4):645–652
- Takai H, Aria V, Borges P, Yeeles JTP, Lange T (2024) CST–Polymerase α -primase solves a second telomere end-replication problem. *Nature* 627:664–670. <https://doi.org/10.1038/s41586-024-07137-1>
- Tesmer VM, Brenner KA, Nandakumar J (2023) Human POT1 protects the telomeric ds-ss DNA junction by capping the 5' end of the chromosome. *Science* 381(6659):771–778. <https://doi.org/10.1126/science.adi2436>
- Vaithiyalingam S, Arnett DR, Aggarwal A, Eichman BF, Fanning E, Chazin WJ (2014) Insights into eukaryotic primer synthesis from structures of the p48 subunit of human DNA primase. *J Mol Biol* 426(3):558–569. <https://doi.org/10.1016/j.jmb.2013.11.007>
- Wan F, Ding Y, Zhang Y, Wu Z, Li S, Yang L, Yan X, Lan P, Li G, Wu J, Lei M (2021) Zipper head mechanism of telomere synthesis by human telomerase. *Cell Res* 31(12):1275–1290. <https://doi.org/10.1038/s41422-021-00586-7>
- Wan M, Qin J, Songyang Z, Liu D (2009) OB fold-containing protein 1 (OBFC1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation. *J Biol Chem* 284(39):26725–26731. <https://doi.org/10.1074/jbc.M109.021105>
- Wang F, Podell ER, Zaug AJ, Yang Y, Baciú P, Cech TR, Lei M (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* 445(7127):506–510. <https://doi.org/10.1038/nature05454>
- Wang F, Stewart JA, Kasbek C, Zhao Y, Wright WE, Price CM (2012) Human CST has independent functions during telomere duplex replication and C-strand fill-in. *Cell Rep* 2(5):1096–1103. <https://doi.org/10.1016/j.celrep.2012.10.007>
- Wang H, Ma T, Zhang X, Chen W, Lan Y, Kuang G, Hsu SJ, He Z, Chen Y, Stewart J, Bhattacharjee A, Luo Z, Price C, Feng X (2023) CTC1 OB-B interaction with TPP1 terminates telomerase and prevents telomere overextension. *Nucleic Acids Res* 51(10):4914–4928. <https://doi.org/10.1093/nar/gkad237>
- Williamson JR (1994) G-quartet structures in telomeric DNA. *Annu Rev Biophys Biomol Struct* 23:703–730. <https://doi.org/10.1146/annurev.bb.23.060194.003415>
- Williamson JR, Raghuraman MK, Cech TR (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* 59(5):871–880. [https://doi.org/10.1016/0092-8674\(89\)90610-7](https://doi.org/10.1016/0092-8674(89)90610-7)
- Wold MS (1997) Replication protein a: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* 66:61–92. <https://doi.org/10.1146/annurev.biochem.66.1.61>
- Wold MS, Kelly T (1988) Purification and characterization of replication protein a, a cellular protein required for in vitro replication of simian virus 40 DNA. *Proc Natl Acad Sci USA* 85(8):2523–2527. <https://doi.org/10.1073/pnas.85.8.2523>
- Wold MS, Weinberg DH, Virshup DM, Li JJ, Kelly TJ (1989) Identification of cellular proteins required for simian virus 40 DNA replication. *J Biol Chem* 264(5):2801–2809
- Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev* 11(21):2801–2809. <https://doi.org/10.1101/gad.11.21.2801>
- Wu P, Takai H, de Lange T (2012) Telomeric 3' overhangs derive from resection by Exo1 and Apollo and fill-in by POT1b-associated CST. *Cell* 150(1):39–52. <https://doi.org/10.1016/j.cell.2012.05.026>
- Wu RA, Upton HE, Vogan JM, Collins K (2017) Telomerase mechanism of telomere synthesis. *Annu Rev Biochem* 86:439–460. <https://doi.org/10.1146/annurev-biochem-061516-045019>
- Xie P (2011) A model for dynamics of primer extension by eukaryotic DNA primase. *Eur Biophys J* 40(10):1157–1165. <https://doi.org/10.1007/s00249-011-0746-8>

- Yan J, Holzer S, Pellegrini L, Bell SD (2018) An archaeal primase functions as a nanoscale caliper to define primer length. *Proc Natl Acad Sci USA* 115(26):6697–6702. <https://doi.org/10.1073/pnas.1806351115>
- Yin Z, Kilkenny ML, Ker D-S, Pellegrini L (2024) CryoEM insights into RNA primer synthesis by the human primosome. *FEBS J* 291(8):1813–1829. <https://doi.org/10.1111/febs.17082>
- Yuan Z, Georgescu R, Li H, O'Donnell ME (2023) Molecular choreography of primer synthesis by the eukaryotic pol alpha-primase. *Nat Commun* 14(1):3697. <https://doi.org/10.1038/s41467-023-39441-1>
- Zaug AJ, Goodrich KJ, Song JJ, Sullivan AE, Cech TR (2022) Reconstitution of a telomeric replicon organized by CST. *Nature* 608(7924):819–825. <https://doi.org/10.1038/s41586-022-04930-8>
- Zaug AJ, Lim CJ, Olson CL, Carilli MT, Goodrich KJ, Wuttke DS, Cech TR (2021) CST does not evict elongating telomerase but prevents initiation by ssDNA binding. *Nucleic Acids Res* 49(20):11653–11665. <https://doi.org/10.1093/nar/gkab942>
- Zerbe LK, Kuchta RD (2002) The p58 subunit of human DNA primase is important for primer initiation, elongation, and counting. *Biochemistry* 41(15):4891–4900. <https://doi.org/10.1021/bi016030b>
- Zhang M, Wang B, Li T, Liu R, Xiao Y, Geng X, Li G, Liu Q, Price CM, Liu Y, Wang F (2019) Mammalian CST averts replication failure by preventing G-quadruplex accumulation. *Nucleic Acids Res* 47(10):5243–5259. <https://doi.org/10.1093/nar/gkz264>
- Zhao Y, Abreu E, Kim J, Stadler G, Eskiocak U, Terns MP, Terns RM, Shay JW, Wright WE (2011) Processive and distributive extension of human telomeres by telomerase under homeostatic and nonequilibrium conditions. *Mol Cell* 42(3):297–307. <https://doi.org/10.1016/j.molcel.2011.03.020>
- Zhao Y, Sfeir AJ, Zou Y, Buseman CM, Chow TT, Shay JW, Wright WE (2009) Telomere extension occurs at most chromosome ends and is uncoupled from fill-in in human cancer cells. *Cell* 138(3):463–475. <https://doi.org/10.1016/j.cell.2009.05.026>
- Zhong FL, Batista LF, Freund A, Pech MF, Venteicher AS, Artandi SE (2012) TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. *Cell* 150(3):481–494. <https://doi.org/10.1016/j.cell.2012.07.012>

Chapter 6

The Core Complex of Yeast COMPASS and Human Mixed-Lineage Leukemia (MLL), Structure, Function, and Recognition of the Nucleosome



Uhn-Soo Cho

Abstract Yeast COMPASS (complex of proteins associated with Set1) and human MLL (mixed-lineage leukemia) complexes are histone H3 lysine 4 methyltransferases with critical roles in gene regulation and embryonic development. Both complexes share a conserved C-terminal SET domain, responsible for catalyzing histone H3 K4 methylation on nucleosomes. Notably, their catalytic activity toward nucleosomes is enhanced and optimized with assembly of auxiliary subunits. In this review, we aim to illustrate the recent X-ray and cryo-EM structures of yeast COMPASS and human MLL1 core complexes bound to either unmodified nucleosome core particle (NCP) or H2B mono-ubiquitinated NCP (H2Bub.NCP). We further delineate how each auxiliary component of the complex contributes to the NCP and ubiquitin recognition to maximize the methyltransferase activity.

Keywords COMPASS · MLL · SET domain · H3 Lys4 methyltransferase · Nucleosomal core particle (NCP) · H2B K120 (K123 in yeast) mono-ubiquitinated nucleosome (H2Bub.NCP) · X-ray crystallography · Single-particle electron cryomicroscopy (cryo-EM)

Introduction

After the discovery of the 11q23 leukemic translocation involving mixed-lineage leukemia (MLL), a human homolog of *Drosophila* Trithorax marked a significant turning point in leukemia research (Djabali et al. 1992; Gu et al. 1992). In acute myeloid leukemia (AML) and acute lymphoid leukemia (ALL), the MLL1 gene frequently undergoes oncogenic gene translocations, resulting in the expression of

U.-S. Cho (✉)

Department of Biological Chemistry, University of Michigan, Ann Arbor, MI, USA

e-mail: uhns00@med.umich.edu

© The Author(s), under exclusive license to Springer Nature

Switzerland AG 2024

J. R. Harris, J. Marles-Wright (eds.), *Macromolecular Protein Complexes V*, Subcellular Biochemistry 104, https://doi.org/10.1007/978-3-031-58843-3_6

chimeric fusion proteins composed of the N-terminal 1400 residues of MLL1 and one of over 90 partner genes (Meyer et al. 2006). MLL-associated leukemia accounts for over 70% of infant ALL cases and contributes significantly to the spectrum of infant AML cases, ranging from 35% to 50% (Chessells et al. 2002; Hilden et al. 2006; Pieters et al. 2007; Tomizawa et al. 2007; van der Linden et al. 2009).

The biological significance of *Drosophila* Trithorax and its yeast and human homologs, COMPASS and MLL, are intricately linked to their functions as histone H3 Lys4 methyltransferases (Nakamura et al. 2002; Sedkov et al. 2003). The methyltransferase activity is mediated by the SET domain located at the C-terminal end of the MLL, COMPASS, or Trx proteins, which specifically targets the side chain of histone H3 lysine 4 (H3K4) (Sedkov et al. 1999; Briggs et al. 2001; Miller et al. 2001; Roguev et al. 2001; Santos-Rosa et al. 2002; Guenther et al. 2005). In a highly regulated and non-processive manner, MLL-SET domain adds mono-, di-, and trimethyl groups to H3K4 using S-adenosylhomocysteine (AdoHcy or SAH) as a methyl donor (Milne et al. 2002; Dillon et al. 2005; Xiao et al. 2006).

H3K4 methylation (H3K4me) exhibits marked enrichment at transcriptional regulatory regions, encompassing the gene promoters and distal enhancers, where it exerts a pivotal influence on the recruitment of essential components of the basic transcription machinery (Shilatifard 2006; Vermeulen et al. 2007; Schuettengruber et al. 2011). Additionally, it orchestrates the recruitment of chromatin remodeling complexes (Nakamura et al. 2002), thereby contributing to complicated outcomes of gene expression regulation. The dynamic interplay between H3K4me and co-transcriptional processes has also been extensively documented (Sims 3rd and Reinberg 2006; Ruthenburg et al. 2007; Buratowski and Kim 2010; Vermeulen and Timmers 2010).

Human genetic studies have provided compelling evidence regarding the functional significance of MLL family enzymes. Heterozygous mutations in MLLs have been implicated in congenital human syndromes, including Kabuki, Wiedemann-Steiner, and Kleefstra spectrum syndromes (Hannibal et al. 2011; Jones et al. 2012; Kleefstra et al. 2012; Shen et al. 2014). Beyond congenital disorders, MLLs feature prominently among the most frequently mutated genes in a wide spectrum of human malignancies (Kandoth et al. 2013; Rao and Dou 2015).

While the SET domain is responsible for the methyltransferase activity of MLL and the COMPASS complex, it exhibits limited methyltransferase activity when acting alone, necessitating the involvement of additional components to achieve full functionality (Dehe et al. 2006; Dou et al. 2006; Steward et al. 2006; Wysocka et al. 2005; Dou et al. 2005; Patel et al. 2009). The C-terminal SET domain of the COMPASS core complex recruits Cps30/Swd3, Cps50/Swd1, Cps60/Bre2, and Cps25/Sdc1, while the MLL core complex incorporates WDR5, RbBP5, ASH2L, and DPY30, to compromise the catalytic efficiency. Therefore, the comprehensive exploration of the structure, function, and regulation of MLL family enzymes, along with a deep understanding of the core complex components, is inevitable to illustrate its underlying molecular mechanism and holds immense promise in deciphering their implications in human diseases. This review aims to provide a comprehensive structural overview of yeast COMPASS and human MLL complex and their

relationship in ubiquitination and nucleosome recognition, thereby shedding light on the multifaceted role of MLL in disease pathogenesis and epigenetic regulation.

The Catalytic SET Domain

The biological activity of histone H3 lysine 4 methyltransferase, represented by COMPASS in yeast and MLL1–4 and SETD1A/B in humans, primarily emanates from the carboxyl-terminal SET (Su(var)3–9, Enhancer-of-zeste, and Trithorax) domain. Previous research has demonstrated that the carboxyl-terminal SET domain, in isolation, exhibits full enzymatic activity both *in vivo* and *in vitro* when coupled with other essential core subunits (Dou et al. 2006; Steward et al. 2006; Patel et al. 2009; Cao et al. 2010; Takahashi et al. 2011). The MLL1-SET domain alone exhibits limited enzymatic activity, suggesting a dependence on additional core subunits to facilitate the comprehensive architectural reorganization required for full enzymatic activation.

The crystal structure of the human MLL1-SET domain, spanning residues 3785–3969, has been determined in conjunction with its cofactor, S-adenosylhomocysteine (AdoHcy or SAH), and a short histone H3 peptide, at a resolution of 2.2 Å (Fig. 6.1) (Southall et al. 2009). This structure encompasses N-flanking, SET-N, SET-I, SET-C, and postSET subdomains, consistent with the structural features observed in other known SET domains (*see* Fig. 6.1a) (Xiao et al.

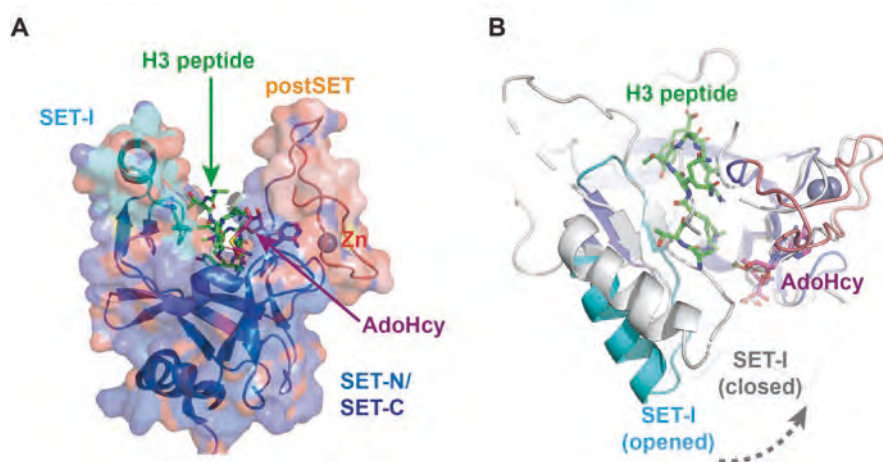


Fig. 6.1 The crystal structure of the human MLL1-SET domain (PDB ID: 2W5Z, residues 3785–3969). (a) Cartoon and surface representation of SET-N, SET-I, SET-C, and postSET subdomains are displayed together with the H3 peptide and S-adenosylhomocysteine (AdoHcy or SAH) (Southall et al. 2009). (b) Structural overlay of opened (cyan) and closed (gray) forms of the SET domain. The crystal structure of Dim-5 (PDB ID: 1PEG) represents the closed form of the SET domain (Zhang et al. 2003)

2003). The SET domain is known to adopt two distinct conformations, an open and a closed form (Xiao et al. 2003). In the case of the crystal structure of the MLL1-SET domain, it adopts an open conformation, characterized by the movement of the SET-I region helix away from the C-flanking region, as compared to the closed conformation observed in Dim-5 (Fig. 6.1b) (Zhang et al. 2003). Importantly, studies have revealed that the purified MLL1-SET domain exhibits minimal methyltransferase activity, potentially attributed to its open conformational state (Dou et al. 2006; Steward et al. 2006; Li et al. 2016). To attain full enzymatic activity, the MLL1-SET domain necessitates a conformational transition, shifting from the open form to the closed form, an alteration that is triggered by association with other essential core components (*see* Fig. 6.1b) (Southall et al. 2009; Li et al. 2016).

The Core Complex of Yeast COMPASS

In the MLL/COMPASS complex, the full activation of the methyltransferase activity associated with the C-terminal SET domain is achieved when this domain assembles with four auxiliary components: WDR5 (Cps30/Swd3 in yeast), RbBP5 (Cps50/Swd1), Ash2L (Cps60/Bre2), and DPY30 (Cps25/sdc1) (Fig. 6.2). This assembly is crucial for the complete activation of MLL/COMPASS methyltransferase activity. To comprehend how these auxiliary components collaborate to achieve this activation, it is imperative to uncover the molecular architecture of the MLL/COMPASS core complex.

Initial insights into the overall architecture of the MLL/COMPASS core complex were provided by Takahashi and colleagues, who employed immunolabeling and 2D analysis of single-particle negative-stain electron microscopy (neg-EM) (Takahashi et al. 2011). Their work revealed that both MLL1 and COMPASS core complexes share a Y-shaped architecture. In this structure, the catalytic SET domain resides in the central position, with RbBP5/Cps50/Swd1 and WDR5/Cps30/Swd3 forming two upper arms, and Ash2L/Cps60/Bre2 and DPY30/Cps25/sdc1 constituting the basement. Subsequently, the researchers attempted single-particle electron cryomicroscopy (sp cryo-EM) to visualize the 3D structure of the MLL/COMPASS core complex. However, due to technical challenges at that time, only low-resolution structures were obtained.

In 2018, two groups provided atomic resolution structures of the COMPASS core complex. The Skiniotis group reported the sp cryo-EM structure of the *Saccharomyces cerevisiae* COMPASS core complex, including the fifth component, Cps40/Spp1, at 4.0 Å resolution (Qu et al. 2018). Concurrently, the Zheng group released the X-ray crystal structure of the *Kluyveromyces lactis* COMPASS core complex at 3.0 Å resolution (Fig. 6.2b) (Hsu et al. 2018). Both structures reaffirmed the Y-shaped architecture previously observed. While the two structures closely resembled each other, the sp cryo-EM structure of *S. cerevisiae* COMPASS core complex included the fifth component, Cps40/Spp1, which is crucial for histone H3 tri-methylation in yeast.

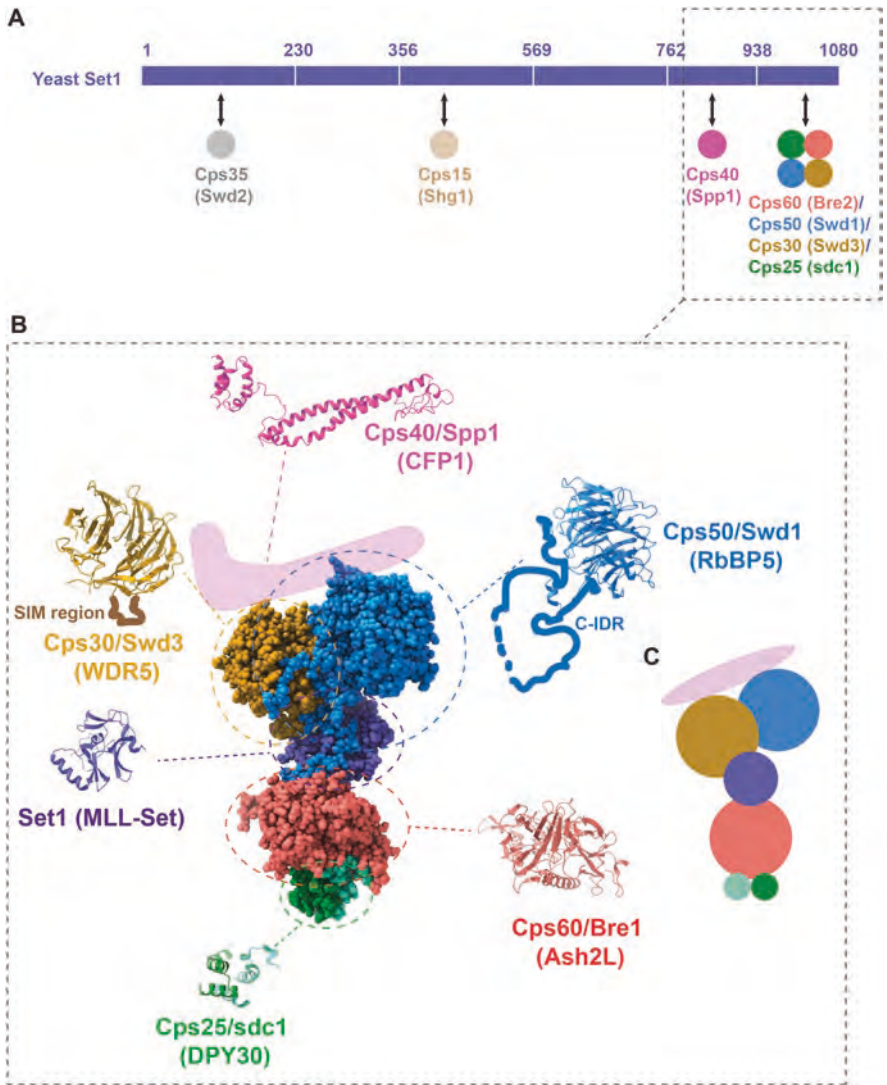


Fig. 6.2 The crystal structure of *K. lactis* COMPASS core complex (PDB ID: 6CHG) (Hsu et al. 2018). (a) Domain architecture of yeast Set1 and its binding partners. The Set domain is located at residues 938–1080. The dashed rectangular box contains the binding partners of the N-terminal extended Set domain (n-SET) and SET domain, which forms a core complex of COMPASS. (b) Surface representation and cartoon views of the *K. lactis* COMPASS core complex. Human counterparts are included in the parenthesis. The same color scheme of individual domains will be used for both yeast COMPASS and human MLL core subunits. Cps50/Swd1 possesses the C-terminal intrinsically disordered region (C-IDR), and Cps30/Swd3 harbors a SET-I interacting motif (SIM) loop. C, A schematic view of the yeast COMPASS core complex. The same color scheme is used for each core subunit of COMPASS and MLL

One of the most noteworthy structural features within the COMPASS core complex is Cps50/Swd1. Cps50/Swd1 contains WD-40 repeats along with short N-terminal and long C-terminal intrinsically disordered regions (IDRs). The C-terminal region of Cps50/Swd1 has a vital role in maintaining the overall core complex architecture by interacting with Cps30/Swd3, Set1, and Cps60/Bre2 through its C-terminal IDR (Fig. 6.1b) (Avdic et al. 2011). This C-terminal region can be further divided into two segments: a conserved segment for organizing Set1, Cps30/Swd3, and Cps60/Bre2 and a variable distal portion that reinforces the binding to Cps30/Swd3. Additionally, the short N-terminal region of Cps50/Swd1 interacts with Cps30/Swd3. Specific acidic residues (D375 and E376) in Cps50/Swd1 form contacts with the SET-I helix of the Set1 domain, and mutations of these residues disrupt the methyltransferase activity of the COMPASS complex significantly. This emphasizes the role of Cps50/Swd1 as a molecular scaffold, orchestrating the entire complex and optimizing the methyltransferase activity of the Set1 catalytic domain.

Cps30/Swd3, the second WD-40 repeat subunit within the COMPASS core complex, is essential for its integrity and function in yeast. Depletion of Cps30/Swd3 leads to defects in histone H3K4 methylation (Steward et al. 2006; Takahashi et al. 2011; Qu et al. 2018). Cps30/Swd3 includes a SET-I interacting motif (SIM) loop that connects β -strands 17 and 18 (Fig. 6.2a). Mutations of residues within this SIM region result in a significant reduction in COMPASS methyltransferase activity particularly for tri-methylation, indicating the importance of the SIM region in reorganizing the SET domain for maximal catalytic activity (Hsu et al. 2018; Qu et al. 2018). This is further supported by experiments involving a chimeric SET1 that disrupts SIM interaction, leading to a dramatic reduction in tri-methylation activity.

Cps60/Bre2, the yeast homologue of human Ash2L, comprises pre-SPRY, SPRY, and a C-terminal SDI (Sdc1 DPY-30 Interacting) domain (Hsu et al. 2018). Cps60/Bre2 forms a hetero complex with Cps25/Sdc1 homodimers via the C-terminal SDI domain, serving as the basement of the Y-shaped COMPASS complex (Fig. 6.2c). The absence of Cps60/Bre2 and Cps25/Sdc1 has a substantial impact, resulting in the complete loss of tri-methylation and an 80% reduction in mono-/di-methylation (Takahashi et al. 2011). The Cps60/Bre2 SPRY domain immobilizes the bottom part of the mobile SET-I region, allowing the SET domain to adopt a closed conformation, maximizing its methyltransferase activity.

Structural alignment of two yeast COMPASS core complexes indicates that the upper arms (Cps50/Swd1 and Cps30/Swd3) and the basement (Cps60/Bre2 and Cps25/Sdc1) exhibit a certain level of conformational dynamics around the SET domain (Hsu et al. 2018; Qu et al. 2018). This is supported by structural comparisons of two major classes of the COMPASS complex in cryo-EM data (Qu et al. 2018). Structural analysis with recent cryo-EM structures of the COMPASS and H2B K123 mono-ubiquitinated NCP (H2Bub.NCP) complex further suggests that the *K. lactis* COMPASS core complex might represent a more stable form (Hsu et al. 2019). In this review, we opt to utilize the X-ray structure to depict the COMPASS core complex due to its high-resolution characteristics, while the contribution of Cps40/Spp1 will be discussed later in the context of the

COMPASS-H2Bub.NCP complex structure (Hsu et al. 2018, 2019; Qu et al. 2018; Worden et al. 2020).

The H2Bub.NCP Recognition of the Yeast COMPASS Core Complex

In yeast, an interplay between two histone modifications, histone H3K4 methylation and H2B K123 mono-ubiquitination, orchestrates the positive regulation of gene expression (Dover et al. 2002; Sun and Allis 2002; Shilatifard 2006). Two distinct lysine to arginine mutations, H2B K123R and H3 K4R, exhibit disparate effects, demonstrating their unique contributions. H2B K123R exerts a profound influence by reducing H3K4 methylation significantly (Sun and Allis 2002). In contrast, H3 K4R demonstrates a tolerance for H2B K123 mono-ubiquitination, suggesting the impact of mono-ubiquitination of H2B K123 on the methyltransferase activity of the COMPASS complex in yeast.

To unravel the intricate mechanisms governing the recognition of the nucleosome core particle (NCP) and its H2B K123 mono-ubiquitinated counterpart (H2Bub.NCP) by the COMPASS core complex, two independent research groups recently reported the *sp* cryo-EM structures of the COMPASS core complex together with NCP or H2Bub.NCP (Hsu et al. 2019; Worden et al. 2020). Despite different yeast species, *S. cerevisiae* and *K. lactis*, these cryo-EM structures exhibit a resemblance in their overall architecture with a root mean square deviation (RMS deviation) of 1.17 Å. For consistency, we will focus on the *K. lactis* COMPASS-NCP and COMPASS-H2Bub.NCP structures for further discussion (Fig. 6.3).

One intriguing finding stems from the work of Hsu et al., wherein they reported a significant disparity in the methyltransferase activity between the recombinantly assembled five-subunit COMPASS complex (Cps50/Swd1-Cps30/Swd3-Cps60/Bre2-Cps25/Sdc1-Set1) and the endogenous COMPASS complex isolated from yeast cells (Hsu et al. 2019). While the recombinant complex exhibits robust methyltransferase activity toward NCPs, the endogenous counterpart attains full activation only in the presence of H2Bub.NCP as a substrate, despite no discernible differences in substrate affinity. This observation strongly suggests the existence of additional negative regulator(s) that modulate the methyltransferase activity of yeast COMPASS, in contrast to the human MLL complex. With two additional elements, the N-terminal extended Set domain (n-SET) and Cps40/Spp1, the recombinant COMPASS core complex mirrors the stimulatory effects observed in the endogenous COMPASS complex (Fig. 6.3c) (Hsu et al. 2019). This underscores the roles of n-SET and Cps40/Spp1 as critical elements conferring sensitivity to H2Bub.NCP in the yeast COMPASS complex (Kim et al. 2013), both of which are not present in the human MLL core complex.

Not surprisingly, the *sp* cryo-EM structures of the COMPASS-NCP and COMPASS-H2Bub.NCP complexes were determined in conjunction with n-SET

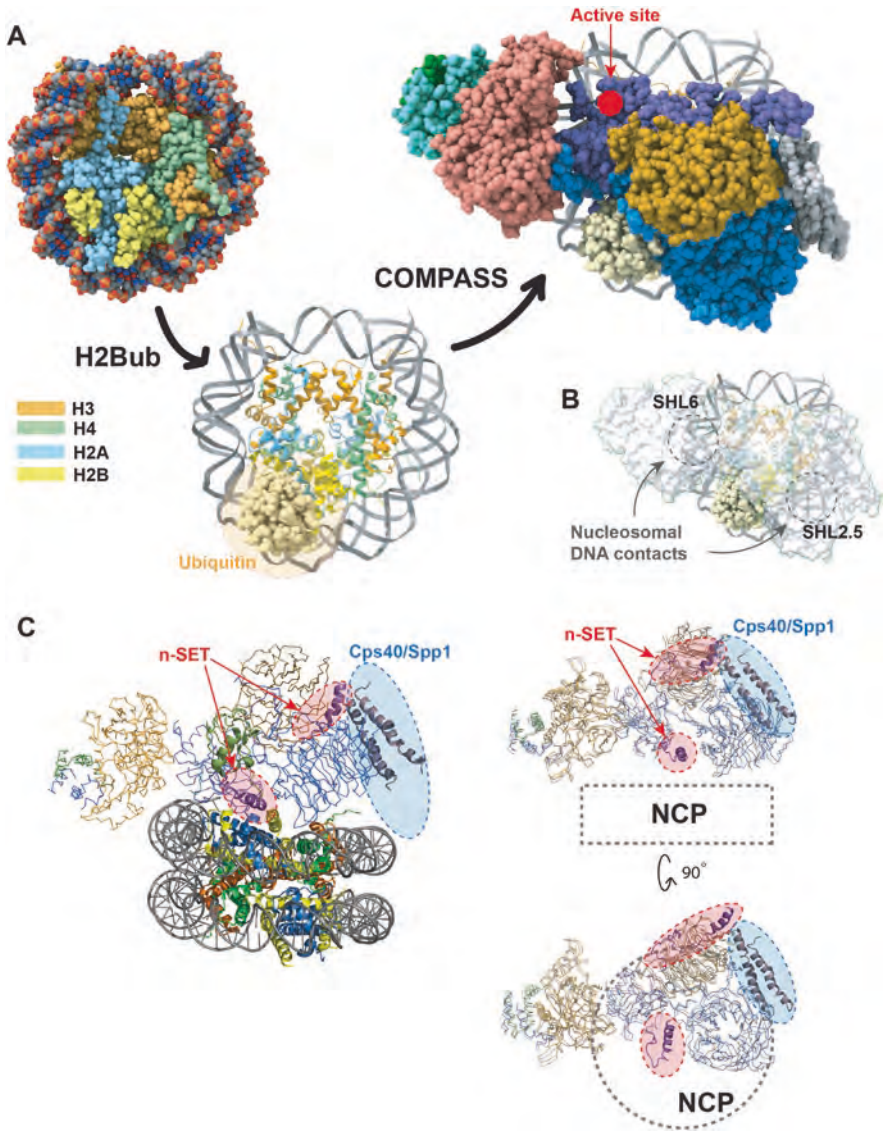


Fig. 6.3 The cryo-EM structures of *K. lactis* COMPASS core complex bound to the NCP or H2Bub.NCP (PDB ID: 6UGM, 6UH5) (Hsu et al. 2019). (a) The NCP is first mono-ubiquitinated at H2B K123 and the COMPASS core complex recognizes H2Bub.NCP. The active site of the Set domain is marked as red circle. (b) Two dashed circles indicate the nucleosomal DNA contact points by the COMPASS core complex when it recognizes H2Bub.NCP. The COMPASS core complex makes major contacts at super-helical locations (SHLs) 2.5 and 6 of the nucleosomal DNA. (c) Two additional components of the COMPASS core complex (n-SET with red dashed circles and Cps40/Spp1 with blue dashed circle) that may contribute to the H2B-monoubiquitination-specific nucleosomal recognition of the COMPASS core complex

and Cps40/Spp1 (Figs. 6.3 and 6.4) (Hsu et al. 2019). Notably, structural alignment with the COMPASS complex alone reveals that the incorporation of these two additional components, along with the recognition of NCP or H2Bub.NCP, does not induce any discernible conformational changes.

Detailed examination of the COMPASS-NCP and COMPASS-H2Bub.NCP complex structures unveils that six core components of COMPASS engage with the nucleosome, identifying specific structural loci, designated as super-helical

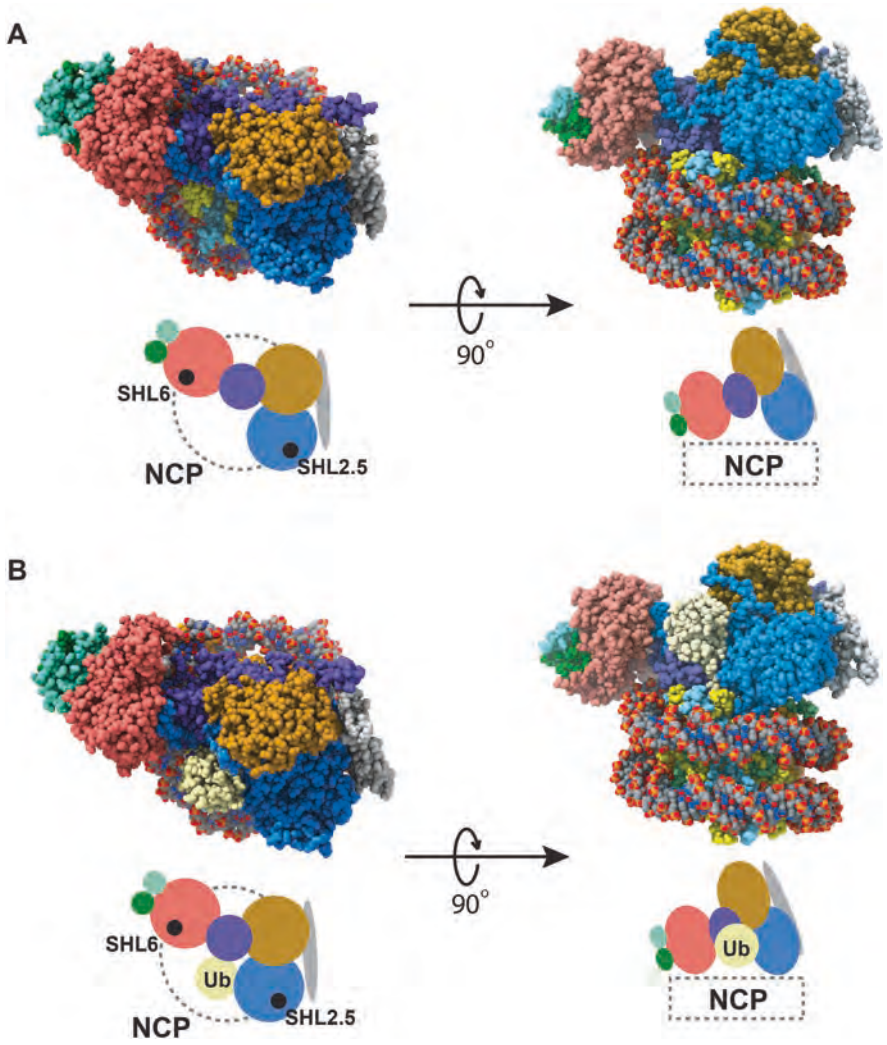


Fig. 6.4 The top and side views of the *K. lactis* COMPASS core complex bound to either NCP (a, PDB ID: 6UGM) or H2Bub.NCP (b, PDB ID: 6UH5). Corresponding schematic views are also displayed

locations (SHLs), for nucleosome recognition (Figs. 6.3 and 6.4). The initial contact point is orchestrated by Cps40/Spp1 and Cps50/Swd1 with the nucleosome at a defined SHL 2.5. In this contact point, Cps50/Swd1 emerges as a central feature, establishing contacts with all four histones, nucleosomal DNA, and ubiquitin (Fig. 6.4b). This underscores the critical role played by Cps50/Swd1 in NCP recognition, in addition to maintaining the integrity of the COMPASS core complex, as previously discussed. The second contacting point of NCP recognition unfolds as Cps60/Bre2 binds to the nucleosome at SHL 6. Interestingly, the catalytic Set1 domain refrains from direct interactions with H2A/H2B but assumes an architectural role, forming an arch-like structure over the NCP. This conformation is maintained through the participation of two WD40 domains and Cps60/Bre2. Additionally, the N-terminal helix of Set1 (ARM helix; residues 819–836) occupies a strategic position at the interface between Cps50/Swd1 and Set1 within the COMPASS complex, engaging with ubiquitin and histones H2A/H2B from the NCP (Fig. 6.3c). This arrangement serves to enhance the recognition of COMPASS toward H2Bub.NCP. Moreover, in the context of the COMPASS-H2Bub.NCP complex, Cps50/Swd1 establishes additional contacts with ubiquitin, facilitated by specific hydrophobic residues within ubiquitin (Ile44-Val70-Leu8) and Cps50/Swd1 (Phe10-Leu13-Val409) (Fig. 6.4b) (Hsu et al. 2019).

These structural insights unveil the intricacies of the orchestration of interactions governing the recognition of NCP and H2Bub.NCP by the COMPASS core complex, providing a deeper understanding of the underlying molecular mechanisms driving this essential epigenetic process.

Human MLL1 Core Complex

The human MLL1 core complex comprises five essential components, MLL1^{SET}, WDR5, ASH2L, RbBP5, and DPY30, collectively referred to the MLL1^{SET}-WARD complex. Recent years have witnessed advances in our understanding of this complex, with several research groups elucidating cryo-EM structures of the MLL1 complex bound to nucleosomes either with or without H2B K120 mono-ubiquitinated (Park et al. 2019; Xue et al. 2019; Lee et al. 2021; Ayoub et al. 2022; Rahman et al. 2022). Xue and colleagues reported cryo-EM structures of the human MLL1 core complex in association with the NCP and H2Bub.NCP (NCP bearing mono-ubiquitination of H2B K120), as well as the MLL3 core complex in similar association with the NCP and H2Bub.NCP (Xue et al. 2019). In parallel, our group contributed the cryo-EM structures of the human MLL1 core complex bound to the NCP (Park et al. 2019). Furthermore, we presented the cryo-EM structure of the human MLL1^{SET}-WARD complex bound to the NCP, in the absence of DPY30 (Lee et al. 2021). Notably, the Wolberger group further contributed key insights by reporting the cryo-EM structure of the MLL1^{SET}-WARD complex in association with the H2Bub.NCP (Rahman et al. 2022).

While the overall architecture of the MLL1^{SET}-WARD core complex bears a resemblance to that of the yeast COMPASS complex, characterized by its Y-shaped

conformation (Fig. 6.5), there exist notable distinctions in how the MLL1^{SET}-WARD complex recognizes the NCP, setting it apart from its yeast counterpart. Firstly, the MLL1 core complexes do not need the Cps40/Spp1 equivalent, which necessitates the yeast COMPASS core complex for NCP and H2Bub.NCP recognition. Secondly, the human MLL1 core complex demonstrates the capability to engage with the NCP in two distinct conformations, denoted as conformation 1 (Fig. 6.5a) and conformation 2 (Fig. 6.5b) (Park et al. 2019; Xue et al. 2019; Ayoub et al. 2022). This divergence in nucleosome recognition mechanisms between yeast COMPASS and human MLL1 complexes was expected. The multiple sequence alignment of RbBP5 shows that key nucleosome recognition residues of human RbBP5_(205-KRK-207) are not conserved in yeast CPS50/Swd1 (Park et al. 2019). Conformation 1 of the MLL1 core-NCP complex closely resembles the interaction observed in the yeast COMPASS complex (Figs. 6.4 and 6.5a). Particularly, the presence of H2Bub populates conformation 1 of the MLL1 core complex (Fig. 6.5c). However, a notable distinction emerges—unlike the yeast COMPASS complex, which requires H2B mono-ubiquitination of the NCP for optimal catalytic activity (Hsu et al. 2019),

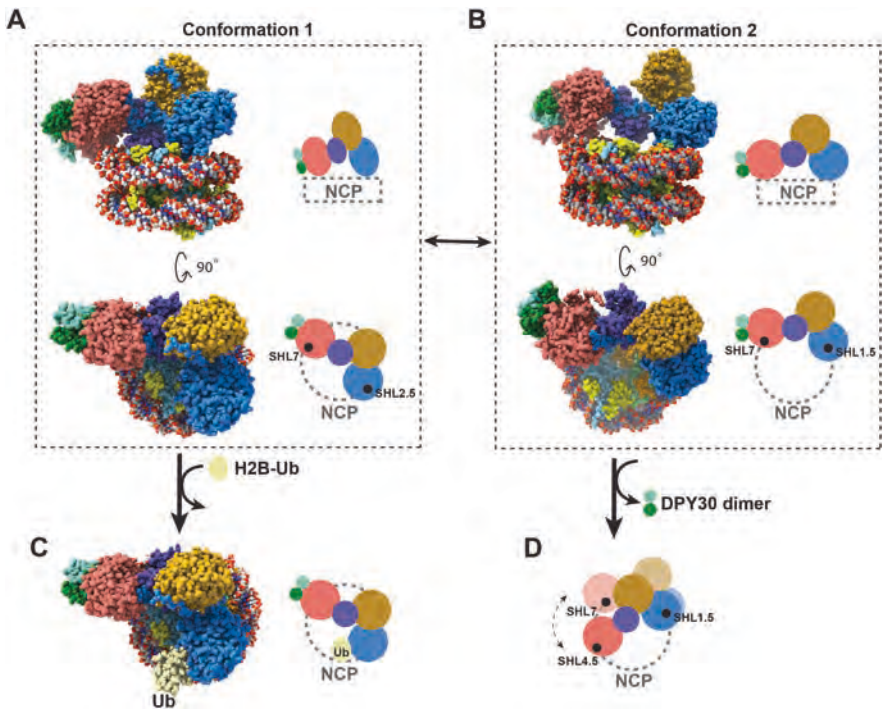


Fig. 6.5 The cryo-EM structures of human MLL1 core complex bound to NCP (PDB ID: 6PWV) (Park et al. 2019) and H2Bub.NCP (PDB ID: 7UD5) (Rahman et al. 2022). When the human MLL1 core complex recognizes the NCP, it adopts two different conformations: conformation 1 (a) and conformation 2 (b). Once the NCP is mono-ubiquitinated at H2B K120, the MLL1 core complex prefers the conformation 1 (c). When DPY30 dimer dissociates from the MLL1 core complex, the nucleosomal DNA recognition site of ASH2L oscillates in between SHL 4.5 and 7 (d)

mono-ubiquitination of the human NCP only modestly enhances the methyltransferase activity of the MLL1 core complex by twofold (Xue et al. 2019). In vitro assays indicate that the MLL1-WARD core complex can perform mono-, di-, and tri-methylation on the NCP in the absence of H2B.ub (Fig. 6.5b) (Park et al. 2019). This is probably related with earlier discovery on COMPASS-H2Bub.NCP showing that the n-SET/Spp1 module acts as an inhibitory element toward the unmodified NCP, which is not present in the human MLL1 core complex (Kim et al. 2013; Hsu et al. 2019).

Two major anchoring points for the MLL1-WARD complex predominantly involve RbBP5 and ASH2L. The nucleosomal DNA recognition capability of ASH2L is notably augmented in the presence of DPY30, which plays a role in reorganizing ASH2L for optimal nucleosomal DNA interaction (Lee et al. 2021). In the absence of DPY30, RbBP5 becomes a major anchoring point at SHL 1.5 and ASH2L unstably positions in between SHL 4.5 and SHL 7 (Fig. 6.5d) (Yang et al. 2016; Lee et al. 2021). Once ASH2L stably anchors at SHL 7 (together with the DPY30 dimer), RbBP5-binding site can be switched in two different positions, conformation 1 at SHL 2.5 and conformation 2 at SHL 1. In vitro methyltransferase assays employing NCP as substrates (not H2Bub.NCP), coupled with mutational analyses, have provided compelling evidence that conformation 2 of the MLL1-WARD complex exhibits enhanced catalytic activity toward the unmodified NCP (Park et al. 2019). This observation underscores the limited contribution of the ubiquitin domain in facilitating the stable anchoring of the MLL1-WARD complex onto the NCP, demonstrated by the dynamic nature of the ubiquitin and MLL1-WARD interactions (Xue et al. 2019). Notably, the coulomb density of the ubiquitin domain is notably weak in MLL1-WARD-H2Bub.NCP complex (Xue et al. 2019; Rahman et al. 2022), different from that of the yeast COMPASS-H2Bub.NCP complex (Hsu et al. 2019; Worden et al. 2020).

Conclusion and Outlook

As depicted above, the recent X-ray and cryo-EM structures of yeast COMPASS and human MLL complex bound to NCP and H2Bub.NCP have provided new insights on how the methyltransferase activity of the SET domain is potentiated through interactions with core auxiliary subunits, the NCP, and H2B mono-ubiquitinated NCP at the molecular level.

Nevertheless, a comprehensive understanding of the structural contribution of the N-terminal regions of yeast Set1 and human MLL proteins remains elusive. These regions encompass combinations of structural domains that are responsible for recognizing specific target sites and orchestrating their functions in concert with the C-terminal methyltransferase core complex. In human, six H3K4 methyltransferases (MLL1, MLL2, MLL3, MLL4, SETD1A, and SETD1B) commonly share a minimal C-terminal SET domain, forming a core complex in conjunction with their auxiliary components (RbBP5, WDR5, ASH2L, and DPY30) (Bochynska et al. 2018; Antunes and Ottersbach 2020). In contrast to the conserved C-terminal SET

domain, the N-terminal fragments of these six H3K4 methyltransferases are highly variable, implying that the N-terminal domain serves as a platform for the recruitment of interaction partners in a target-specific manner. For example, MLL1 and MLL2 interact with Menin and LEDGF (Hughes et al. 2004; Yokoyama and Cleary 2008; Murai et al. 2014), while MLL3 and MLL4 engage PTIP and NCOA6 through their N-terminal domains (Goo et al. 2003; Cho et al. 2007; Patel et al. 2007). Notably, MLL1 and MLL2 undergo proteolytic cleavage by Taspase 1 (Takeda et al. 2006) and form non-covalent dimers via FYRN and FYRC domains, which physically break and reconnect the N- and C-terminal domains (Yokoyama et al. 2002).

To unravel the contributions of the N-terminal domains of the six MLL isoforms to target-specific recognition and their subsequent influence on methyltransferase activity located at the C-terminal domain, it is imperative to determine cryo-EM structures of either the isolated N-terminal domains or the full-length MLL holo-complex. The former study showed that the full-length COMPASS complex exhibits reduced catalytic activity toward NCP and H2Bub.NCP compared to the C-terminal COMPASS core complex, suggesting that the N-terminal portion of COMPASS may play a negative role in enzymatic activity (Hsu et al. 2019). It is anticipated that a significant portion of the N-terminal domain will exhibit a lack of structural order, or remain largely unstructured. Although efforts have been made to elucidate the native architecture of the *S. cerevisiae* COMPASS holo-complex, only a low-resolution cryo-EM structure has been achieved, offering limited insights into the overall architecture of the yeast COMPASS holo-complex (Wang et al. 2018). Attaining near-atomic resolution structures of the yeast COMPASS and human MLL holo-complexes would provide a comprehensive view of how chromatin binding and methyltransferase activity synergize.

Furthermore, it has been shown that the fusion form of MLL1 lacks the C-terminal SET domain and fuses with one of the partner genes (Meyer et al. 2006). MLL1 fusion chimeras like MLL1-AF9 still maintain the expression of some of the MLL1 target genes in hematopoietic stem cells and drive leukemogenesis (Jing et al. 2016; Mishra et al. 2014). These observations suggest the existence of alternative mechanisms for the direct regulation of MLL1 target genes mediated by MLL1 fusion chimeras (Mishra et al. 2014; Chopra and Bohlander 2015). The cryo-EM structure of MLL1 fusion chimeras has potential to furnish new molecular-level insights into the regulation of gene expression at MLL1 target genes.

Acknowledgment I thank Dr. Yali Dou at the University of Southern California for comments and suggestions. This work is supported by the NCI grant (CA250329) to Y.D. and U.S.C.

References

- Antunes ETB, Ottersbach K (2020) The MLL/SET family and haematopoiesis. *Biochim Biophys Acta Gene Regul Mech* 1863(8):194579
- Avdic V, Zhang P, Lanouette S, Groulx A, Tremblay V, Brunzelle J, Couture JF (2011) Structural and biochemical insights into MLL1 core complex assembly. *Structure* 19(1):101–108

- Ayoub A, Park SH, Lee YT, Cho US, Dou Y (2022) Regulation of MLL1 methyltransferase activity in two distinct nucleosome binding modes. *Biochemistry* 61(1):1–9
- Bochynska A, Luscher-Firzlaff J, Luscher B (2018) Modes of interaction of KMT2 histone H3 lysine 4 methyltransferase/COMPASS complexes with chromatin. *Cells* 7(3)
- Briggs SD, Bryk M, Strahl BD, Cheung WL, Davie JK, Dent SY, Winston F, Allis CD (2001) Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev* 15(24):3286–3295
- Buratowski S, Kim T (2010) The role of cotranscriptional histone methylations. *Cold Spring Harb Symp Quant Biol* 75:95–102
- Cao F, Chen Y, Cierpicki T, Liu Y, Basrur V, Lei M, Dou Y (2010) An Ash2L/RbBP5 heterodimer stimulates the MLL1 methyltransferase activity through coordinated substrate interactions with the MLL1 SET domain. *PLoS One* 5(11):e14102
- Chessells JM, Harrison CJ, Kempinski H, Webb DK, Wheatley K, Hann IM, Stevens RF, Harrison G, Gibson BE, M. R. C. C. L. w. party (2002) Clinical features, cytogenetics and outcome in acute lymphoblastic and myeloid leukaemia of infancy: report from the MRC Childhood Leukaemia working party. *Leukemia* 16(5):776–784
- Cho YW, Hong T, Hong S, Guo H, Yu H, Kim D, Guszczynski T, Dressler GR, Copeland TD, Kalkum M, Ge K (2007) PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* 282(28):20395–20406
- Chopra M, Bohlander SK (2015) Disturbing the histone code in leukemia: translocations and mutations affecting histone methyl transferases. *Cancer Genet* 208(5):192–205
- Dehe PM, Dichtl B, Schaft D, Roguev A, Pamblanco M, Lebrun R et al (2006) Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *J Biol Chem* 281(46):35404–35412. <https://doi.org/10.1074/jbc.M603099200>.
- Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* 6(8):227
- Djabali M, Selleri L, Parry P, Bower M, Young BD, Evans GA (1992) A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. *Nat Genet* 2(2):113–118
- Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Wysocka J, Allis CD, Chait BT, Hess JL, Roeder RG (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* 121(6):873–885. <https://doi.org/10.1016/j.cell.2005.04.031>. PMID: 15960975
- Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD, Roeder RG (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* 13(8):713–719
- Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A (2002) Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* 277(32):28368–28371
- Goo YH, Sohn YC, Kim DH, Kim SW, Kang MJ, Jung DJ, Kwak E, Barlev NA, Berger SL, Chow VT, Roeder RG, Azorsa DO, Meltzer PS, Suh PG, Song EJ, Lee KJ, Lee YC, Lee JW (2003) Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins. *Mol Cell Biol* 23(1):140–149
- Gu Y, Nakamura T, Alder H, Prasad R, Canaani O, Cimino G, Croce CM, Canaani E (1992) The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to *Drosophila* trithorax, to the AF-4 gene. *Cell* 71(4):701–708
- Guenther MG, Jenner RG, Chevalier B, Nakamura T, Croce CM, Canaani E, Young RA (2005) Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci USA* 102(24):8603–8608
- Hannibal MC, Buckingham KJ, Ng SB, Ming JE, Beck AE, McMillin MJ, Gildersleeve HI, Bigham AW, Tabor HK, Mefford HC, Cook J, Yoshiura K, Matsumoto T, Matsumoto N, Miyake N, Tonoki H, Naritomi K, Kaname T, Nagai T, Ohashi H, Kurosawa K, Hou JW, Ohta T, Liang D, Sudo A, Morris CA, Banka S, Black GC, Clayton-Smith J, Nickerson DA, Zackai EH, Shaikh TH, Donnai D, Niikawa N, Shendure J, Bamshad MJ (2011) Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome. *Am J Med Genet A* 155A(7):1511–1516

- Hilden JM, Dinndorf PA, Meerbaum SO, Sather H, Villaluna D, Heerema NA, McGlennen R, Smith FO, Woods WG, Salzer WL, Johnstone HS, Dreyer Z, Reaman GH, Children's Oncology G (2006) Analysis of prognostic factors of acute lymphoblastic leukemia in infants: report on CCG 1953 from the Children's Oncology Group. *Blood* 108(2):441–451
- Hsu PL, Li H, Lau HT, Leonen C, Dhall A, Ong SE, Chatterjee C, Zheng N (2018) Crystal structure of the COMPASS H3K4 methyltransferase catalytic module. *Cell* 174(5):1106–1116 e1109
- Hsu PL, Shi H, Leonen C, Kang J, Chatterjee C, Zheng N (2019) Structural basis of H2B ubiquitination-dependent H3K4 methylation by COMPASS. *Mol Cell* 76(5):712–723 e714
- Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, Lee JC, Hayes DN, Shanmugam KS, Bhattacharjee A, Biondi CA, Kay GF, Hayward NK, Hess JL, Meyerson M (2004) Menin associates with a trithorax family histone methyltransferase complex and with the *hoxc8* locus. *Mol Cell* 13(4):587–597
- Jing X, Li L, Xiong J, den Dekker A, Ye A, Karatas H, Liu L, Wang H, Qin ZS, Wang S, Dou Y (2016) MLL1 and MLL1 fusion proteins have distinct functions in regulating leukemic transcription program. *Cell Discov* 2:16008. <https://doi.org/10.1038/celldisc.2016.8>. eCollection 2016
- Jones WD, Dafou D, McEntagart M, Woollard WJ, Elmslie FV, Holder-Espinasse M, Irving M, Sagar AK, Smithson S, Trembath RC, Deshpande C, Simpson MA (2012) De novo mutations in MLL cause Wiedemann-Steiner syndrome. *Am J Hum Genet* 91(2):358–364
- Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MDM, Miller CA, Welch JS, Walter MJ, Wendl MC, Ley TJ, Wilson RK, Raphael BJ, Ding L (2013) Mutational landscape and significance across 12 major cancer types. *Nature* 502(7471):333–339
- Kim J, Kim J-A, McGinty RK, Nguyen UTT, Muir TW, David Allis C, Roeder RG (2013) The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. *Mol Cell* 49(6):1121–1133. <https://doi.org/10.1016/j.molcel.2013.01.034>. Epub 2013 Feb 28. PMID: 23453808
- Kleefstra T, Kramer JM, Neveling K, Willemsen MH, Koemans TS, Vissers LE, Wissink-Lindhout W, Fenckova M, van den Akker WM, Kasri NN, Nillesen WM, Prescott T, Clark RD, Devriendt K, van Reeuwijk J, de Brouwer AP, Gilissen C, Zhou H, Brunner HG, Veltman JA, Schenck A, van Bokhoven H (2012) Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am J Hum Genet* 91(1):73–82
- Lee YT, Ayoub A, Park SH, Sha L, Xu J, Mao F, Zheng W, Zhang Y, Cho US, Dou Y (2021) Mechanism for DPY30 and ASH2L intrinsically disordered regions to modulate the MLL/SET1 activity on chromatin. *Nat Commun* 12(1):2953
- Li Y, Han J, Zhang Y, Cao F, Liu Z, Li S, Wu J, Hu C, Wang Y, Shuai J, Chen J, Cao L, Li D, Shi P, Tian C, Zhang J, Dou Y, Li G, Chen Y, Lei M (2016) Structural basis for activity regulation of MLL family methyltransferases. *Nature* 530(7591):447–452
- Meyer C, Schneider B, Jakob S, Strehl S, Attarbaschi A, Schnittger S, Schoch C, Jansen MW, van Dongen JJ, den Boer ML, Pieters R, Ennas MG, Angelucci E, Koehl U, Greil J, Griesinger F, Zur Stadt U, Eckert C, Szczepanski T, Niggli FK, Schafer BW, Kempinski H, Brady HJ, Zuna J, Trka J, Nigro LL, Biondi A, Delabesse E, Macintyre E, Stanulla M, Schrappe M, Haas OA, Burmeister T, Dingermann T, Klingebiel T, Marschalek R (2006) The MLL recombinome of acute leukemias. *Leukemia* 20(5):777–784
- Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A (2001) COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci USA* 98(23):12902–12907
- Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* 10(5):1107–1117
- Mishra BP, Zaffuto KM, Artinger EL, Org T, Mikkola HK, Cheng C, Djabali M, Ernst P (2014) The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. *Cell Rep* 7(4):1239–1247
- Murai MJ, Pollock J, He S, Miao H, Purohit T, Yokom A, Hess JL, Muntean AG, Grembecka J, Cierpicki T (2014) The same site on the integrase-binding domain of lens epithelium-derived growth factor is a therapeutic target for MLL leukemia and HIV. *Blood* 124(25):3730–3737

- Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM, Canaani E (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell* 10(5):1119–1128
- Park SH, Ayoub A, Lee YT, Xu J, Kim H, Zheng W, Zhang B, Sha L, An S, Zhang Y, Cianfrocco MA, Su M, Dou Y, Cho US (2019) Cryo-EM structure of the human MLL1 core complex bound to the nucleosome. *Nat Commun* 10(1):5540
- Patel A, Dharmarajan V, Vought VE, Cosgrove MS (2009) On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *J Biol Chem* 284(36):24242–24256
- Patel SR, Kim D, Levitan I, Dressler GR (2007) The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex. *Dev Cell* 13(4):580–592
- Pieters R, Schrappe M, De Lorenzo P, Hann I, De Rossi G, Felice M, Hovi L, LeBlanc T, Szczepanski T, Ferster A, Janka G, Rubnitz J, Silverman L, Stary J, Campbell M, Li CK, Mann G, Suppiah R, Biondi A, Vora A, Valsecchi MG (2007) A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. *Lancet* 370(9583):240–250
- Qu Q, Takahashi YH, Yang Y, Hu H, Zhang Y, Brunzelle JS, Couture JF, Shilatifard A, Skiniotis G (2018) Structure and conformational dynamics of a COMPASS histone H3K4 methyltransferase complex. *Cell* 174(5):1117–1126 e1112
- Rahman S, Hoffmann NA, Worden EJ, Smith ML, Namitz KEW, Knutson BA, Cosgrove MS, Wolberger C (2022) Multistate structures of the MLL1-WRAD complex bound to H2B-ubiquitinated nucleosome. *Proc Natl Acad Sci USA* 119(38):e2205691119
- Rao RC, Dou Y (2015) Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* 15(6):334–346
- Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilm M, Aasland R, Stewart AF (2001) The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J* 20(24):7137–7148
- Ruthenburg AJ, Allis CD, Wysocka J (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* 25(1):15–30
- Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* 419(6905):407–411
- Schuettengruber B, Martinez AM, Iovino N, Cavalli G (2011) Trithorax group proteins: switching genes on and keeping them active. *Nat Rev Mol Cell Biol* 12(12):799–814
- Sedkov Y, Benes JJ, Berger JR, Riker KM, Tillib S, Jones RS, Mazo A (1999) Molecular genetic analysis of the *Drosophila* trithorax-related gene which encodes a novel SET domain protein. *Mech Dev* 82(1–2):171–179
- Sedkov Y, Cho E, Petruk S, Cherbas L, Smith ST, Jones RS, Cherbas P, Canaani E, Jaynes JB, Mazo A (2003) Methylation at lysine 4 of histone H3 in ecdysone-dependent development of *Drosophila*. *Nature* 426(6962):78–83
- Shen E, Shulha H, Weng Z, Akbarian S (2014) Regulation of histone H3K4 methylation in brain development and disease. *Philos Trans R Soc Lond Ser B Biol Sci* 369(1652):20130514
- Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* 75:243–269
- Sims RJ 3rd, Reinberg D (2006) Histone H3 Lys 4 methylation: caught in a bind? *Genes Dev* 20(20):2779–2786
- Southall SM, Wong PS, Odho Z, Roe SM, Wilson JR (2009) Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. *Mol Cell* 33(2):181–191
- Steward MM, Lee JS, O'Donovan A, Wyatt M, Bernstein BE, Shilatifard A (2006) Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat Struct Mol Biol* 13(9):852–854

- Sun ZW, Allis CD (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 418(6893):104–108
- Takahashi YH, Westfield GH, Oleskie AN, Trievel RC, Shilatifard A, Skiniotis G (2011) Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc Natl Acad Sci USA* 108(51):20526–20531
- Takeda S, Chen DY, Westergard TD, Fisher JK, Rubens JA, Sasagawa S, Kan JT, Korsmeyer SJ, Cheng EH, Hsieh JJ (2006) Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes Dev* 20(17):2397–2409
- Tomizawa D, Koh K, Sato T, Kinukawa N, Morimoto A, Isoyama K, Kosaka Y, Oda T, Oda M, Hayashi Y, Eguchi M, Horibe K, Nakahata T, Mizutani S, Ishii E (2007) Outcome of risk-based therapy for infant acute lymphoblastic leukemia with or without an MLL gene rearrangement, with emphasis on late effects: a final report of two consecutive studies, MLL96 and MLL98, of the Japan Infant Leukemia Study Group. *Leukemia* 21(11):2258–2263
- van der Linden MH, Valsecchi MG, De Lorenzo P, Moricke A, Janka G, Leblanc TM, Felice M, Biondi A, Campbell M, Hann I, Rubnitz JE, Stary J, Szczepanski T, Vora A, Ferster A, Hovi L, Silverman LB, Pieters R (2009) Outcome of congenital acute lymphoblastic leukemia treated on the Interfant-99 protocol. *Blood* 114(18):3764–3768
- Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131(1):58–69
- Vermeulen M, Timmers HT (2010) Grasping trimethylation of histone H3 at lysine 4. *Epigenomics* 2(3):395–406
- Wang Y, Ding Z, Liu X, Bao Y, Huang M, Wong CCL, Hong X, Cong Y (2018) Architecture and subunit arrangement of the complete *Saccharomyces cerevisiae* COMPASS complex. *Sci Rep* 8(1):17405
- Worden EJ, Zhang X, Wolberger C (2020) Structural basis for COMPASS recognition of an H2B-ubiquitinated nucleosome. *elife* 9
- Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121(6):859–872
- Xiao B, Gamblin SJ, Wilson JR (2006) 6 structure of SET domain protein lysine methyltransferases. *Enzyme* 24:155–178
- Xiao B, Wilson JR, Gamblin SJ (2003) SET domains and histone methylation. *Curr Opin Struct Biol* 13(6):699–705
- Xue H, Yao T, Cao M, Zhu G, Li Y, Yuan G, Chen Y, Lei M, Huang J (2019) Structural basis of nucleosome recognition and modification by MLL methyltransferases. *Nature* 573(7774):445–449
- Yang Z, Shah K, Khodadadi-Jamayran A, Jiang H (2016) Dpy30 is critical for maintaining the identity and function of adult hematopoietic stem cells. *J Exp Med* 213(11):2349–2364
- Yokoyama A, Cleary ML (2008) Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* 14(1):36–46
- Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML, Ohki M (2002) Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. *Blood* 100(10):3710–3718
- Zhang X, Yang Z, Khan SI, Horton JR, Tamaru H, Selker EU, Cheng X (2003) Structural basis for the product specificity of histone lysine methyltransferases. *Mol Cell* 12(1):177–185

Chapter 7

Oligomerization of Monoamine Transporters



Arnav Paul and Diwakar Shukla

Abstract Transporters of the monoamine transporter (MAT) family regulate the uptake of important neurotransmitters like dopamine, serotonin, and norepinephrine. The MAT family functions using the electrochemical gradient of ions across the membrane and comprises three transporters, dopamine transporter (DAT), serotonin transporter (SERT), and norepinephrine transporter (NET). MAT transporters have been observed to exist in monomeric states to higher-order oligomeric states. Structural features, allosteric modulation, and lipid environment regulate the oligomerization of MAT transporters. NET and SERT oligomerization are regulated by levels of PIP2 present in the membrane. The kink present in TM12 in the MAT family is crucial for dimer interface formation. Allosteric modulation in the dimer interface hinders dimer formation. Oligomerization also influences the transporters' function, trafficking, and regulation. This chapter will focus on recent studies on monoamine transporters and discuss the factors affecting their oligomerization and its impact on their function.

Keywords Monoamine transporters · DAT · SERT · NET · PIP2 · Oligomers

A. Paul

Department of Chemistry, University of Illinois Urbana-Champaign, Urbana, IL, USA
e-mail: paul30@illinois.edu

D. Shukla (✉)

Department of Chemical and Biomolecular Engineering, Department of Bioengineering, Center for Biophysics and Quantitative Biology, Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
e-mail: diwakar@illinois.edu

Introduction

Neuronal excitability is the process of generating an alteration in membrane potential, which is used by the neuron to propagate information across the synaptic cleft. Changes in the membrane potential result in the generation of an ionic current across the membrane. The membrane potential can be modulated by neurotransmitter release at synapses. Presynaptic neurons release neurotransmitters that bind to receptors, which need to be removed efficiently to generate new signals or modulate neuron activity. This membrane potential is generated through reuptake of neurotransmitter molecules by selectively permeable ion channels and dedicated active membrane transporters. Improper reuptake of neurotransmitters from the synapses has been linked to multiple neurodegenerative diseases and neurological disorders (e.g., schizophrenia, depression, Parkinson's disease, and attention deficit hyperactivity disorder) (Kristensen et al. 2011). These transporters are also the target of illegal recreational drugs and narcotics that interfere with the reuptake of substrates (Sitte and Freissmuth 2015). The largest family of neurotransmitter transporters is the solute carrier 6 (SLC6) family of transporters. SLC6 transporters use the electrochemical gradient caused by differential sodium ion concentrations across the membrane bilayer to transport substrates. The SLC6 family of transporters mainly transport monoamines and include norepinephrine transporter (NET), dopamine transporter (DAT), serotonin transporter (SERT), glycine transporter (GLYT), proline transporter (PROT), γ -amino butyric acid transporter (GAT), taurine transporter (TauT), and creatine transporter (CRT).

A subfamily of SLC6 is the monoamine transporter (MAT) family, comprising the dopamine transporter, the serotonin transporter, and the norepinephrine transporter. The plasmalemmal norepinephrine transporter (NET) is a neurotransmitter transporter that belongs to the sodium solute carrier (SLC6) family (Pacholczyk et al. 1991). NET controls the uptake of norepinephrine from the synaptic cleft. NET's function is linked with various important physiological processes, and its regulation is controlled by various proteins, kinases, and lipids (Apparsundaram et al. 1998a, b; Liu et al. 2018). Many diseases, including neurodegenerative diseases like Alzheimer's disease, have been linked to altered norepinephrine signaling. NET is also a target of prescription drugs and illegal psychostimulant drugs (Kristensen et al. 2011; Luethi and Liechti 2020). The phospholipid phosphatidylinositol 4,5-bisphosphate (PIP2) plays a key role in regulating substrate efflux and dimerization of the NET transporter (Jayaraman et al. 2021; Buchmayer et al. 2013; Anderluh et al. 2017; Belovich et al. 2021; Hamilton et al. 2014).

Serotonin transporter (SERT) is present in the synaptic cleft of the neuron that regulates the uptake of serotonin in the synapse for neurotransmission. Like NET, SERT is also known to exist in oligomeric forms, usually as tetramers, in the forms of pairs of dimers, but is present in monomeric and dimeric states as well (Schmid et al. 2001; Just et al. 2004; Anderluh et al. 2014; Kilic and Rudnick 2000; Jess et al. 1996). SERT is a member of the neurotransmitter sodium symporter (NSS) family or the solute carrier (SLC6) family. SERT is a target of antidepressant drugs that

block the reuptake of serotonin as well as psychostimulant drugs like amphetamines that reverse the transport of SERT, releasing excess serotonin in the synaptic cleft (Kristensen et al. 2011; Schloss and Williams 1998).

Another member of MAT family is the dopamine transporter DAT. DAT is involved in dopamine uptake from the synaptic cleft and is an important target for psychostimulants (Rothman and Baumann 2003; Giros and Caron 1993). Altered DAT function can lead to diseases like depression, schizophrenia, bipolar disease, and Parkinson's disease (Torres et al. 2003b). DAT is found in multiple oligomeric forms like monomers, dimers, and tetramers (Milner et al. 1994; Berger et al. 1994; Torres et al. 2003a; Chen and Reith 2008; Egaña et al. 2009; Sorkina et al. 2003).

These transporters follow the alternate access model proposed by Jardetzky, where the transporter alternates between being accessible to the substrate binding site from the intracellular side or extracellular side, via a closed-occluded state (Jardetzky 1966). The general principles of alternate access in the context of ion-dependent transport have been exquisitely reviewed in this article (Beckstein and Naughton 2022).

Transporters of the SLC6 family have 12 transmembrane helices forming two domains, scaffold and bundle domains (Forrest and Rudnick 2009). The scaffold domain is larger and stays stationary, anchoring the protein to the membrane, while the smaller bundle domain moves to provide access to the substrate-binding site alternatively to intracellular and extracellular sides of the membrane. This type of domain-coupled transporter movement is referred to as a rocker-bundle mechanism (Drew and Boudker 2016; Henderson et al. 2019; Forrest et al. 2011). The crystal structure of leucine transporter LeuT from *Aquifex aeolicus* first shed light on the structure of SLC6 family of transporters (Forrest and Rudnick 2009). The transporters go through five major conformations: outward open, outward occluded, occluded, inward occluded, and inward facing (Krishnamurthy and Gouaux 2012; Singh et al. 2008; Yamashita et al. 2005; Gotfryd et al. 2020). The resolved structures of MAT subfamily transporters DAT, SERT, and NET share a conserved fold across different species, and all transporters follow the conformational cycle described above. However, small differences in sequences between the transporters lead to large differences in their function and oligomerization behavior. A recent review by Jayaraman and coworkers reviewed the advances in SLC6 transporter oligomerization (Jayaraman et al. 2021). This chapter will explore the factors that regulate the oligomerization of MAT transporters.

Oligomeric States of Monoamine Transporters

Monoamine transporters can form oligomers, which are complexes made up of multiple protein subunits. Oligomerization is a common phenomenon among membrane proteins, including neurotransmitter transporters (Anderluh et al. 2017; Luethi et al. 2022; Das et al. 2019). The exact nature of the oligomers and their functional significance can vary among different types of transporters.

Oligomerization of NET

A study by Sitte and coworkers examined oligomerization of NET transporter and transporter kinetics (Luethi et al. 2022). To determine oligomer distribution of hNET, the authors used single-molecule brightness analysis method called “thinning out clusters while conserving stoichiometry of labeling” (TOCCSL) (Moertelmaier et al. 2005). In this method, an area of the cell membrane is photobleached by a high-intensity pulse. Due to Brownian motion, unbleached molecules enter the photobleached area. Unbleached molecules can be tracked by monitoring fluorescence spots. Oligomer distribution states can be determined by comparing the brightness of these spots with a single NET molecule. The probability distribution function of the oligomer size was obtained by looking at the brightness distribution. The distribution showed that NET exists as monomers and dimers at the plasma membrane. Around 60% of the molecules are observed to exist as a monomer and the rest 40% as dimers. The authors then studied if the dimeric NETs exchange subunits with other dimers or not. If there is exchange of subunits between photobleached dimer and non-photobleached dimer, the resulting pairs would have a photobleached subunit and one non-photobleached subunit (Fig. 7.1). This would show up as increasing concentrations of monomers. The authors observed that in a time point experiment of 0, 10, and 20 min, the concentration of dimer systems decrease, and monomers increase as time progresses. This indicates that there is exchange of subunits among the dimers of NET molecules with increase in time, suggesting that dimers are not very stable over the timescale of minutes.

Oligomerization of SERT

Schütz and coworkers studied the oligomerization of SERT using the TOCCSL method to determine the distribution of oligomers (Anderluh et al. 2017; Moertelmaier et al. 2005). They observed that in the plasma membrane, the oligomer distribution remained constant as SERT expression increases and did not depend on SERT surface density. However, in the endoplasmic reticulum (ER), the average oligomer size increases with an increase in expression levels of SERT in the ER. To further study the stability of the oligomers formed, a time point TOCCSL run was done to examine subunit exchange between oligomers. If the higher-order oligomers dissociate or exchange units with photobleached oligomers, the apparent number of smaller oligomers would increase. The authors observed that in the ER, the average number of smaller oligomers increases with time, suggesting that rapid subunit exchange occurs at the ER and the oligomers are not very stable. In the plasma membrane, it was observed that the oligomer distribution remained constant over time, suggesting that larger oligomers remain stable in the plasma membrane.

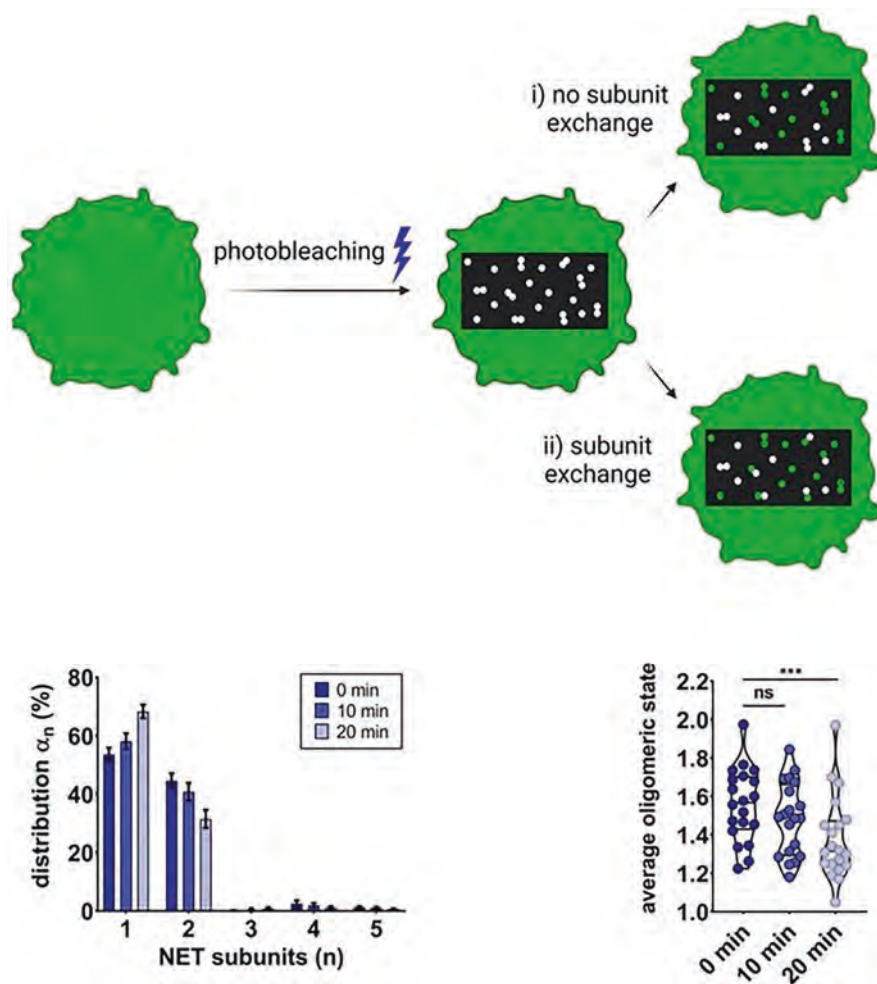


Fig. 7.1 (Top) Sketch showing two potential scenarios. (i) Scenario 1: dimers are stable; repeated runs result in a decreased number of active fluorophores but no altered distribution. (ii) Scenario 2: dimers exchange subunits; the exchange of bleached (white) and unbleached (green) dimers results in an increasing fraction of apparent monomers. (Bottom left) Repeated TOCCSL runs on the same cell indicate that a small portion of NET dimers exchange subunits over time. A total of 20 cells were repeatedly measured three times over the course of 20 min. Data show means $\pm$ SD. (Bottom right) Average oligomeric state of NET molecules for each of the repeatedly measured cells. Reproduced from Dino Luethi, Julian Maier, et al. *Communications Biology*, 5(1):1259, 2022 (Luethi et al. 2022). This article was originally published and distributed under the terms of the Creative Commons Attribution License, permitting unrestricted reproduction and adaptation provided proper crediting to author and source

Oligomerization of DAT

Sitte and coworkers investigated the dimerization of DAT similar to their study on NET and SERT (Anderluh et al. 2017; Luethi et al. 2022; Das et al. 2019). They used the TOCCSL approach to determine the distribution of oligomers present in the membrane (Moertelmaier et al. 2005). It was observed that 55% of the DATs were present in monomeric form and 35% in dimeric form. A small fraction of higher-order oligomers were also identified. Further, the stability of dimers was studied by looking at the subunit exchange rates between the dimers. TOCCSL was run at different time points from 5 to 50 min, and if subunit exchange happens between photobleached dimers and fluorescence active dimers, the number of apparent monomers would increase, making dimers effectively decrease. It was observed that the DAT dimer stays stable for at least several minutes, before they dissociate for subunit exchange or dissociate into monomers.

Structural Features of Dimer Interface of MATs

The dimerization interface contributes to the overall stability of the protein complex. The interaction between subunits at the dimer interface can affect the conformation of the transporters and, consequently, their functional properties.

Dimer Interface of SERT

A computational study by Schjøtt and coworkers investigated the dimerization property of serotonin transporter SERT (Periole et al. 2018). Molecular dynamics (MD) was used to determine the intrinsic dynamics of the transporter and its self-assembly. The strength of the dimeric interfaces of SerT was investigated, and the self-assembly of the proteins in a membrane environment was observed. TM1/2 and TM11/12 helices were previously observed to be involved in symmetric dimer interface formation (Just et al. 2004). To verify the experimental results and to gain insights into the dimerization process, multiple long coarse-grained simulations were performed with 16 to 64 SerT proteins in a membrane, and their self-assembly was observed (Fig. 7.2). Multiple SerT proteins formed dimers and higher-order oligomers. Four main dimer interfaces were identified in a cluster analysis, out of which 40% of the dimer interfaces involved transmembrane helix 12 (TM12). The largest cluster was formed by the interaction of TM7-TM12 and constituted around 22% of all the dimers formed. The second cluster observed was a symmetric cluster formed by TM12-TM12 interactions and accounted for 18% of the clusters. The third cluster was an asymmetrical cluster involving TM4/9 and TM2/11 constituting for 18% of the clusters, and the last cluster was a symmetric cluster between TM3/4

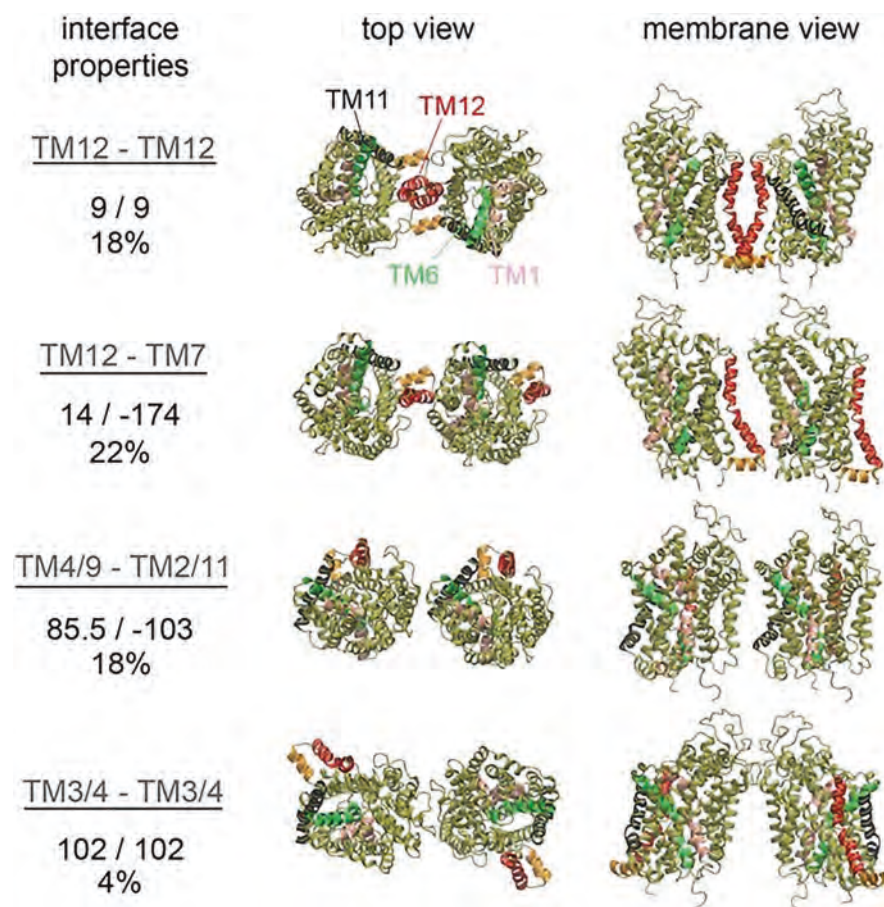


Fig. 7.2 Representative structures of the four most populated interfaces formed in the self-assembly CG simulations of hSERT in a POPC membrane. From top to bottom: TM12-TM12, TM12-TM7, TM4/9-TM2/11, and TM3/4-TM3/4. For each interface, the values of the $\phi 1$ and $\phi 3$ angles used to define their relative orientation in the VBA and their contributions to the first ten clusters are indicated. Atomistic structures of hSERT were fitted onto the CG configurations. Reproduced from Xavier Periole, Talia Zeppelin, and Birgit Schiøtt. Scientific reports, 8(1):5080, 2018 (Periole et al. 2018). This article was originally published and distributed under the terms of the Creative Commons Attribution License, permitting unrestricted reproduction and adaptation provided proper crediting to author and source

and TM3/4 and involves the extracellular loop 2 (ECL2). TM12 has a major contribution in oligomerization of SerT protein, and to understand its contribution in stabilizing the dimer interface, potential of mean force (PMF) calculations using replica exchange molecular dynamics (REMD) were performed (Sugita et al. 2000). From the PMF calculations, it was found out that the symmetric TM12 dimer is the most stable dimer orientation, which is similar to the bacterial homologue of SERT, LeuT dimer interface (Just et al. 2004; Yamashita et al. 2005; Neubauer et al. 2006).

Dimer Interface of DAT

The structure of SLC6 family members has the same 12-transmembrane helix topology; however, there are small differences that make each of the transporter behave in a different way. Bahar and coworkers published a study where they compared the structural features of LeuT dimer with a computational model of hDAT dimer (Cheng et al. 2017). The LeuT structure has been characterized in the outward-facing (OF) and inward-facing (IF) states, as well as the OF-OF dimer (Krishnamurthy and Gouaux 2012; Singh et al. 2008; Yamashita et al. 2005; Forrest 2013). The authors used a docking-based computational method called ClusPro to model the dimeric structure of LeuT in the IF-IF state and OF-OF state (Comeau et al. 2004). After validating the results from the computational method with the resolved crystal structure of the OF-OF LeuT dimer, the computational method was applied on dopamine transporters to obtain a dimer model of hDAT. The predicted model of the hDAT dimer using the computational method was observed to be significantly different to the LeuT dimer. LeuT forms a dimer interface using TM12-TM12 with support from TM9 and EL2 loop. However, in the hDAT dimer, they observed that the association of TM2, TM6, and TM11 formed the dimer interface (Fig. 7.3). The residues that are involved in the dimer interface are V100-P112 from TM2, F302-W311 from TM6a, and S517-F543 from TM11. Previous experimental results indicated that a leucine zipper in TM2 can stabilize the dimer interface of hDAT, which was also observed in the experimental model (Torres et al. 2003a). Experimental studies suggest that a potential Cys306 cross-linking interaction between the hDATs could be important for dimer formation (Hastrup et al. 2001, 2003). The distance between Cys306 of both the subunits is within cross-linking range. This validates the computational model prediction of TM2, TM6, and TM11 forming the dimer interface instead of TM12 for hDAT. Another reason for TM12 not forming the dimer interface could be attributed to the sharp kink present in TM12 of hDAT, which prevents the helices from interacting. To observe the effect of the TM12 kink, the authors mutated the kink-forming residues to their counterpart in the LeuT. The surface expression levels of mutant hDATs reduced significantly. To study the stability of the predicted models, MD simulations were performed to measure the RMSD and gating dynamics. They observed that the predicted models remain stable over 100 ns simulations, and the gating dynamics resemble monomer dynamics. Further the authors also validated that Cys306 cross-linking is necessary for dimerization. They showed that promoting disulfide bridge formation with the help of oxidizing agent leads to an increase in dimerization of hDAT. Nearby charged residues were also observed to form interfacial salt bridges.

Further study on hDAT dimer formation by Schiøtt and coworkers shed light on the effect of palmitoylation on the dimer interface (Cheng et al. 2017). The authors used coarse-grained molecular dynamics simulations to observe the aggregation of hDAT monomers in a plasma membrane. Two sets of simulations were performed, one with 16 hDAT monomers and one with 16 S-palmitoylated hDAT on TM12 (hDAT-palm) monomers. Palmitoylation of DATs has been shown to regulate

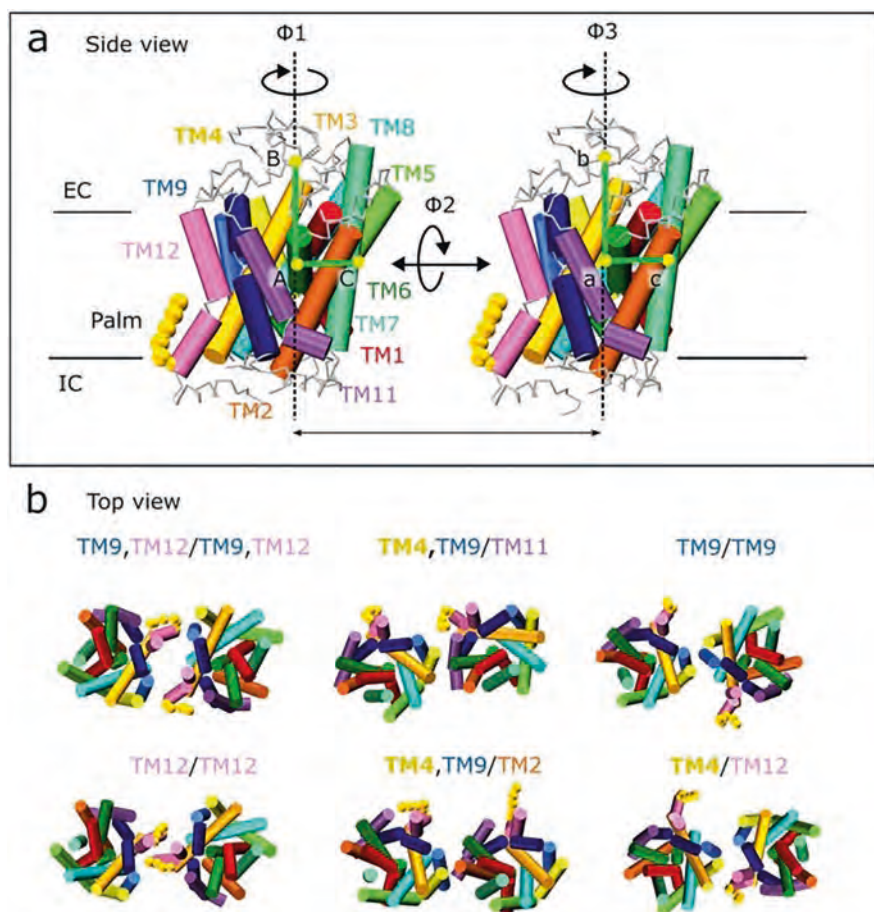


Fig. 7.3 Dimer cluster comparison between the two systems, hDAT and hDAT-palm. Depicted in (a) is a transmembrane (TM) overview of hDAT shown from the side with respect to the membrane normal. The helices are individually color-coded using the same coloring scheme as applied throughout the article. Highlighted in bold spheres is the palmitoyl group (palm) covalently bound to TM12. The two angles, ϕ_1 and ϕ_3 , are used for describing the individual protomer's relative orientation. Depicted in (b) are the common top clusters observed in both the hDAT and hDAT-palm simulations, with the exception of the TM12/TM12 dimer. Reproduced from Talia Zeppelin, Kasper B Pedersen, Nils A Berglund, Xavier Periole, and Birgit Schiøtt. Scientific Reports, 11(1):4164, 2021 (Zeppelin et al. 2021). This article was originally published and distributed under the terms of the Creative Commons Attribution License, permitting unrestricted reproduction and adaptation provided proper crediting to author and source

surface expression and turnover (Rastedt et al. 2017; Foster and Vaughan 2011). From the simulations, the authors observed cluster formation of hDAT to form dimers by self-assembly.

Multiple dimer conformations were observed, out of which TM9, TM12/TM9, and TM12 dimers were observed to be the most stable. LeuT has also been observed

to stabilize in a similar dimer conformation (Yamashita et al. 2005). The authors also observed dimeric interface formation between TM4/TM4 and TM6/TM6; however, they were not very stable. The dimer interface involved was investigated further due to it being structurally different to TM12 of hSERT. TM12 of hDAT possessed a kink that prevents it from forming a direct TM12/TM12 dimer. The authors added the palmitoyl group on the TM12 of hDAT and observed that even though the energy of dimer formation changed for hDAT-palm when compared to hDAT, the distribution of the dimer interface formed remained constant. TM9 and TM11 were observed to form the dimer interface in most of the cases. In this study, it was observed that hDAT dimer formation involves interaction of TM9, TM11, and/or TM12. Palmitoylation of hDAT decreases the free energy of association, thus regulating the dimerization of hDATs.

Figure 7.4 shows the kink present in transporters of the MAT family compared to LeuT. Multiple sequence alignment was done on residues forming TM12 in the four transporters to identify the residues responsible for the kink in monoamine transporters. Serine residues present in the center of TM12 are responsible for the kink, which is highlighted by the multisequence alignment.

Effect of PIP2 on Oligomerization of MAT Transporters

Phosphatidylinositol 4,5-bisphosphate (PIP2) is a phospholipid present in the plasma membrane. It is present in 1% of the plasma membrane and is involved in various cellular functions. PIP2 plays an important role in signaling, membrane dynamics, regulating ion channels, and modulating oligomerization of membrane proteins (Picas et al. 2014; McLean et al. 2019; Liepiņa et al. 2003; Dutta Chowdhury et al. 2023). In this section, we will focus on the role of PIP2 in modulating the oligomerization of monoamine transporters.

Effect of PIP2 on NET

Sitte and coworkers examined the effect of lipid environment on NET oligomerization (Luethi et al. 2022). To study the effect of cholesterol and PIP2 in the plasma membrane, cholesterol/PIP2 concentrations were varied, and the effect on NET dimerization was followed. It was observed that oxidizing and depleting the cholesterol molecules did not affect the oligomerization of wild-type NET. To check for the effect of PIP2 on oligomerization of NET, a PLC activator was used to hydrolyze PIP2 (Bae et al. 2003). In the presence of PLC, it was observed that as PIP2 concentration decreases, the average density of dimers decreases for NET. This indicates that cholesterol does not play a role in oligomerization, but PIP2 is necessary for oligomerization of NET.

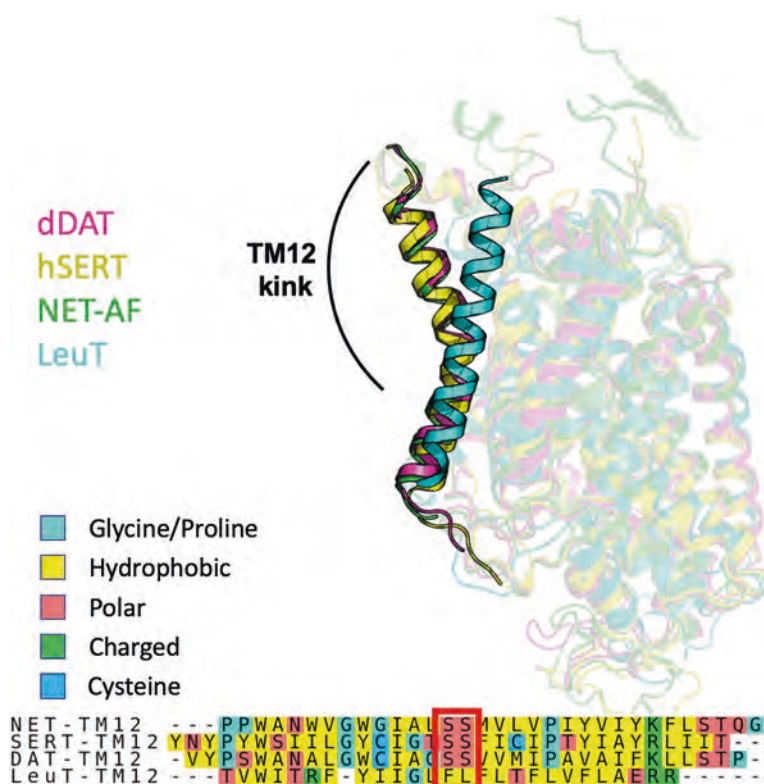


Fig. 7.4 Crystal structure of LeuT (PDB: 3TT1), dDAT (PDB: 4XP4), hSERT (PDB: 5I6X), and alphaFold prediction of NET overlaid together to highlight the kink present in TM12 of the MAT subfamily of transporters (Krishnamurthy and Gouaux 2012; Wang et al. 2015; Coleman et al. 2016; Jumper et al. 2021). The PDB structures were aligned using PyMOL (DeLano et al. 2002). Multiple sequence alignment of TM12 residues of SERT, DAT, NET, and LeuT was done using the Kalign web server to show the residues responsible for kink formation (Lassmann and Sonnhammer 2006). The red box in multiple sequence alignment points out the residues in TM12 that differ from LeuT, and the MAT transporters and are the reason for the kink in TM12

Sitte and coworkers also studied the dimeric interface of the proteins and mutated multiple positively charged residues that were probable PIP₂-binding sites to observe the effects on dimerization of NET (Luethi et al. 2022). Three residues were modified to alanines to make a triple mutant (R121A/K334A/R440A). These mutations were chosen because they are present at the dimer interface. It was observed that PIP₂-binding sites along the dimer interface do not play a key role in dimerization as the triple mutant also showed similar propensity to form dimers as the wild type. It was also found that PIP₂ also modulates substrate efflux, and in PIP₂-depleted systems, substrate uptake was not affected, but substrate efflux was reduced substantially. This study shows that NET forms stable dimers, and dimer formation is modulated by the level of PIP₂ present in the plasma membrane.

Effect of PIP2 on SERT

Schütz and coworkers studied the role of PIP2 in stabilizing the oligomers of SERT (Anderluh et al. 2017). The plasma membrane has a much higher concentration of PIP2 compared to ER, and it was suspected that PIP2 impacts SERT oligomerization (Buchmayer et al. 2013). A phospholipase C activator was used to modulate the level of PIP2 at the plasma membrane (Bae et al. 2003). In systems with lower density of PIP2, a higher percentage of smaller oligomers was observed. This reinforced the idea that PIP2 helps in modulating the oligomerization of SERT. Residues in the PIP2-binding site were mutated to alanines (K352A–K460A). The double mutant SERT was observed to form only monomers and dimers, whereas WT SERT also formed higher-order oligomers (Fig. 7.5). This suggests that PIP2 interaction at the cytosolic side of SERT stabilizes higher-order oligomers. PIP2 levels decide the kinetics of dissociation of oligomers into monomers, and a higher concentration of PIP2 slows the dissociation rate as in plasma membrane, and low concentration of PIP2 leads to faster dissociation of oligomers like in the ER. This study found the distribution of higher-order oligomers of SERT and observed that lack of PIP2 hinders the formation of higher-order SERT oligomers to dimers.

Schiøtt and coworkers investigated the effect of 1-palmitoyl-2-oleoyl-PIP2 (POPIP2) perturbation on the interface of symmetric TM12-TM12 dimer using computational tools (Zeppelin et al. 2021). From PMF calculations, it was observed that in presence of PPIP2, the proteins are stabilized at a TM12-TM12 distance of 0.3 nm, this stabilization was not observed in pure POPC membranes. However, it was observed that PPIP2 density at a distance of 0.58 nm increased the energy barrier for dimer formation, suggesting that PPIP2 prevents the association of monomers as well as separation of dimers. After identifying the role of PIP2 in the dimer interface, the effect of cholesterol in the TM12 interface was also investigated. It was observed that, with cholesterol in the membrane, the interaction strength decreases between the SERTs. It was observed that cholesterol densities at the dimer interface around TM12 is higher than monomer SERT. This shows that cholesterol at the dimer interface is necessary for dimerization. This work showed that membrane composition is important for dimerization, and different lipids can affect the oligomerization of SERTs differently.

Effect of PIP2 on DAT

Sitte and coworkers investigated the effect of lipids on dimerization of DAT. To study the effect of cholesterol, a cholesterol oxidase was used to oxidize the cholesterol molecules in the membrane (Christian et al. 1997). The altered membrane did not have any effect on the concentration of dimers. Further, the role of PIP2 in the stability of DAT dimer was investigated. PIP2 is known to interact with the N-terminus of hDAT electrostatically, which leads to impaired dopamine efflux

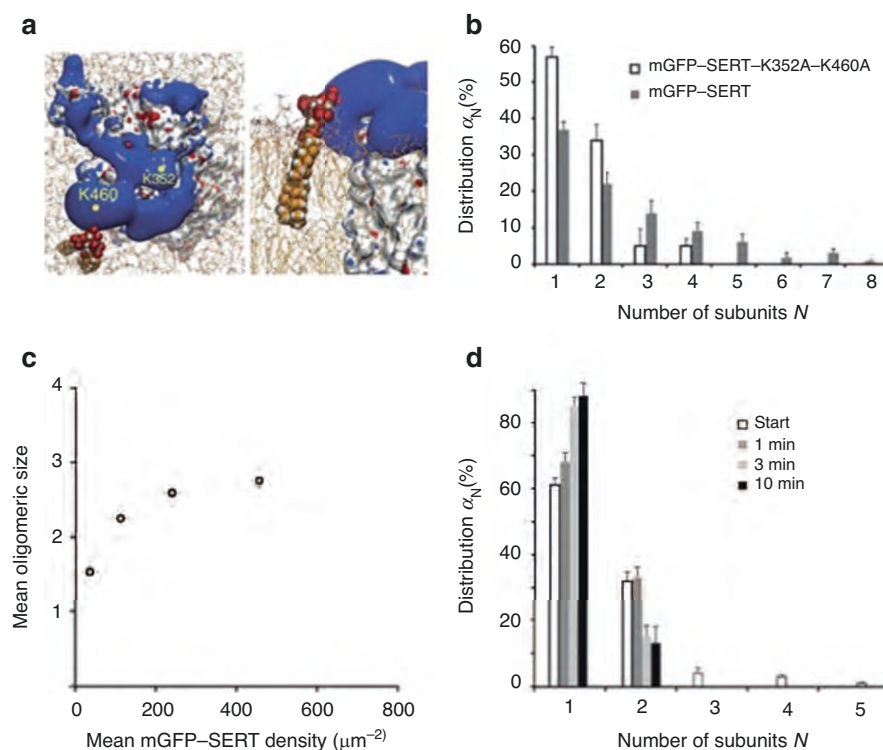


Fig. 7.5 Analysis of the electrostatic field generated by SERT. **(a)** The final structure of a 100 ns simulation of a membrane-inserted SERT is shown as viewed from the cytosol (left) or in the side view (right). The SERT surface is shown in white and the electrostatic isosurfaces are shown in red (negative potential) and blue (positive potential). For illustration purposes, a PIP2 molecule was placed into the membrane (in space-filled representation) in close proximity to the large positively charged area that includes residue K460. **(b)** The quaternary assembly of the mutant mGFP-SERT-K352A-K460A at the plasma membrane (white bars). **(c)** A pronounced dependence of mean oligomeric state on mGFP-SERT-K352A-K460A surface density was observed, which saturates around 2.8 transporter molecules per oligomer. **(d)** Repetitive TOCCSL runs revealed rapid protomer exchange kinetics for this mutant ($n = 18$ cells). Anderluh, A., Hofmaier, T., Klotzsch, E. et al. “Direct PIP2 binding mediates stable oligomer formation of the serotonin transporter.” *Nature communications* 8, no. 1: 14089, 2017 (Anderluh et al. 2017). This article was originally published and distributed under the terms of the Creative Commons Attribution License, permitting unrestricted reproduction and adaptation provided proper crediting to author and source

(Hamilton et al. 2014). A PLC activator was used to hydrolyze the PIP2 molecule (Bae et al. 2003). The concentration of dimerized DAT did not change in the absence of PIP2. Also, the total concentration of DAT expressed on the membrane did not change. This indicated that DAT forms stable dimers but PIP2 does not have any major role in the stability of DAT dimers.

Functional Role of Oligomerization in MATs

The formation of oligomers by the MAT transporters has established that the transporters are present in monomer, dimer, and higher oligomeric forms in the plasma membrane. To study the functionality of the dimers and compare this with monomeric structures, the Bahar group conducted multiple studies on dopamine transporter (DAT) (Gur et al. 2017). In a computational study, they studied the effect of dimerization on the dynamics of neurotransmitter sodium symporter (NSS), LeuT, and hDAT dimers (Gur et al. 2017; Cheng et al. 2017). hDAT does not have a crystal structure and was modeled using dDAT and LeuT structures as a template. To assess the significance of oligomerization in monomeric and dimeric forms, an anisotropic network model (ANM) was used (Atilgan et al. 2001). ANM predicts the global modes that correspond to the principal modes of motion observed in long timescale MD simulations. When applied to the dimeric transporter, ANM predicted modes that correspond to transition from OF to IF states. However, in the monomeric state, ANM does not predict transition from OF to IC states, which suggests that dimerization favors the transition of each unit from OF to IF states. This suggests that dimeric transporters function more efficiently than the monomeric transporters.

In another study by Bahar and coworkers, allosteric modulation of the dimer interface of DAT revealed the factors regulating the function of the transporter (Cheng et al. 2017). In the section “Structural Features of Dimer Interface of MATs,” we have discussed the kink present in TM12 of MAT transporters that parallel alignment of the protomers. In this study, the residues forming the kink at TM12 were mutated to their counterpart in LeuT (G561A, S567F, S568L, or P573A). In all the mutant hDATs, a decrease in surface expression was observed. Significant loss in uptake activity of the transport was observed for G561A and P573A mutants; however, the uptake activity was observed to increase for S567F and S568L mutants. The study suggested that even though S567F and S568L mutants have higher uptake activity, they cause an impairment in the trafficking of DAT to the plasma membrane. This suggested that mutations that hinder the formation of dimeric DAT also impede the function of the transporter. This suggests that dimer DAT functions more efficiently in comparison to monomer DAT and allosteric modulation that hinders dimerization of the transporter also regulates its transport activity in a negative manner. Zhen and Reith (2018) studied the functional properties of DAT oligomers after copper linking. Copper phenanthroline (CuP) was used as a cross-linking tool to measure the functional uptake properties of DAT monomers and oligomers. The effect of increasing cross-linking of CuP was studied on the distribution of dimers present and on the uptake of activity of the dimers. It was observed that increasing the concentrations of CuP led to an increase in the fraction of dimers present; however, the uptake activity decreased. It was suggested that Cu⁺ plays an inhibitory role when linked to Cys306, which is an important site for dimer interface formation, and the dimer transporter functions at a diminished rate compared to a monomer transporter. The study also concluded that in a dimeric DAT, only one protomer

remains active at a time, similar to SRT and NET (Kilic and Rudnick 2000; Kocabas et al. 2003; Larsen et al. 2011).

Conclusions and Future Directions

The MAT family of transporters has been observed to form oligomers at the plasma membrane via various experimental and computational techniques. This chapter focused on the effect of membrane composition on dimerization of transporters. All the members of the MAT subfamily have been shown to form stable dimers by subunit exchange studies. NET is observed to form dimers in the presence of PIP₂, SERT can form higher-order oligomers, and DAT oligomerization does not depend on the presence of PIP₂ molecules in the plasma membrane. Single-molecule microscopy studies showed that PIP₂ lipid modulates the oligomerization of the transporters. Further studies are required to answer why DAT does not depend on PIP₂ for dimerization even though it shares a similar fold as NET and SERT.

Computational and experimental studies on the dimer interface of the transporters revealed multiple possible oligomeric conformations for DAT and SERT. SERT and DAT form dimers via TM12-TM12 interaction, similar to LeuT despite having a kink in TM12. The pocket formed by the kinks of TM12-TM12 dimer interface acts as a lipid-binding site that can regulate the oligomerization of the transporters. Computational studies have observed higher densities of cholesterol in the dimer interface pockets, but further studies on the role of lipids binding in the dimer pocket are required. Experimental studies show that palmitoylation of TM12 also modulates the energy of dimer formation, suggesting that modifications to TM12 that alters the pocket reduce the propensity to form dimers. Experimental techniques like deep mutational scans which have been used to investigate transporter function could also be applied to the study of the dimeric interface that can help elucidate the residues important for the formation of stable dimers (Park et al. 2019; Young et al. 2021; Braberg et al. 2022).

The studies on oligomerization of MAT transporters have provided evidence that the MAT family of transporters can form oligomers and many factors have been identified that affect the oligomerization. However, the question about the role of oligomerization in increasing the transport efficiency of the transporters remains to be answered. This question can be answered using molecular simulation techniques that have been applied to investigate the transport cycle (Chan et al. 2022b). In particular, long timescale MD simulations of the transport dimers could provide a detailed atomistic view of the impact of oligomerization on the transport rate. Furthermore, atomistic molecular dynamics could be validated using protein-protein association results obtained by experiments (Jacobs et al. 2022). Similarly, the effect of posttranslational modifications (PTMs) (Weigle et al. 2022) such as phosphorylation (Chan et al. 2022a) and glycosylation (Chan and Shukla 2023) has been investigated for the monomeric MATs, but the interplay between these PTMs is not well established. Machine learning methods are also useful in characterizing

protein-protein interactions based on sequence, structure, domain-domain interactions, etc. (Tang et al. 2023; Shamsi et al. 2020; Horne and Shukla 2022).

Acknowledgments We would like to thank Austin T. Weigle and Xuenan Mi from Shukla Group for their valuable comments on this book chapter. DS acknowledges support from the National Institutes of Health Award R35GM142745.

References

- Anderluh A, Klotzsch E, Reismann AW, Brameshuber M, Kudlacek O, Newman AH, Sitte HH, Schutz GJ (2014) Single molecule analysis reveals coexistence of stable serotonin transporter monomers and oligomers in the live cell plasma membrane. *J Biol Chem* 289(7):4387–4394
- Anderluh A, Hofmaier T, Klotzsch E, Kudlacek O, Stockner T, Sitte HH, Schütz GJ (2017) Direct PIP₂ binding mediates stable oligomer formation of the serotonin transporter. *Nat Commun* 8(1):14089
- Apparsundaram S, Galli A, DeFelice LJ, Hartzell HC, Blakely RD (1998a) Acute regulation of norepinephrine transport: I. Protein kinase c-linked muscarinic receptors influence transport capacity and transporter density in SK-N-SH cells. *J Pharmacol Exp Ther* 287(2):733–743
- Apparsundaram S, Schroeter S, Giovanetti E, Blakely RD (1998b) Acute regulation of norepinephrine transport: II. PKC-modulated surface expression of human norepinephrine transporter proteins. *J Pharmacol Exp Ther* 287(2):744–751
- Atilgan AR, Durell S, Jernigan RL, Demirel MC, Keskin O, Bahar I (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 80(1):505–515
- Bae Y-S, Lee TG, Park JC, Hur JH, Kim Y, Heo K, Kwak J-Y, Suh P-G, Ryu SH (2003) Identification of a compound that directly stimulates phospholipase C activity. *Mol Pharmacol* 63(5):1043–1050
- Beckstein O, Naughton F (2022) General principles of secondary active transporter function. *Biophys Rev* 3(1)
- Belovich AN, Aguilar JI, Mabry SJ, Cheng MH, Zanella D, Hamilton PJ, Stanislawski DJ, Shekar A, Foster JD, Bahar I et al (2021) A network of phosphatidylinositol (4,5)-bisphosphate (PIP₂) binding sites on the dopamine transporter regulates amphetamine behavior in drosophila melanogaster. *Mol Psychiatry* 26(8):4417–4430
- Berger SP, Farrell K, Conant D, Kempner ES, Paul SM (1994) Radiation inactivation studies of the dopamine reuptake transporter protein. *Mol Pharmacol* 46(4):726–731
- Braberg H, Echeverria I, Kaake RM, Sali A, Krogan NJ (2022) From systems to structure—using genetic data to model protein structures. *Nat Rev Genet* 23(6):342–354
- Buchmayer F, Schicker K, Steinkellner T, Geier P, Stübiger G, Hamilton PJ, Jurik A, Stockner T, Yang J-W, Montgomery T et al (2013) Amphetamine actions at the serotonin transporter rely on the availability of phosphatidylinositol-4,5-bisphosphate. *Proc Natl Acad Sci* 110(28):11642–11647
- Chan MC, Shukla D (2023) The effects of n-linked glycosylation on slc6 transporters. *J Chem Inf Model* 63(9):2748–2758
- Chan MC, Procko E, Shukla D (2022a) Structural rearrangement of the serotonin transporter intracellular gate induced by Thr276 phosphorylation. *ACS Chem Neurosci* 13(7):933–945
- Chan MC, Selvam B, Young HJ, Procko E, Shukla D (2022b) The substrate import mechanism of the human serotonin transporter. *Biophys J* 121(5):715–730
- Chen N, Reith ME (2008) Substrates dissociate dopamine transporter oligomers. *J Neurochem* 105(3):910–920

- Cheng MH, Garcia-Olivares J, Wasserman S, DiPietro J, Bahar I (2017) Allosteric modulation of human dopamine transporter activity under conditions promoting its dimerization. *J Biol Chem* 292(30):12471–12482
- Christian A, Haynes MP, Phillips M, Rothblat G (1997) Use of cyclodextrins for manipulating cellular cholesterol content. *J Lipid Res* 38(11):2264–2272
- Coleman JA, Green EM, Gouaux E (2016) X-ray structures and mechanism of the human serotonin transporter. *Nature* 532(7599):334–339
- Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* 20(1):45–50
- Das AK, Kudlacek O, Baumgart F, Jaentsch K, Stockner T, Sitte HH, Schütz GJ (2019) Dopamine transporter forms stable dimers in the live cell plasma membrane in a phosphatidylinositol 4,5-bisphosphate-independent manner. *J Biol Chem* 294(14):5632–5642
- DeLano WL et al (2002) PyMOL: an open-source molecular graphics tool. *CCP4 Newsl Protein Crystallogr* 40(1):82–92
- Drew D, Boudker O (2016) Shared molecular mechanisms of membrane transporters. *Annu Rev Biochem* 85:543–572
- Dutta Chowdhury U, Paul A, Bhargava B (2023) Phosphatidylinositol (PI) lipids modulate the binding of tau fibrils on lipid bilayers. *bioRxiv*. <https://doi.org/10.1101/2023.09.20.558589>
- Egaña LA, Cuevas RA, Baust TB, Parra LA, Leak RK, Hochendoner S, Peña K, Quiroz M, Hong WC, Dorostkar MM et al (2009) Physical and functional interaction between the dopamine transporter and the synaptic vesicle protein synaptogyrin-3. *J Neurosci* 29(14):4592–4604
- Forrest LR (2013) (Pseudo-)symmetrical transport. *Science* 339(6118):399–401
- Forrest LR, Rudnick G (2009) The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters. *Physiology* 24(6):377–386
- Forrest LR, Krämer R, Ziegler C (2011) The structural basis of secondary active transport mechanisms. *Biochim Biophys Acta* 1807(2):167–188
- Foster JD, Vaughan RA (2011) Palmitoylation controls dopamine transporter kinetics, degradation, and protein kinase c-dependent regulation. *J Biol Chem* 286(7):5175–5186
- Giros B, Caron MG (1993) Molecular characterization of the dopamine transporter. *Trends Pharmacol Sci* 14(2):43–49
- Gotfryd K, Boesen T, Mortensen JS, Khelashvili G, Quick M, Terry DS, Missel JW, LeVine MV, Gourdon P, Blanchard SC et al (2020) X-ray structure of LeuT in an inward-facing occluded conformation reveals mechanism of substrate release. *Nat Commun* 11(1):1005
- Gur M, Cheng MH, Zomot E, Bahar I (2017) Effect of dimerization on the dynamics of neurotransmitter: sodium symporters. *J Phys Chem B* 121(15):3657–3666
- Hamilton PJ, Belovich AN, Khelashvili G, Saunders C, Erreger K, Javitch JA, Sitte HH, Weinstein H, Matthies HJ, Galli A (2014) PIP2 regulates psychostimulant behaviors through its interaction with a membrane protein. *Nat Chem Biol* 10(7):582–589
- Hastrup H, Karlin A, Javitch JA (2001) Symmetrical dimer of the human dopamine transporter revealed by cross-linking Cys-306 at the extracellular end of the sixth transmembrane segment. *Proc Natl Acad Sci* 98(18):10055–10060
- Hastrup H, Sen N, Javitch JA (2003) The human dopamine transporter forms a tetramer in the plasma membrane: cross-linking of a cysteine in the fourth transmembrane segment is sensitive to cocaine analogs. *J Biol Chem* 278(46):45045–45048
- Henderson RK, Fendler K, Poolman B (2019) Coupling efficiency of secondary active transporters. *Curr Opin Biotechnol* 58:62–71
- Horne J, Shukla D (2022) Recent advances in machine learning variant effect prediction tools for protein engineering. *Ind Eng Chem Res* 61(19):6235–6245
- Jacobs MI, Bansal P, Shukla D, Schroeder CM (2022) Understanding supramolecular assembly of supercharged proteins. *ACS Cent Sci* 8(9):1350–1361
- Jardetzky O (1966) Simple allosteric model for membrane pumps. *Nature* 211(5052):969–970
- Jayaraman K, Das AK, Luethi D, Szöllösi D, Schütz GJ, Reith ME, Sitte HH, Stockner T (2021) SLC6 transporter oligomerization. *J Neurochem* 157(4):919–929

- Jess U, Betz H, Schloss P (1996) The membrane-bound rat serotonin transporter, SERT1, is an oligomeric protein. *FEBS Lett* 394(1):44–46
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A et al (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873):583–589
- Just H, Sitte HH, Schmid JA, Freissmuth M, Kudlacek O (2004) Identification of an additional interaction domain in transmembrane domains 11 and 12 that supports oligomer formation in the human serotonin transporter. *J Biol Chem* 279(8):6650–6657
- Kilic F, Rudnick G (2000) Oligomerization of serotonin transporter and its functional consequences. *Proc Natl Acad Sci* 97(7):3106–3111
- Kocabas A, Rudnick G, Kilic F (2003) Functional consequences of homo- but not hetero-oligomerization between transporters for the biogenic amine neurotransmitters. *J Neurochem* 85(6):1513–1520
- Krishnamurthy H, Gouaux E (2012) X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* 481(7382):469–474
- Kristensen AS, Andersen J, Jørgensen TN, Sørensen L, Eriksen J, Loland CJ, Strømgaard K, Gether U (2011) SLC6 neurotransmitter transporters: structure, function, and regulation. *Pharmacol Rev* 63(3):585–640
- Larsen MB, Sonders MS, Mortensen OV, Larson GA, Zahniser NR, Amara SG (2011) Dopamine transport by the serotonin transporter: a mechanistically distinct mode of substrate translocation. *J Neurosci* 31(17):6605–6615
- Lassmann T, Sonnhammer EL (2006) Kalign, kalignvu and mumsa: web servers for multiple sequence alignment. *Nucleic Acids Res* 34(suppl 2):W596–W599
- Liepiņa I, Czaplowski C, Janmey P, Liwo A (2003) Molecular dynamics study of a gelsolin-derived peptide binding to a lipid bilayer containing phosphatidylinositol 4,5-bisphosphate. *Pept Sci* 71(1):49–70
- Liu JJ, Hezghia A, Shaikh SR, Cenido JF, Stark RE, Mann JJ, Sublette ME (2018) Regulation of monoamine transporters and receptors by lipid microdomains: implications for depression. *Neuropsychopharmacology* 43(11):2165–2179
- Luethi D, Liechti ME (2020) Designer drugs: mechanism of action and adverse effects. *Arch Toxicol* 94(4):1085–1133
- Luethi D, Maier J, Rudin D, Szöllösi D, Angenooth TJ, Stankovic S, Schittmayer M, Burger I, Yang J-W, Jaentsch K et al (2022) Phosphatidylinositol 4,5-bisphosphate (PIP₂) facilitates nor-epinephrine transporter dimerization and modulates substrate efflux. *Commun Biol* 5(1):1259
- McLean MA, Stephen AG, Sligar SG (2019) PIP₂ influences the conformational dynamics of membrane-bound KRAS4b. *Biochemistry* 58(33):3537–3545
- Milner HE, Beliveau R, Jarvis SM (1994) The in situ size of the dopamine transporter is a tetramer as estimated by radiation inactivation. *Biochim Biophys Acta* 1190(1):185–187
- Moertelmaier M, Brameshuber M, Linimeier M, Schutz GJ, Stockinger H (2005) Thinning out clusters while conserving stoichiometry of labeling. *Appl Phys Lett* 87(26)
- Neubauer HA, Hansen CG, Wiborg O (2006) Dissection of an allosteric mechanism on the serotonin transporter: a cross-species study. *Mol Pharmacol* 69(4):1242–1250
- Pacholczyk T, Blakely RD, Amara SG (1991) Expression cloning of a cocaine- and antidepressant-sensitive human noradrenaline transporter. *Nature* 350(6316):350–354
- Park J, Selvam B, Sanematsu K, Shigemura N, Shukla D, Procko E (2019) Structural architecture of a dimeric class C GPCR based on co-trafficking of sweet taste receptor subunits. *J Biol Chem* 294(13):4759–4774
- Periole X, Zeppelin T, Schjøtt B (2018) Dimer interface of the human serotonin transporter and effect of the membrane composition. *Sci Rep* 8(1):5080
- Picas L, Viaud J, Schauer K, Vanni S, Hnia K, Fraissier V, Roux A, Bassereau P, Gaits-Iacovoni F, Payrastré B et al (2014) BIN1/M-amphiphysin2 induces clustering of phosphoinositides to recruit its downstream partner dynamin. *Nat Commun* 5(1):5647

- Rastedt DE, Vaughan RA, Foster JD (2017) Palmitoylation mechanisms in dopamine transporter regulation. *J Chem Neuroanat* 83:3–9
- Rothman RB, Baumann MH (2003) Monoamine transporters and psychostimulant drugs. *Eur J Pharmacol* 479(1–3):23–40
- Schloss P, Williams DC (1998) The serotonin transporter: a primary target for antidepressant drugs. *J Psychopharmacol* 12(2):115–121
- Schmid JA, Scholze P, Kudlacek O, Freissmuth M, Singer EA, Sitte HH (2001) Oligomerization of the human serotonin transporter and of the rat GABA transporter 1 visualized by fluorescence resonance energy transfer microscopy in living cells. *J Biol Chem* 276(6):3805–3810
- Shamsi Z, Chan M, Shukla D (2020) TLMutation: predicting the effects of mutations using transfer learning. *J Phys Chem B* 124(19):3845–3854
- Singh SK, Piscitelli CL, Yamashita A, Gouaux E (2008) A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* 322(5908):1655–1661
- Sitte HH, Freissmuth M (2015) Amphetamines, new psychoactive drugs and the monoamine transporter cycle. *Trends Pharmacol Sci* 36(1):41–50
- Sorkina T, Doolen S, Galperin E, Zahniser NR, Sorkin A (2003) Oligomerization of dopamine transporters visualized in living cells by fluorescence resonance energy transfer microscopy. *J Biol Chem* 278(30):28274–28283
- Sugita Y, Kitao A, Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. *J Chem Phys* 113(15):6042–6051
- Tang T, Zhang X, Liu Y, Peng H, Zheng B, Yin Y, Zeng X (2023) Machine learning on protein–protein interaction prediction: models, challenges and trends. *Brief Bioinform* 24(2):bbad076
- Torres GE, Carneiro A, Seamans K, Fiorentini C, Sweeney A, Yao W-D, Caron MG (2003a) Oligomerization and trafficking of the human dopamine transporter: mutational analysis identifies critical domains important for the functional expression of the transporter. *J Biol Chem* 278(4):2731–2739
- Torres GE, Gainetdinov RR, Caron MG (2003b) Plasma membrane monoamine transporters: structure, regulation and function. *Nat Rev Neurosci* 4(1):13–25
- Wang KH, Penmatsa A, Gouaux E (2015) Neurotransmitter and psychostimulant recognition by the dopamine transporter. *Nature* 521(7552):322–327
- Weigle AT, Feng J, Shukla D (2022) Thirty years of molecular dynamics simulations on posttranslational modifications of proteins. *Phys Chem Chem Phys* 24(43):26371–26397
- Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E (2005) Crystal structure of a bacterial homologue of Na⁺/Cl[−]-dependent neurotransmitter transporters. *Nature* 437(7056):215–223
- Young HJ, Chan M, Selvam B, Szymanski SK, Shukla D, Procko E (2021) Deep mutagenesis of a transporter for uptake of a non-native substrate identifies conformationally dynamic regions. *bioRxiv*. <https://doi.org/10.1101/2021.04.19.440442>
- Zeppelin T, Pedersen KB, Berglund NA, Periole X, Schiøtt B (2021) Effect of palmitoylation on the dimer formation of the human dopamine transporter. *Sci Rep* 11(1):4164
- Zhen J, Reith ME (2018) Functional properties of dopamine transporter oligomers after copper linking. *J Neurochem* 144(2):162–171

Chapter 8

Macromolecular Interactions of Lipoprotein Lipase (LPL)



Anna Wheless, Kathryn H. Gunn, and Saskia B. Neher

Abstract Lipoprotein lipase (LPL) is a critical enzyme in humans that provides fuel to peripheral tissues. LPL hydrolyzes triglycerides from the cores of lipoproteins that are circulating in plasma and interacts with receptors to mediate lipoprotein uptake, thus directing lipid distribution via catalytic and non-catalytic functions. Functional losses in LPL or any of its myriad of regulators alter lipid homeostasis and potentially affect the risk of developing cardiovascular disease—either increasing or decreasing the risk depending on the mutated protein. The extensive LPL regulatory network tunes LPL activity to allocate fatty acids according to the energetic needs of the organism and thus is nutritionally responsive and tissue dependent. Multiple pharmaceuticals in development manipulate or mimic these regulators, demonstrating their translational importance. Another facet of LPL biology is that the oligomeric state of the enzyme is also central to its regulation. Recent structural studies have solidified the idea that LPL is regulated not only by interactions with other binding partners but also by self-associations. Here, we review the complexities of the protein-protein and protein-lipid interactions that govern LPL structure and function.

Keywords Lipoprotein lipase · Apolipoprotein · Triglyceride · Angiopoietin-like protein · Lipase maturation factor 1 · Syndecan-1 · Helix · CryoEM · Heparin

A. Wheless · S. B. Neher (✉)

University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

e-mail: awheless@unc.edu; neher@email.unc.edu

K. H. Gunn

University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Stony Brook University, Stony Brook, USA

e-mail: kathryn.gunn@stonybrook.edu

© The Author(s), under exclusive license to Springer Nature

Switzerland AG 2024

J. R. Harris, J. Marles-Wright (eds.), *Macromolecular Protein Complexes V*,
Subcellular Biochemistry 104, https://doi.org/10.1007/978-3-031-58843-3_8

List of Abbreviations

Albumin	Human serum albumin
Aliases	serum albumin, HSA
ANGPTL3	Angiopoietin-like protein 3
Aliases	ANGL3, ANL3, ANG-5, ANG5, ANGPT5
ANGPTL4	Angiopoietin-like protein 4
Aliases	ANGL4, Fasting-induced adipose factor, FIAF, ARP4, HFARP, PGAR
ANGPTL8	Angiopoietin-like protein 8
Aliases	ANGL8, Lipasin, Betatrophin, TD26, Refeeding-induced fat and liver protein (RIFL)
APOA5	Apolipoprotein A5
Aliases	ApoA5, ApoA-5, ApoAV, ApoA-V, Apolipoprotein A-5, Apolipoprotein AV, Apolipoprotein A-V
APOC1	Apolipoprotein C1
Aliases	ApoC1, ApoC-1, ApoCI, ApoC-I, Apolipoprotein C-1, Apolipoprotein CI, Apolipoprotein C-I, Apolipoprotein-serine (obsolete), R-Ser (obsolete)
APOC2	Apolipoprotein C2
Aliases	ApoC2, ApoC-2, ApoCII, ApoC-II, Apolipoprotein C-2, Apolipoprotein CII, Apolipoprotein C-II, Apolipoprotein-glutamic acid (obsolete), ApoLp-Glu (obsolete), R-Glu (obsolete)
APOC3	Apolipoprotein C3
Aliases	ApoC3, ApoC-3, ApoCIII, ApoC-III, Apolipoprotein C-3, Apolipoprotein CIII, Apolipoprotein C-III, Apolipoprotein-alanine (obsolete), ApoLp-Ala (obsolete), R-Ala (obsolete)
APOE	Apolipoprotein E
GPIHBP1	Glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1
LDLR	Low-density lipoprotein receptor
Aliases	LDL receptor, LDLr
LMF1	Lipase maturation factor 1
LPL	Lipoprotein lipase
LRP	Low-density lipoprotein receptor-related protein
Aliases	LDL receptor-related protein, LRP1
SDC1	Syndecan-1
SEL1L	Suppressor/enhancer of Lin-12-like
VLDLR	Very low-density lipoprotein receptor
Aliases	VLDL receptor, VLVLr

Introduction to LPL

The idea of an enzyme such as LPL was initially recognized by Paul Hahn in 1943. He observed that the intravenous administration of heparin could clear lipemic plasma, and hence there must be a plasma lipid “clearing factor” (Hahn 1943). Later studies showed that this clearing factor reduced the density of lipoprotein particles (Graham et al. 1951). Indeed, the heparin-treated plasma could clear lipidemic plasma from another individual, indicating that heparin released a factor into the bloodstream capable of acting on lipoproteins (Graham et al. 1951; Anderson and Fawcett 1950). Multiple groups subsequently suggested that the lipid clearing action of the post-heparin plasma was caused by an enzyme acting on the lipoproteins (Anfinsen et al. 1952; Boyle et al. 1952). In 1954, Edward Korn extracted the clearing factor from rat heart acetone powder and named the enzyme lipoprotein lipase, as it “denotes both its specificity and the type of reaction which it catalyzes (Korn 1954).” Additional groups endeavored to purify the enzyme from blood or tissue (Korn 1955; Brown et al. 1953; Hollett and Meng 1956; Fielding 1969; Reichl 1969), with the eventual discovery that heparin resin could be used to purify the enzyme to homogeneity (Olivecrona et al. 1971). It was not until 1986 that a partial protein sequence for LPL was determined (Ben-Avram et al. 1986; Reddy et al. 1986; Bengtsson-Olivecrona et al. 1986). Shortly thereafter, in 1987, the cDNA sequence for LPL was resolved (Enerback et al. 1987).

LPL has direct control over the level of triglycerides (TGs) in circulation in humans and other vertebrates. As LPL hydrolyzes the TGs within triglyceride-rich lipoproteins (TRLs), the fatty acid products of the reaction are taken up by nearby cells and stored or used. Thus, the reaction catalyzed by LPL has two functions critically important to our lipid metabolism: maintaining healthy levels of the reactants (TGs) and distribution of the products (non-esterified fatty acids, NEFAs). The *location* and *timing* of LPL activity determine which tissues do or do not receive fatty acids. The *level* of LPL activity determines circulating TG levels and, more specifically, the distribution of lipoproteins of varying densities (lipid profile), which has implications for cardiovascular disease (Le and Walter 2007). Fine tuning of LPL activity by endogenous macromolecular activators and inhibitors is crucial for a healthy lipid profile.

LPL and lipoproteins meet on the surface of capillary endothelial cells. There, LPL is bound by heparin sulfate proteoglycans and/or its anchoring partner GPI-anchored high-density lipoprotein binding protein 1 (GPIHBP1) (Beigneux et al. 2007). When LPL encounters a hydrophobic substrate, it undergoes interfacial activation in which the lid peptide swings away from the active site (van Tilbeurgh et al. 1993). With the active site accessible, LPL catalyzes the hydrolysis of the TG component of the lipoproteins. This reaction yields 2-monoacylglycerol and NEFAs. These NEFAs can be used by the underlying tissues, as they can be oxidized to provide energy or re-esterified for storage in adipose tissue. Following hydrolysis of their TG component, chylomicron remnants are taken up by the liver for recycling, whereas VLDL remnants have greater density and are known as intermediate density lipoproteins (IDLs). Further TG hydrolysis of IDLs yields low-density lipoproteins (LDLs).

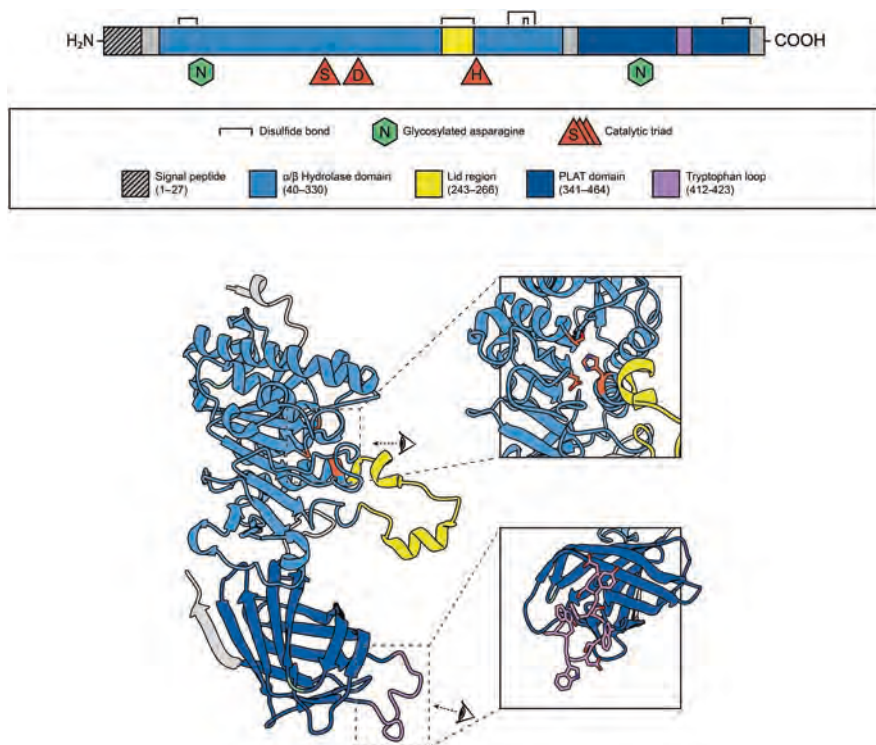


Fig. 8.1 LPL Domain Architecture. LPL is comprised of 475 amino acids, the first 27 of which are the signal sequence, which is cleaved prior to LPL secretion from the cell. Mature LPL contains two N-linked glycosylations (residues 70 and 386) and five disulfide bonds (54–67, 243–266, 291–310, 302–305, 445–465) (Hayne et al. 2018; Ben-Zeev et al. 1994). The ABHD (40–330) of LPL is characterized by an α -helix/ β -sheet interaction containing 8 β -strands and 6 α -helices (Ollis et al. 1992), and the PLAT/LH2 domain (341–464) is composed of 2 β -sheets, forming a β -sandwich (Bateman and Sandford 1999). The LPL active site is in the ABHD and is comprised of three conserved residues (S159, D183, H268) (Emmerich et al. 1992). The LPL lid peptide (243–266), also located in the ABHD, covers the entry to the active site residues (Dugi et al. 1995). LPL has heparin-binding residues that span both the ABHD and PLAT domains (R306, K307, R309, K323, R324, K430, R432, K434, K440, K441) (Sendak and Bensadoun 1998; Hata et al. 1993). The LPL tryptophan loop (412–423) is located in the PLAT domain and interacts with triglyceride-rich lipoproteins to play a role in substrate recognition (Lookene et al. 1997). Where stick representations of side chains are shown, oxygen is represented in red and nitrogen in blue. Protein structures rendered with ChimeraX (Pettersen et al. 2021)

The mature LPL amino acid sequence is approximately 50.5 kDa (without glycans) with an N-terminal α/β hydrolase domain (ABHD) and a C-terminal polycystin-1, lipoxigenase, and alpha-toxin (PLAT) domain, also called a lipoxigenase homology (LH2) domain as illustrated in Fig. 8.1. Within these two domains, LPL has features that play important roles in both activity and binding. In this chapter, we will use amino acid numbering for LPL and all other secreted proteins that includes the signal sequence. Residue numbers will come from human sequences for consistency and clarity, although some of these proteins are conserved across species.

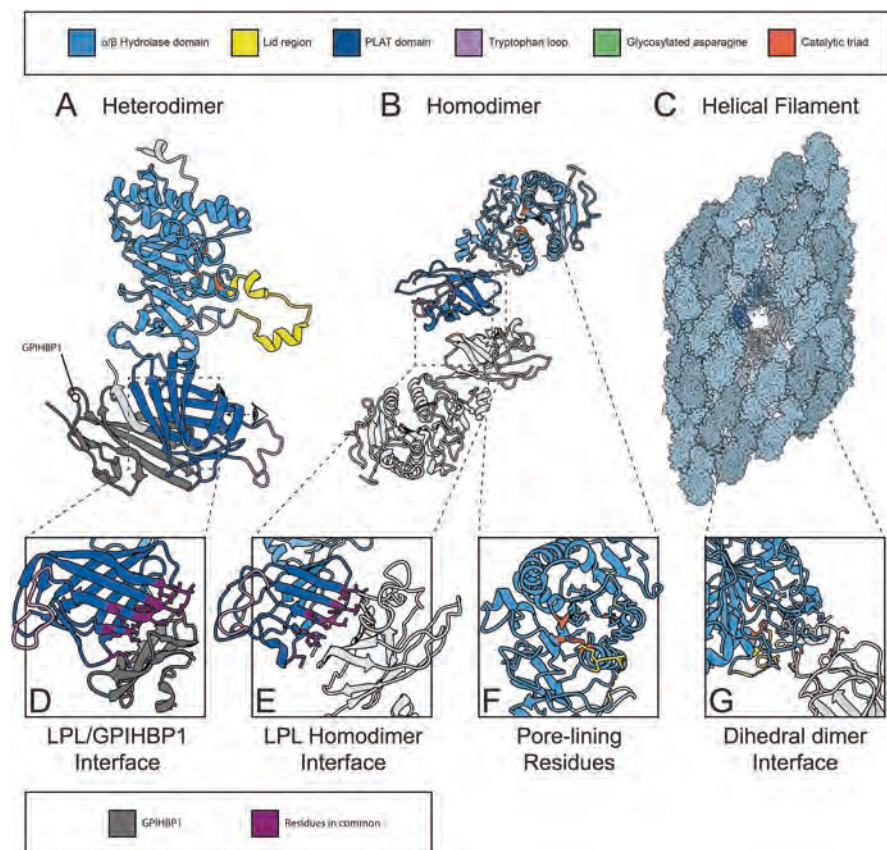


Fig. 8.2 LPL Oligomers. Three structures of LPL have been solved (Gunn and Neher 2023; Gunn et al. 2020; Birrane et al. 2019; Arora et al. 2019). Where stick representations of side chains are shown, oxygen is represented in red, nitrogen in blue, and sulfur in yellow. (a) The LPL/GPIHBP1 heterodimer (6OB0), (b) the LPL homodimer (8ERL), and (c) the LPL helical filament (6U7M). (d–e) The LPL/GPIHBP1 interaction interface overlaps with the LPL homodimer interface, suggesting homodimerization and GPIHBP1 binding are mutually exclusive. (f) The active sites pore is mostly lined with hydrophobic side chains, suggesting a hydrophobic substrate or product may travel through it. (g) In the inactive helical filament, the tryptophan loop of one LPL is in close contact with the active site residues of the opposing LPL in the dihedral dimer

Structural Studies of LPL Through the Years

Structures of three LPL-containing complexes have been determined: the LPL/GPIHBP1 heterodimer, the LPL helical filament, and the LPL homodimer (Fig. 8.2) (Gunn and Neher 2023; Gunn et al. 2020; Birrane et al. 2019; Arora et al. 2019). These structures have upended the previous paradigms of LPL activity and oligomerization set by biochemical investigation. For decades, it was believed that LPL was active primarily as a homodimer. The original basis for this belief was a radiation inactivation experiment in 1983 that assessed the molecular weight of the active LPL oligomer. The researchers found the size of active LPL was 127 ± 15 kDa,

which most closely corresponds to an LPL homodimer (the molecular weight of a single LPL is approximately 50 kDa) (Garfinkel et al. 1983). This data was followed by sedimentation analysis in 1985 which found enzymatic activity was solely correlated with the LPL homodimer, and activity was not observed in the presence of LPL monomers or higher-order LPL oligomers (Osborne et al. 1985). This work also suggested that the LPL homodimers were in equilibrium with monomers. These monomers could convert back into active homodimers or decay to an inactive form (Osborne et al. 1985).

The idea that LPL homodimers are active and monomers inactive remained undisputed for over 30 years until crystal structures were solved of LPL in complex with GPIHBP1 (Birrane et al. 2019; Arora et al. 2019). Initial analysis of the first of these structures suggested that LPL and GPIHBP1 formed a heterodimer, and two of these complexes came together to form a heterotetramer. In this structure, the LPL subunits interacted in a head-to-tail fashion (the N-terminal domain of one subunit interacts with the C-terminal domain of the other) (Birrane et al. 2019). A second structure of LPL/GPIHBP1 was then solved, which resolved the residues facilitating the reported LPL dimerization interface at the center of the putative heterotetramer (Arora et al. 2019). This led to the conclusion that the LPL-to-LPL association in the crystal structure was an effect of crystallization, not a physiological interface (Arora et al. 2019). This analysis was supported by size-exclusion chromatography multi-angle light scattering (SEC-MALS) data showing that LPL/GPIHBP1 is a heterodimer (Fig. 8.2a), not a heterotetramer, in solution (Arora et al. 2019). Biochemically, it has been shown that the LPL/GPIHBP1 complex is active (Sonnenburg et al. 2009), and it had been hypothesized following the discovery of GPIHBP1 that it would interact with the active LPL homodimer, so the finding of an LPL/GPIHBP1 heterodimer was unexpected.

In 1990, it was reported that LPL could form a concentration-dependent, inactive oligomer in cells prior to secretion (Pradines-Figuères et al. 1990). This theory was developed due to the observation that upon dilution, lysate from cells expressing LPL showed an unexpected increase in LPL enzymatic activity (Pradines-Figuères et al. 1990). In 2020, cryogenic electron microscopy (cryoEM) revealed the structure of a filament composed of repeating LPL dimer subunits related by helical symmetry (Fig. 8.2c) (Gunn et al. 2020). Within the helical structure, the LPL dimers adopt a head-to-tail interaction resembling the crystallographic interface seen in the LPL/GPIHBP1 structures (Gunn et al. 2020; Birrane et al. 2019; Arora et al. 2019). However, in the helical filament, the interface was highly energetically favorable, as it brought together the hydrophobic residues from LPL's active site pocket and the substrate recognition loop of each LPL and buried them, creating an inactive structure (Fig. 8.2g) (Gunn et al. 2020). Residues in the helical interfaces overlap with the LPL/GPIHBP1 interface indicating these two structures are not compatible—GPIHBP1 must bind separately from the LPL helical filament (Gunn et al. 2020). In this oligomeric form, LPL is folded correctly but inactive. Accompanying analysis by fluorescence microscopy supported the hypothesis that this oligomer may be a key element of LPL storage in secretory vesicles.

The most recent structure of LPL available to date (also determined by cryoEM) is that of the active LPL homodimer (Fig. 8.2b). Figure 8.2e shows this structure's

novel homodimerization interface located between the C-terminal domains of each LPL, creating a tail-to-tail interface (Gunn and Neher 2023). The interacting residues in the tail-to-tail interface overlap with both the LPL/GPIHBP1 interface and the helical interfaces, suggesting that these three interactions cannot occur simultaneously in a single complex (Gunn and Neher 2023).

LPL Oligomers

LPL Monomer

As mentioned above, sedimentation assays showed that LPL monomers are inactive (Osborne et al. 1985). However, it was also demonstrated that LPL monomers and dimers are in equilibrium and that a portion of LPL monomers remained competent to become active dimers (Lookene et al. 2004). The finding that the LPL/GPIHBP1 complex involved a single LPL subunit rather than a homodimer reignited debate on the topic (Birrane et al. 2019; Arora et al. 2019). Density gradient ultracentrifugation was used to compare the LPL activity peak to known molecular weight standards (Beigneux et al. 2019). At low concentrations, the location of LPL in the gradient matched most closely to a monomeric molecular weight running slightly after the BSA standard (66 kDa) and before muscle phosphorylase b (97 kDa) (Beigneux et al. 2019). The ability of both BSA (Levi and Gonzalez Flecha 2002) and muscle phosphorylase b (Chebotareva et al. 2001) to form dimers in solution, however, makes the results difficult to confidently interpret. Recent investigations with mass photometry suggest that at low concentrations (16 nM) LPL is a monomer in solution (Gunn and Neher 2023). The structure and potential activity of LPL monomers remains an area of investigation.

LPL Homodimer

Before the structure of LPL was first determined, researchers drew on the crystal structure of a LPL homolog, pancreatic lipase, to create a homology model to approximate the dimeric structure of LPL (van Tilbeurgh et al. 1994). Pancreatic lipase is known to be active as a monomer but had formed a dimer during crystallization, providing an ideal starting point for LPL modeling (van Tilbeurgh et al. 1992, 1993). This modeling helped visualize the architecture of the LPL dimer forming a head-to-tail (N-terminal to C-terminal; ABHD to PLAT) dimerization interface. Further modeling was undertaken to understand how LPL might rearrange when activated (Kobayashi et al. 2002). It was also shown biochemically that an LPL with an active site mutation and an LPL with a substrate recognition mutation were able to dimerize to recover LPL activity (Kobayashi et al. 2002). As the structure of the LPL dimer still proved elusive, further work was performed combining biochemical data with molecular modeling and single-molecule fluorescence

distance measurements to generate a more accurate dimeric model (Hayne et al. 2018). In this model, the LPL dimer still took on a head-to-tail formation, but the surfaces of LPL which interacted with each other flipped (Hayne et al. 2018).

The first structure of the LPL homodimer, shown in Fig. 8.2b, was solved in 2023 and revealed a tail-to-tail interface where the two PLAT domains interact (367–372, 403–406, 447–463) (Gunn and Neher 2023). In the structure, LPL has undergone interfacial activation, wherein the lid peptide, which normally covers the active site, slides to the side (Gunn and Neher 2023). Interfacial activation is triggered by association with a nonpolar-aqueous interface, such as a triglyceride-rich lipoprotein (Khan et al. 2017). In addition to having undergone interfacial activation, the LPL homodimer structure revealed a hydrophobic pore adjacent to the active site in the ABHD (Fig. 8.2f) (Gunn and Neher 2023). This is the first active site pore seen in mammalian lipase structures and suggests an added layer of substrate discrimination. It resembles the tunnel which controls substrate access to the *Candida rugosa* lipase active site (Grochulski et al. 1994; Kingsley and Lill 2015). The pore could also facilitate unidirectional transfer of the acyl chain following hydrolysis, so LPL can stay associated with a lipoprotein and still pass NEFAs to the surrounding environment (Gunn and Neher 2023).

LPL Helical Filament

In 2020, the structure of a helical filament of LPL was solved by cryoEM (Gunn et al. 2020). This filament was both concentration dependent and inactive, suggesting it could be the higher order form found in the cell (Gunn et al. 2020). The LPL helix had dihedral symmetry and was composed of repeating subunits of LPL dimers (Gunn et al. 2020). The dimerization interface rendered the protein inactive, with the substrate binding tryptophan loop of one LPL blocking access to the active site of the other LPL. The repeating subunits interact in a series of helical interfaces creating a stable filament with a diameter of 254 Å. Super-resolution microscopy of endogenous LPL in adipocytes showed LPL forming filament-like structures inside of vesicles (Gunn et al. 2020). When mutations were made in the helical interface, the filament structures were no longer observed in vesicles (Gunn et al. 2020).

Dynamics of LPL Oligomers

In solution, LPL has been shown to be in rapid equilibrium between monomers and dimers (Lookene et al. 2004; Osborne et al. 1985). Beyond these dynamics, LPL's oligomeric state depends on a great number of factors that are still not fully understood, including LPL concentration, the presence of substrate, additives, and binding partners.

LPL Concentration

LPL concentration appears to play a role in oligomerization at both high and low concentrations. At a low concentration (16 nM), mass photometry shows that LPL forms a single molecular weight peak corresponding to a monomer of LPL. However, when the concentration is increased to 100 nM, the molecular weight shifts, suggesting LPL is predominantly forming dimers (Gunn and Neher 2023). At high concentrations of LPL, such as 2 μ M, the LPL helical oligomer is observed. The filament exists in the presence of other oligomers, and there is a concentration threshold for filament formation which can be modulated by the addition of heparin (Gunn et al. 2020). Both heparin and albumin have both also been observed to play a role in the concentration dependent formation of reversibly inactive LPL oligomers (see section “LPL and Extracellular Binding Partners”) (Gunn et al. 2020; Risti et al. 2023).

Substrate

LPL's native substrates are the TGs packaged inside of ApoB-containing lipoproteins (Zambon et al. 1996). Dietary TGs are assembled into chylomicrons in the intestine. These largest and most buoyant of lipoprotein particles are comprised of a neutral lipid core, a phospholipid monolayer coat, and the presence of the truncated form of apolipoprotein B, ApoB48 (Kane et al. 1980). Endogenously synthesized TGs are assembled into VLDLs in the liver. These large and buoyant lipoprotein particles are marked by the full-length form of apolipoprotein B, ApoB100 (Kane et al. 1980). Substrate affects LPL's oligomeric form. When intralipid, a substrate mimetic, is added to LPL filaments, they disassemble and are no longer observed as higher-order oligomers (Gunn et al. 2020). The presence of TRLs or NEFAs has also been shown to drive dissociation of LPL from GPIHBP1 (Shetty et al. 2020). Indirect evidence for the effect of substrate on oligomeric state comes from the structure of the LPL homodimer. In solution, cross-linking of LPL reveals an array of oligomers, including monomer, homodimer, homotrimer, and homotetramer (Gunn and Neher 2023). However, this plethora of oligomeric states is at odds with the homodimers observed in cryoEM micrographs, where LPL was only observed adopting homodimers or homotetramers (2 homodimers coming together) (Gunn and Neher 2023). One theory is that the air-water interface (AWI) formed during grid preparation is creating an aqueous-nonpolar interface reminiscent of LPL substrate. In this way, LPL could have interacted with the AWI as though it was a substrate mimic, offering an explanation for why the protein has undergone interfacial activation, and suggests that substrate may trigger preferential formation of the LPL homodimer (Gunn and Neher 2023). The potential for a substrate or substrate mimic to alter the oligomeric state of LPL regardless of concentration should be considered, particularly when interpreting activity assay data.

Bile Salt (Deoxycholate)

Deoxycholate is a bile salt that binds to and stabilizes LPL from thermal inactivation *in vitro* (Bengtsson and Olivecrona 1979). However, there has been debate in the field that deoxycholate may not represent a physiological interactor of LPL because the circulating deoxycholate concentration is ~350 nM (Frommherz et al. 2016) and can vary greatly between individuals (Fiamoncini et al. 2017). Recent work using cross-linking has shown that deoxycholate-treated LPL adopts a wide range of oligomers in solution (monomer, dimer, trimer, tetramer), which are identical to those seen in LPL that has not been treated with deoxycholate (Gunn and Neher 2023). Mass photometry experiments have also shown that 16 nM LPL mixed with deoxycholate forms monomers, but, as with LPL alone (and LPL with heparin), 100 nM LPL treated with deoxycholate adopts a dimeric form (Gunn and Neher 2023). This stabilizing additive, then, does not appear to influence the distribution of non-helical oligomers of LPL. Deoxycholate has, however, been found to disrupt the formation of LPL filaments (Gunn et al. 2020). The structure of the LPL homodimer was determined in the presence of deoxycholate to prevent filament formation, but individual deoxycholate molecules were not resolved bound to the structure (Gunn and Neher 2023).

Regulation of LPL Folding

LPL is a particularly labile protein that is prone to misfolding and aggregation, perhaps because of the same characteristics that enable its self-association and affinity for hydrophobic substrates. To avert misfolding and accumulation of aggregates, multiple chaperones assist with LPL folding in the ER. Loss of either of two ER-resident proteins, suppressor/enhancer of Lin-12-like (SEL1L) and lipase maturation factor 1 (LMF1), results in decreased LPL secretion and thus causes hypertriglyceridemia in mice and humans (Peterfy et al. 2007; Sha et al. 2014). The precise mechanism by which SEL1L and LMF1 promote LPL ER exit is not known, but studies suggest that both proteins prevent and resolve LPL aggregation, as outlined below, and illustrated in Fig. 8.3.

LMF1

Prior to the discovery of the gene for *LMF1*, a naturally occurring truncation mutation in *LMF1* was known as the *cld* mutation for combined lipase deficiency (Paterniti et al. 1983). This name came from the observation that the genes for LPL and hepatic lipase were unaffected, and the gene products were synthesized, but no lipase was secreted in *cld* mice (Oka et al. 1989b; Olivecrona et al. 1986; Masuno

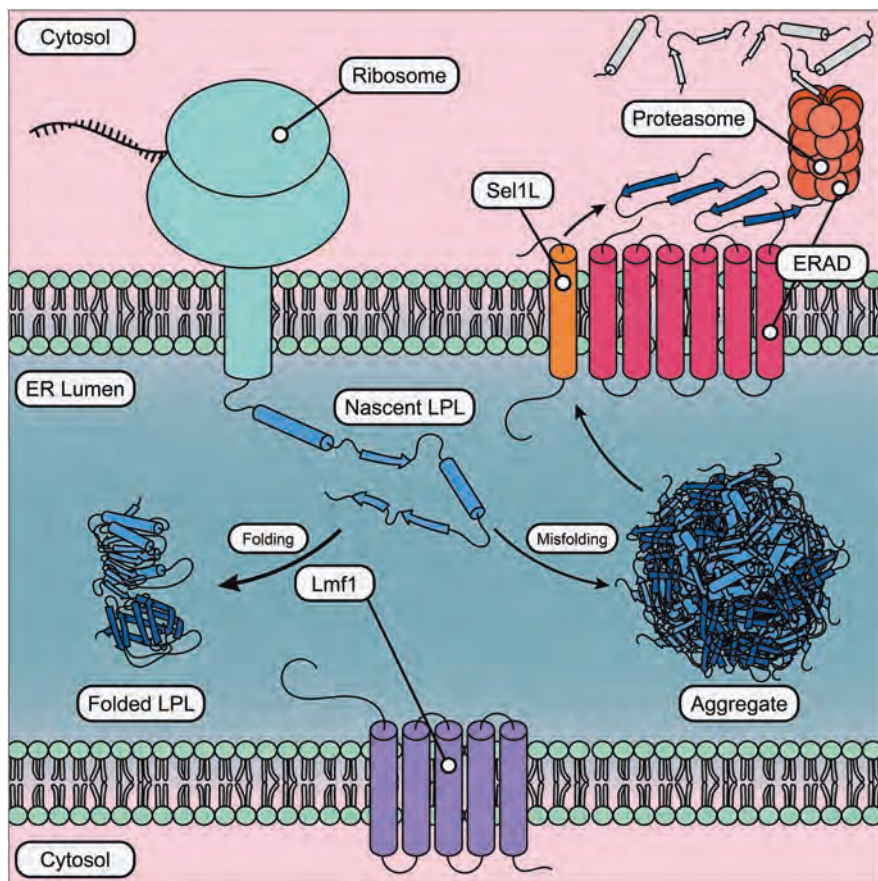


Fig. 8.3 LPL folding. After LPL is translated, it can either fold correctly and be exported for secretion or misfold and aggregate in the lumen of the endoplasmic reticulum. Lmf1 promotes proper LPL folding. Sel1L in complex with other ERAD components facilitates the removal and degradation of aggregated LPL

et al. 1990). As a result, newborn mice homozygous for the mutation died of hyperchylomicronemia shortly after birth (Paterniti et al. 1983). Later genetic mapping revealed that the *cld* mutation was a C-terminal gene truncation caused by insertion of the long terminal repeat of a murine endogenous retrovirus into intron 7 (Peterfy et al. 2007). The gene containing the *cld* mutation was renamed *LMF1*, for lipase maturation factor 1 (Peterfy et al. 2007). *LMF1* encodes a multi-pass membrane protein that is localized in the endoplasmic reticulum (Peterfy et al. 2007; Babilonia-Rosa and Neher 2014). Expression of *LMF1* is induced by ER stress through the ATF6 arm of the unfolded protein response (Mao et al. 2014). By mass spectrometry, *LMF1* associates with multiple ER-resident chaperones and redox proteins such as protein disulfide isomerases (Roberts et al. 2018). Loss of *LMF1* results in accumulation of large, disulfide-bonded aggregates of LPL (Roberts et al. 2018).

Additionally, the ER in cells lacking LMF1 is more oxidized than the ER of WT cells (Roberts et al. 2018). It is thus clear that LMF1 promotes the proper folding and ER exit of LPL. Current evidence suggests that LMF1 does so by promoting an ER environment that helps to resolve mis-paired disulfides in LPL.

SEL1L

SEL1L was identified in humans in 1997, as a homologue of the suppressor of the *lin-12* mutant in *Caenorhabditis elegans* (Biunno et al. 1997). SEL1L is an ER-resident, single-pass transmembrane protein, with most of the protein in the ER lumen. In mice, loss of SEL1L is embryonic lethal, and analysis of embryonic tissue from knockout animals revealed ER stress and morphological changes to the ER (Francisco et al. 2010). Indeed, SEL1 is upregulated by the unfolded protein response (UPR) and contributes to the degradation of unfolded proteins in the ER (Francisco et al. 2010). Specifically, SEL1L nucleates a multiprotein complex that contributes to the dislocation of misfolded glycoproteins from the ER lumen (Mueller et al. 2006, 2008). SEL1L is the mammalian homolog of the yeast protein HRD3, which associates with HRD1, a transmembrane ubiquitin ligase (Mueller et al. 2006). Recent structural studies show how the HRD1 complex (comprised of HRD1, SEL1L, Der1, Usa1, and Yos9) functions to translocate ERAD substrates from the ER lumen to the cytosol, where they are polyubiquitinated (Wu et al. 2020).

SEL1L was implicated in LPL processing in the ER when mice with an adipocyte-specific knockout of Sel1L were found to have postprandial hypertriglyceridemia (Sha et al. 2014). The mice had both smaller adipocytes and hepatic steatosis, and these phenotypes were due to LPL retention in the ER of the adipocytes, resulting in a lack of secreted LPL (Sha et al. 2014). In this study, the authors concluded that Sel1L plays a role independent of ERAD in regulating LPL secretion. However, their subsequent studies support the hypothesis that LPL's propensity to misfold causes a critical need for SEL1L-HRD1 ERAD to clear the misfolded LPL and allow nascent LPL to fold and become active (Wu et al. 2021, 2023).

LPL Trafficking

Syndecan 1

Syndecan-1 (SDC1) is a transmembrane heparan sulfate proteoglycan (HSPG) with roles in cell proliferation, cell migration, and intracellular trafficking. SDC1 is comprised of an extracellular ectodomain, a membrane spanning helix, and a small cytoplasmic domain that can interact with regulatory factors. The ectodomain contains heparan-sulfate chains that bind many ligands, and this domain can be cleaved from

the membrane domain by the action of matrix metalloproteinases (Rapraeger 2001; Fitzgerald et al. 2000). Sundberg et al. showed that SDC1 plays a role in sorting LPL in the trans-Golgi network (TGN) (Sundberg et al. 2019). They discovered that SDC1 is found in the sphingomyelin-rich microdomains of the TGN that bud exocytic vesicles (Sundberg et al. 2019). Additionally, they found that the heparan sulfate chains of SDC1 bind to known heparan-interacting regions of LPL and recruit it to specific exocytic vesicles (Sundberg et al. 2019). LPL present in the vesicles is likely to be organized as a helical filament, as super-resolution microscopy studies showed filament-like LPL surrounded by SDC1 in adipocytes (Gunn et al. 2020). SDC1 may play additional roles after LPL is trafficked out of the TGN. Reizes and colleagues found that adipocytes express both SDC-1 and syndecan-4 on their cell surface and when insulin triggers shedding of the syndecan ectodomains, both ectodomains bound to and stabilized LPL (Reizes et al. 2006).

HSPGs and Heparin

LPL has several heparin-binding sites that interact with HSPGs in vivo (Lookene et al. 1996; Lutz et al. 2001). In the cell, LPL interacts with the HSPG SDC1 (Sundberg et al. 2019), as well as HSPGs in the interstitial space and in the capillaries (Allan et al. 2017; Sarrazin et al. 2011; Meneghetti et al. 2015). Injecting heparin into the blood causes release of LPL from the capillary walls, greatly increasing the amount of LPL collected in a blood draw (Vilella et al. 1993; Zambon et al. 1996). The addition of heparin to purified LPL alters the way LPL travels through a sedimentation gradient, creating a broader peak of LPL mass (Beigneux et al. 2019), suggesting that heparin could be altering the oligomeric state of LPL. It has also been observed that heparin alters the concentration dependence of the LPL filament, specifically that LPL filaments bound to heparin persist at lower concentrations than filaments without heparin (Gunn et al. 2020). Mass photometry experiments also show that LPL mixed with heparin at 16 nM forms monomers, but as with LPL alone at 100 nM LPL with heparin adopts a predominantly dimeric form (Gunn and Neher 2023). The addition of heparin to LPL activity assays has been shown to stabilize LPL activity—interestingly, this affect does not seem to necessarily stem from an oligomeric affect but rather stabilization of the monomer form of LPL, preventing inactivation (Lookene et al. 1996, 2004).

GPIHBP1

Following its release into the interstitial space, LPL is taken up by GPIHBP1. This protein was named for its discovery as a protein that caused HDL to bind to the cell surface (Ioka et al. 2003). GPIHBP1 was later found to play a key role in the processing of chylomicrons, as mice deficient in this gene had severe chylomicronemia

because of reduced post-heparin LPL activity (Beigneux et al. 2007). GPIHBP1 is a glycan-anchored protein that resides on the capillary endothelial cell membrane (Davies et al. 2010). GPIHBP1 binds with LPL in the interstitial space and is essential for LPL transport across the capillary endothelium and into the capillary itself, where LPL can interact with substrates (Davies et al. 2010). This finding differed from prior studies that showed that LPL bound to the luminal surface of endothelial cells via heparin-like glycosaminoglycan residues (Oka et al. 1989a; Cheng et al. 1981; Shimada et al. 1981).

Highlighting the importance of GPIHBP1 to LPL transport into the capillary, LPL is mislocalized to the interstitial spaces in *Gpihbp1*-deficient mice (Davies et al. 2010). Several crystal structures resolved the interaction of the GPIHBP1 Ly6/uPar (LU) domain with LPL's PLAT domain (367–374, 401–406, and 443–471) (Fig. 8.2d) (Arora et al. 2019; Birrane et al. 2019). In both studies, the asymmetric unit is a LPL/GPIHBP1 heterotetramer (two heterodimers), with the LPL subunits interacting with each other in a “head-to-tail” fashion (with the N-terminus of one lined up with the C-terminus of the other). The second of the two studies found that the LPL/LPL interface at the center of the tetramer was likely to be an effect of crystallization and not a physiological interface. A second known binding interface with the N-terminal acidic domain of GPIHBP1 and LPL has not been structurally determined (Reimund et al. 2015). The acidic domain is predicted to be unstructured (Mysling et al. 2016a), although it may adopt a set fold when associated with LPL. The acidic domain interacts with the heparin-binding sites on LPL (Reimund et al. 2015), disrupting its interaction with the HSPGs in the interstitial space, which is essential in moving LPL into the capillary (Song et al. 2022).

LPL and Extracellular Binding Partners

Albumin

Recent work using surface plasmon resonance has shown that albumin directly interacts with LPL (Risti et al. 2023). It had previously been thought that the role of albumin was indirect—as a molecular crowder and binder of NEFA products in the capillary. In combination, Raster image correlation spectroscopy and electron microscopy revealed that, at 200 nM LPL, albumin induces the formation of reversible LPL aggregates (Risti et al. 2023). These albumin-induced aggregates are concentration dependent. When examined using electron microscopy, the aggregates were amorphous and unstructured (Risti et al. 2023). Like the concentration-dependent LPL filament, the activity of the LPL from the LPL/albumin aggregates is recovered upon dilution to lower concentration (5 nM) (Risti et al. 2023). Of note, when heparin and albumin are both added to 200 nM LPL, no aggregates either helical or amorphous were observed (Risti et al. 2023). Further work on the structural details of the interaction of LPL and albumin remains to be done, but this suggests

that the presence of albumin and HSPGs in the interstitial space could trigger LPL filament dissociation, allowing transit of LPL across the interstitial space to reach GPIHBP1; see Fig. 8.4 (Risti et al. 2023).

Angiopoietin-Like Proteins

The extensive posttranslational macromolecular regulatory network of LPL includes three of the angiopoietin-like (ANGPTL) proteins: ANGPTL3, ANGPTL4, and ANGPTL8 (Fig. 8.5). These proteins downregulate LPL activity in response to nutritional cues to meet tissue-specific metabolic demands (Zhang 2016). ANGPTL3 and ANGPTL4 both have a N-terminal coiled-coil domain (CCD) and a C-terminal fibrinogen-like domain (FLD) (Fig. 8.5a). The CCD mediates self-oligomerization as well as inhibition of lipoprotein lipase, whereas the FLD is implicated in receptor binding and angiogenesis (Camenisch et al. 2002; Bini et al. 2022; Yoshida et al. 2002; Ono et al. 2003).

ANGPTL3

In 1999, ANGPTL3 was identified by searching expressed sequence tag databases for secreted proteins that form amphipathic helices (Conklin et al. 1999). The link between ANGPTL3 and lipid metabolism was subsequently determined by genetic mapping in hypolipidemic mice (Shimizu et al. 2002). Due to loss of ANGPTL3's inhibitory effect on LPL, mice deficient for ANGPTL3 not only had lower plasma triglycerides but also had decreased atherosclerotic burden (Ando et al. 2003). In human patients, individuals with complete deficiency in ANGPTL3 were found to have less atherosclerosis, as compared to unaffected controls (Stitzel et al. 2017). With the powerful inhibitory effect of ANGPTL3 on LPL clearly established, ANGPTL3 has become an attractive therapeutic target for treatment of hypertriglyceridemia. Vupanorsen, an antisense oligonucleotide against ANGPTL3, was shown to reduce non-HDL-C in humans before it was discontinued (Bergmark et al. 2022). Evinacumab, a monoclonal antibody against ANGPTL3, is a currently approved treatment for familial hypercholesterolemia and hypertriglyceridemia (Sosnowska et al. 2022; Dewey et al. 2017).

The N-terminal CCD of ANGPTL3 is necessary and sufficient for LPL inhibition, but mutations throughout the C-terminal FLD are also linked to lowered triglyceride levels and are thought to destabilize the structure or decrease secretion (Biterova et al. 2018). The CCD of ANGPTL3 forms trimers in solution (Fig. 8.5b) and small-angle X-ray scattering (SAXS) envelopes show that it is elongated and flexible (SASDJL8) (Gunn et al. 2021). Within the CCD, a region encompassing residues 32-55 was reported to be essential for LPL inhibition (Lee et al. 2009).

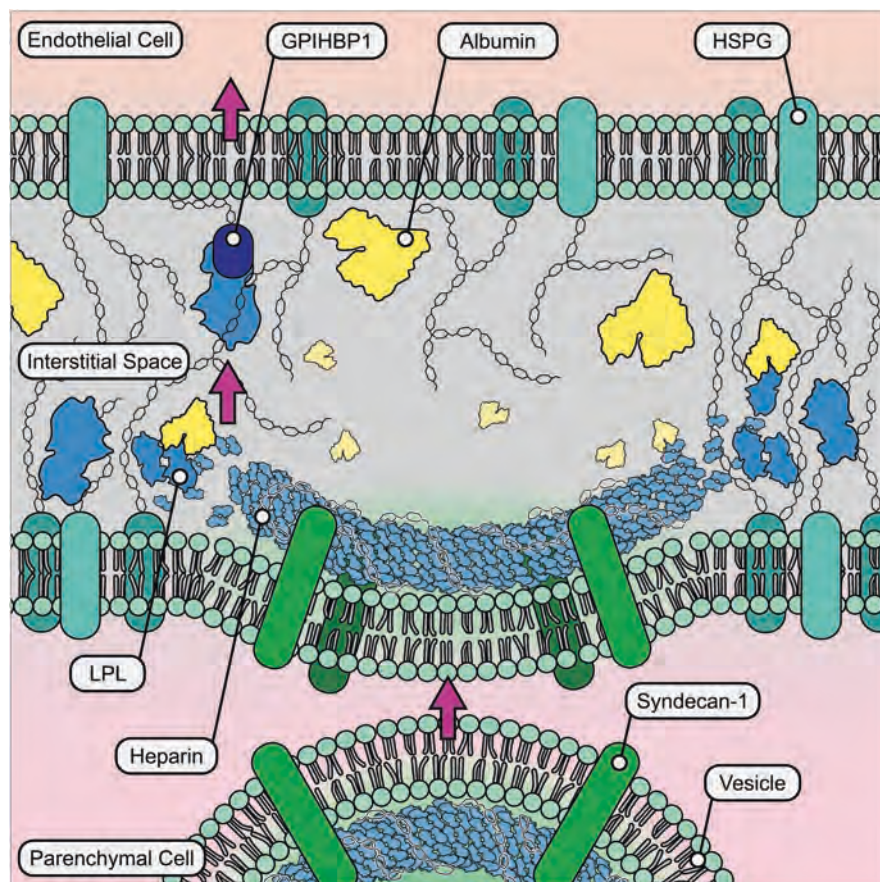


Fig. 8.4 LPL trafficking. LPL-containing vesicles fuse with the cell membrane and release their cargo. The heparin chains attached to syndecan-1, an HSPG, may be stabilizing LPL in a filamentous form during transport. Once LPL helices are in the interstitial space, they dissolve. The change in concentration and the presence of albumin may facilitate filament dissociation. LPL monomers or dimers are bound by GPIHBP1 on the endothelial cell surface before being transported across the cell to reach the capillary lumen

ANGPTL4 and ANGPTL8 are nutritionally regulated, whereas ANGPTL3 is constitutively secreted (Quagliarini et al. 2012; Ge et al. 2005). ANGPTL3 forms a complex with ANGPTL8, and mass spectrometry analyses of proteins pulled down from human serum suggest that ANGPTL3 and ANGPTL8 interact in a 3:1 ratio in vivo (Chen et al. 2020), which agrees with the SEC-MALS observations that ANGPTL3 forms a trimer in vitro (Gunn et al. 2021). Initial reports of ANGPTL3 versus ANGPTL3/8 inhibition of LPL concluded that ANGPTL3/8 is a significantly more potent inhibitor than ANGPTL3 alone (Chi et al. 2017). A new purification protocol to remove contaminating DNA from recombinantly expressed ANGPTL3 greatly increased its potency, showing in vitro that ANGPTL3 can inhibit LPL with

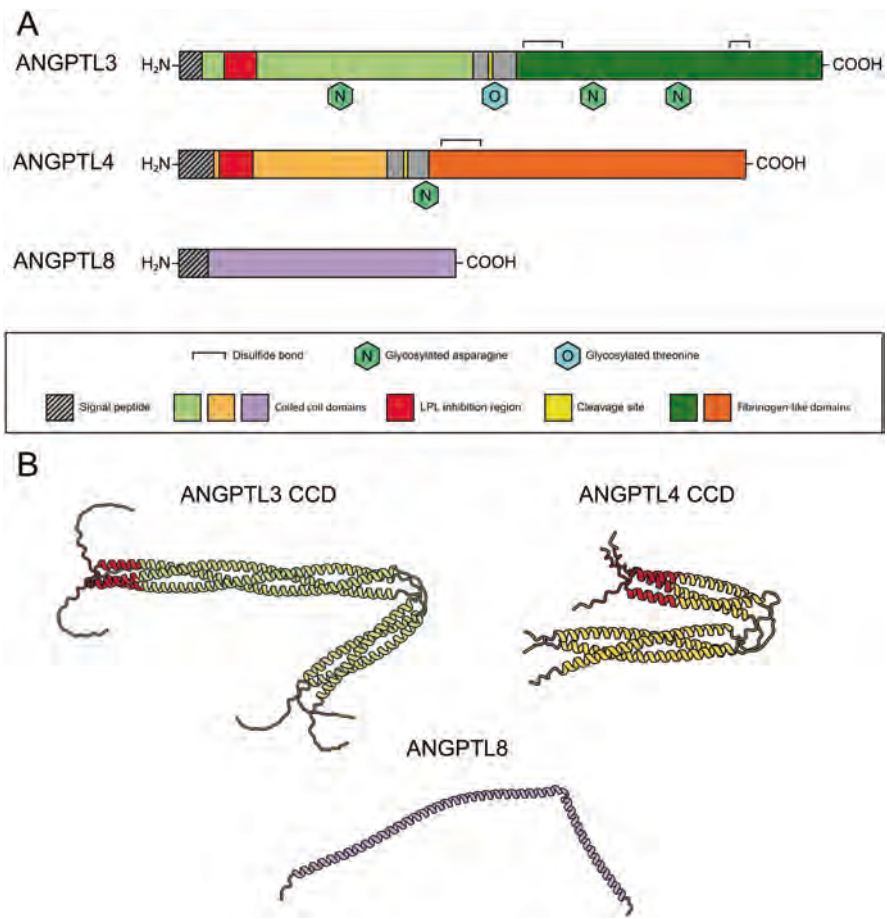


Fig. 8.5 ANGPTL Domain Architecture. (a) ANGPTL3, ANGPTL4, and ANGPTL8 domain maps. ANGPTL3 and ANGPTL4 both contain N-terminal coiled coil domains (CCDs) and C-terminal fibrinogen-like domain (FLDs). The CCDs and FLDs of each are separated by an unstructured region that contains a cleavage site. ANGPTL8 is the only member of the ANGPTL family that lacks the FLD. Regions of ANGPTL3 and ANGPTL4 implicated in LPL inhibition are shown in red and are located in the N-terminal chains. ANGPTL8 has not been shown to inhibit LPL directly on its own, and so no LPL inhibition region is indicated on its map. (b) AlphaFold (Mirdita et al. 2022) structural predictions for the CCDs of ANGPTL3, ANGPTL4, and ANGPTL8. Structures are colored to match domain architecture maps. Regions of interest include the LPL inhibitory regions of ANGPTL3 and ANGPTL4 shown in red and the single N-linked glycosylation site in the CCD of ANGPTL3

a nanomolar K_i without ANGPTL8 present (Gunn et al. 2021). Jin and colleagues observed that evinacumab protected LPL from ANGPTL3/8-mediated inhibition in vitro, indicating that the antibody is potentially effective against ANGPTL3/8 in addition to ANGPTL3 for which it was designed (Jin et al. 2021). Other antibodies produced against ANGPTL3 do not necessarily block the action of ANGPTL3/8. In

2017, with endpoint experiments performed in cell culture medium, Haller and colleagues showed that an ANGPTL3-blocking antibody did not block LPL inhibition when ANGPTL8 was coexpressed with ANGPTL3. They also found that when ANGPTL8 was coexpressed with mutated, inactive ANGPTL3, LPL was still inhibited. ANGPTL8 alone or mutant ANGPTL3 alone had little to no effect on LPL activity, but in combination, these proteins could inhibit LPL (Haller et al. 2017). Interestingly, although LPL is homologous to endothelial lipase (EL) and ANGPTL3 inhibits both lipases, ANGPTL8 does not affect ANGPTL3 inhibition of EL, potentially pointing to a key difference in the mechanisms of inhibition of the two lipases (Sylvers-Davie et al. 2021; Balasubramaniam et al. 2022).

Investigations of the mechanism by which ANGPTL3 inhibits LPL are complicated by, and interwoven with, changing notions of LPL activity and deactivation. Prior to the very recent discovery of catalytically active LPL monomers (Beigneux et al. 2019), inhibition of LPL was often interpreted as dissociation of “active LPL dimers” to “inactive LPL monomers” (Sukonina et al. 2006; Osborne et al. 1985). Today, there are different hypotheses for mechanisms of LPL inhibition by the ANGPTL proteins, including promotion of furin cleavage of LPL (Jin et al. 2021), destabilization and unfolding of the LPL hydrolase domain (Kristensen et al. 2020), and reversible noncompetitive inhibition (Gunn et al. 2021). An HDX-MS study found that ANGPTL3/8 protects the LPL furin cleavage site from deuterium exchange, and a follow-up cell culture experiment showed that ANGPTL3/8 promoted cleavage of LPL after secretion from HEK cells (Jin et al. 2021). A study using native LPL substrate and a standard Michaelis-Menton kinetics framework showed that ANGPTL3 inhibits LPL via a reversible and noncompetitive mechanism (Gunn et al. 2021). The HDX-MS study focused exclusively on ANGPTL3/8 and is a method designed to measure conformational dynamics, while the Michaelis-Menton kinetics experiment measured the effect of ANGPTL3 alone on LPL catalytic activity. The two analyses are thus not entirely comparable. A structure of the complex between LPL and ANGPTL3 or ANGPTL3/8 is still unknown, but a recent study modeled a complex of LPL and ANGPTL3/8 (Balasubramaniam et al. 2022).

ANGPTL4

ANGPTL4 was identified in 2000 as a protein secreted from adipose tissue in response to fasting and as a target of peroxisome proliferator-activated receptor-gamma (PPAR- γ) signaling (Kersten et al. 2000; Kim et al. 2000; Yoon et al. 2000). Its function in LPL inhibition was revealed by Yoshida and colleagues in 2002 after they observed that mice injected with ANGPTL4 had significantly increased triglyceride levels and confirmed the inhibitory effect on LPL in vitro (Yoshida et al. 2002). Similarly, adenoviral overexpression of ANGPTL4 increased plasma TG levels in mice (Ge et al. 2004). ANGPTL4 knockout mice showed higher VLDL clearance and thus lower plasma TGs levels (Desai et al. 2007). Human carriers of an E40K mutation in ANGPTL4 have a cardioprotective phenotype of lower

circulating triglyceride levels, lower LDL-C, and higher HDL-C (Romeo et al. 2007; Dewey et al. 2016; Talmud et al. 2008; McQueen et al. 2017). Thus, mABs against ANGPTL4 were explored as novel drugs to treat hypertriglyceridemia. However, these treatments were ultimately deemed unsuccessful following the observation of lipid accumulation in the mesenteric lymph nodes in nonhuman primates (Dewey et al. 2016). Recently, ANGPTL4 silencing via antisense oligonucleotides has been tested in mice, and this study did not report the lymphatic lesions previously observed with knockout mice and anti-ANGPTL4 antibody treatment (Deng et al. 2022). Today, multiple aspects of ANGPTL4 function are contentious, including its oligomerization, cleavage, mechanism of LPL inhibition, and the connections between these facets.

The linker between the CCD and FLD of ANGPTL4 contains a proprotein convertase cleavage motif (Lei et al. 2011) (Fig. 8.5a). Both the full-length and cleaved forms of ANGPTL4 are found in circulation (Chen et al. 2020), and the FLD is linked to receptor binding, angiogenesis, and intracellular TG lipolysis (Biterova et al. 2018; Kim et al. 2000). Like in ANGPTL3, the CCD as characterized by SAXS is an elongated and flexible structure (SASDJM8). SEC-MALS analysis suggests that the CCD of ANGPTL4 forms a homotrimer in solution (Fig. 8.5b) (Gunn et al. 2021). The CCD of ANGPTL4 contains two cysteine residues, C76 and C80, that can form intermolecular disulfide bonds, though whether they do so in vivo is uncertain. Two studies showed that ANGPTL4 could form disulfide-bonded dimers and tetramers when analyzed by gel electrophoresis and that mutation of either C76 or C80 to alanine allowed disulfide-bonded dimers, but mutation of both residues blocked disulfide formation and ANGPTL4 inhibition of LPL (Yin et al. 2009; Ge et al. 2004). A later study making the more conservative mutation of either C76 or C80 to serine yielded ANGPTL4 that did not form disulfide-bonded oligomers but that did inhibit LPL (Shan et al. 2009). Beyond E40, additional ANGPTL4 residues identified as critical for LPL inhibition include others located on the same N-terminal alpha helix: H46, Q50, and Q53 (Yau et al. 2009; Lee et al. 2009).

As with ANGPTL3, the precise inhibitory mechanism of ANGPTL4 is disputed. HDX-MS has been used to determine that in the presence of substoichiometric amounts of the ANGPTL4 CCD, LPL unfolds more quickly than it does on its own (Mysling et al. 2016b). Additional tests with ANGPTL4 CCD treated with a reducing agent revealed that disulfide-linked oligomers are not essential to this unfolding mechanism but that the E40 residue is (Mysling et al. 2016b). Other investigations looking at the effect of ANGPTL4 on initial rate of catalysis on native and non-native substrates have concluded that ANGPTL4 inhibits LPL noncompetitively (Gunn et al. 2021; Gutsell et al. 2019). Additionally, a study showed that LPL regains its activity upon dissociation from ANGPTL4, suggesting that inhibition is reversible (Lafferty et al. 2013). Unlike ANGPTL3, the CCD of ANGPTL4 lacks a heparin-binding region, and binding to polyanions like heparin does not reduce ANGPTL4 inhibitory activity (Gunn et al. 2021). ANGPTL4 inhibits LPL in the capillaries, but there are reports of ANGPTL4 inhibition of LPL in both the intracellular and interstitial space (Dijk and Kersten 2016; Dijk et al. 2018; McQueen et al. 2017). Like ANGPTL3, ANGPTL4 may interact with ANGPTL8. However, the

effect of ANGPTL8 on ANGPTL4 is reported to be opposite. In 2020, a ANGPTL4/8 complex was discovered in human plasma and displayed over 100-fold *lower* inhibitory activity on LPL (Chen et al. 2020).

ANGPTL8

In 2012, ANGPTL8 was identified through a database search for ANGPTL homologs (Quagliarini et al. 2012). ANGPTL8 is an atypical member of the ANGPTL family. It has a distinct structure and does not inhibit LPL on its own but rather binds to and modulates the effects of ANGPTL3 and ANGPTL4 on LPL, as described above (Chen et al. 2020). Mice lacking ANGPTL8 have decreased plasma triglyceride levels, whereas mice with ANGPTL8 overexpression via adenovirus show elevated plasma triglyceride levels (Wang et al. 2013; Zhang 2012). Structurally, ANGPTL8 is unlike any other known ANGPTL protein in that it lacks the larger fibrinogen-like domain that the others possess (Fig. 8.5a). Complexes of both ANGPTL3/8 and ANGPTL4/8 can be found in human serum (Chen et al. 2020). A mass spectrometry (MS) analysis of ANGPTL3, 4, and 8 pulled down from human serum showed that ANGPTL8 interacts with ANGPTL3 in an approximately 1:3 ratio (Chen et al. 2020). ANGPTL3 has been shown to form trimers in solution (Gunn et al. 2021), so the MS data is consistent with the model of an ANGPTL3 trimer interacting with a single unit of ANGPTL8 (Balasubramaniam et al. 2022). In the context of the ANGPTL3/8 complex, a region encompassing ANGPTL8 residues 37-50 was protected from deuterium exchange by LPL (Balasubramaniam et al. 2022). The interaction of ANGPTL8 and ANGPTL4 was found to be at a 1:1 ratio (Chen et al. 2020). ANGPTL4/8 is a significantly worse inhibitor of LPL than ANGPTL4, suggesting that ANGPTL8 may downregulate ANGPTL4 activity (Chen et al. 2020). The structural details of the ANGPTL3/8 and ANGPTL4/8 complexes are still unclear.

LPL and Lipoprotein Binding Partners

Apolipoproteins

LPL activity is modulated in part by apolipoproteins, which are found in circulation on the surfaces of lipoproteins. From the lipoprotein surface, apolipoproteins interact with other factors (such as enzymes and receptors) and mark the lipoprotein for uptake or processing. Apolipoproteins are either exchangeable (they dynamically dissociate from and reassociate to other lipoproteins) or structural (non-exchangeable) (Hsu et al. 2023). The exchangeable apolipoproteins linked to LPL activity include APOA5, APOC1, APOC2, APOC3, and APOE (Kersten 2014),

which can be found on LPL substrates VLDL and chylomicrons, as well as on HDL. Although HDL is not primarily an LPL substrate, it serves as a circulating reservoir for apolipoproteins that can reassociate to TRLs when they become available (Miller 1990; Mahley 2016; Forte and Ryan 2015). All of the apolipoproteins have similarly amphipathic helical secondary structures that mediate their association to the lipoprotein surfaces (Segrest et al. 1992) but vary widely in sequence length, allelic diversity, and apparent affinity for and effect on LPL (Table 8.1). The actions of the apolipoproteins on LPL might be indirect, or they may be direct and transient, nonspecific, or dependent on the presence of lipids or other molecules. The precise mechanisms by which these apolipoproteins modulate triglyceride homeostasis and affect LPL activity are still under study.

Table 8.1 Apolipoproteins linked to LPL activity

Protein	LPL enhancer/suppressor	Mass and sequence length	Circulating concentration (normolipidemic)
APOA5	Enhancer	366aa (translated) 343aa (mature) ~39,000 Da (mature)	0.002–0.04 mg/dL (O’Brien et al. 2005)
APOC1	Suppressor	83aa (translated) 57aa (mature) ~6600 Da (mature)	4–12 mg/dL (Cohn et al. 2003)
APOC2	Enhancer	101aa (translated) 79aa (mature) ~8900 Da (mature)	4 mg/dL (Hortin et al. 2008)
APOC3	Suppressor	99aa (translated) 79aa (mature) ~8800 Da (mature)	~10 mg/dL in males (Cohn et al. 2004; Huff and Hegele 2013)
APOE	Suppressor (isoform-dependent) (Whitacre et al. 2022)	317aa (translated) 299aa (mature) ~34,200 Da (mature)	3–7 mg/dL (Mahley et al. 1984)

Apolipoproteins associated with LPL function. The apolipoproteins vary in size and circulating concentration. The sizes and masses of the mature proteins (after signal peptide cleavage) are provided. The table uses the terms “enhancer” and “suppressor” to describe the apolipoproteins’ effects on LPL because “activator” and “inhibitor” imply direct interaction, which remains uncertain. It should be noted that in addition to any possible direct contact interactions these proteins have with LPL, they may (or demonstrably do) have additional mechanisms by which they modulate lipid metabolism, including lipoprotein uptake receptor binding, stimulation or suppression of lipoprotein production, interaction with other apolipoproteins, and regulation of enzymes other than LPL. These apolipoproteins are included due to the possibility that they may directly interact with LPL (and form hetero-oligomers), but there are no structures solved to date of such complexes. Abbreviations: amino acids (aa), Daltons (Da), milligrams per deciliter (mg/dL)

Apolipoprotein A5

APOA5 is an exchangeable apolipoprotein of 39 kDa found on chylomicrons, VLDL, and HDL (O'Brien et al. 2005). It is present at very low concentrations in the serum, with estimates that only 1 in 1000 APOB-containing lipoproteins and 1 in 20,000 APOA1 containing lipoproteins also have APOA5 (O'Brien et al. 2005). The connection between APOA5 and lipid levels was established in 2001 when it was shown that VLDL triglyceride levels increased in APOA5 knockout mice and decreased in mice overexpressing human APOA5 (Pennacchio et al. 2001; van der Vliet et al. 2002). These results in mice are mirrored in population-level genetic studies, which revealed several variants of APOA5 associated with hypertriglyceridemia (Guardiola and Ribalta 2017). Notably, the point mutation G185C was associated with a nearly 12X greater risk of hypertriglyceridemia and a significantly greater risk of developing cardiovascular disease (Hsu et al. 2006).

Data from multiple studies shows that APOA5 affects plasma triglyceride levels by enhancing LPL activity. Additional studies suggest a role for APOA5 in enhancing remnant lipoprotein uptake (Nilsson et al. 2007). Here, we focus on APOA5 effects on LPL activity. In 2004, Fruchart-Najib and colleagues demonstrated that APOA5 directly interacts with LPL using a combination of VLDL activity assays and surface plasmon resonance studies with purified bovine LPL and recombinant APOA5 (Fruchart-Najib et al. 2004). Two studies suggested that heparan sulfate proteoglycans are needed for APOA5-mediated enhancement of LPL activity. Assays using VLDL as the substrate with HSPGs added showed that APOA5 was only able to enhance VLDL hydrolysis by LPL in the presence of HSPGs (Merkel et al. 2005). The second study showed that APOA5 bound to lipoproteins may promote their association with HSPGs, thus indirectly enhancing LPL activity (Lookene et al. 2005).

A more recent study agreed that the APOA5 mechanism of LPL activation is indirect but suggested a different mechanism. Assays were conducted on cells expressing LPL. Within this context, the authors found that APOA5 binds the ANGPTL3/8 complex with high affinity and thus counteracts the inhibitory effect of the ANGPTL3/8 complex (Chen et al. 2021). In another study that used isothermal titration calorimetry (ITC) to measure LPL activity, the authors found that APOA5 at near-physiological concentrations did not enhance LPL activity at all (Reimund et al. 2017). The lipase substrate in this case was whole plasma, which may have contained endogenous apolipoproteins or other factors that abrogated the effect. Thus, there is contrasting in vitro evidence for whether APOA5 binds directly to LPL to activate it and the precise mechanism of LPL activation is unknown.

Apolipoprotein C1

At only about 7 kDa, APOC1 is the smallest of the human apolipoproteins. The primary impact of APOC1 in lipid metabolism is through its modulation of factors other than LPL, such as cholesteryl ester transfer protein, lecithin cholesterol acyl transferase, hepatic lipase, phospholipase A₂, as well as by blocking uptake of lipoprotein remnants by hepatic receptors (Gautier et al. 2000; Sehayek and Eisenberg 1991; Fuior and Gafencu 2019). However, a combination of in vitro and in vivo evidence also points to APOC1 as a suppressor of LPL activity. As early as the 1970s, activity assays showed that high concentrations of APOC1 inhibited LPL lipolysis of intralipid, though the authors noted the effect was potentially nonspecific (Havel et al. 1973). Another group that reported hypertriglyceridemia in mice expressing human APOC1 also found that APOC1 inhibited LPL lipolysis of tritiated triolein emulsified particles in vitro (Berbee et al. 2005). More recently, a study of APOC1, C2, and C3 effects on LPL found that APOC1 and APOC3 displace LPL from substrate particles, but the mechanism by which this occurs is unknown (Larsson et al. 2013). Thus, APOC1 is linked to a suppressive effect on LPL activity and has been demonstrated to promote LPL dissociation from substrate-like particles, but its primary role in triglyceride homeostasis is likely through other mechanisms.

Apolipoprotein C2

APOC2 potentially activates LPL hydrolysis of TRLs and has therapeutic potential for the treatment of hypertriglyceridemia (Wolska et al. 2020; Sviridov et al. 2023). Mature APOC2 has a molecular weight of 8.9 kDa. NMR studies show that APOC2 secondary structure consists mostly of amphipathic alpha helices, which are critical for lipid contact and are interspersed with flexible disordered regions (Zdunek et al. 2003; MacRaid et al. 2004). These NMR structures demonstrate that the secondary structures differ depending on the lipid/lipid mimetic with which APOC2 is paired.

There is consensus that the C-terminal region of APOC2 is responsible for directly interacting with and activating LPL. Specifically, Y85, I88, D91, and Q92 were important to LPL activation, and Y85F and Q92E mutations of APOC2 resulted in near-complete loss of LPL activation (Shen et al. 2002). In this study, wild-type (dansylated) APOC2 bound to LPL with an apparent K_D of about 0.29 μ M. The authors found that the inactive APOC2 mutants could still bind LPL with affinity equal to wild type APOC2 in the absence of hydrolyzable lipids. These data agree with a 2023 study using HDX-MS that found that APOC2 residues 83–101 were protected against deuterium uptake in the presence of LPL (Kumari et al. 2023). Furthermore, peptide therapeutics based on the C-terminus of APOC2 (residues 62–101) are highly effective in activating LPL. In all the peptides described, the last 19 residues were identical to the native APOC2 sequence

corresponding to the C-terminal helix (residues 83–101). Modifications to the peptides that altered their efficacy in activating LPL were mostly located in the region that corresponds to the “random coil” region of APOC2, residues 62–80 (Wolska et al. 2020).

Multiple sites on LPL have been reported to interact with APOC2. Cross-linking studies from the early 1990s–2000s identified both C- and N-terminal regions of LPL that interact with an activating fragment of APOC2, including residues 119–140 and 416–475 (Davis et al. 1992; Dichek et al. 1993; Hill et al. 1998; McIlhargey et al. 2003). In 2023, HDX-MS pointed to several regions of LPL that are “bound and/or stabilized” by APOC2, including residues 77–87, 114–126, 212–244, and 245–263 (lid), which do not overlap with the areas of interest identified by the previous cross-linking studies. The difference could be in the buffer conditions—the earlier studies did not describe a lipid or detergent component in their protein buffers, but the later HDX study was conducted in the presence of DMPC.

There are multiple hypotheses for the mechanism by which APOC2 activates LPL. One hypothesis is that APOC2 stabilizes the LPL lid domain in an open conformation via direct contact with LPL (MacRaild et al. 2004; Razzaghi et al. 2001). The 2023 HDX-MS study reported that APOC2 binds the area near the LPL catalytic pocket, which includes the LPL lid domain. The same study concluded that as part of its mechanism of activation, APOC2 enhances LPL’s thermal stability and protects it from unfolding. For these experiments, the authors employed a construct of LPL with two mutations in its catalytic triad to prevent LPL from catabolizing the DMPC they used to stabilize APOC2 in solution (Kumari et al. 2023). Another hypothesis is that APOC2 makes the triglycerides in the core of TRLs more accessible at the lipoprotein surface (interdigitation of triglycerides with phospholipids), “activating” LPL indirectly (MacRaild et al. 2004). Biophysical examination of APOC2 on POPC/triolein/water and triolein/water interfaces has suggested that the conformation of APOC2 and the mechanism of its LPL activation are dependent on lipoprotein surface pressure (Meyers et al. 2015), a value related to interfacial tension. Interfacial tension and surface pressure of lipoproteins are potentially influenced by surface curvature/shape, adsorption of proteins to the surface, and the properties of the lipids at the interface (Meyers et al. 2015; Ollila et al. 2012), all of which are relevant to APOC2’s biological context. A complete understanding of APOC2’s mechanism will likely need to incorporate these considerations.

The extent to which APOC2 activates LPL *in vitro* depends on the amount and identity of substrate used in the assay system. APOC2 does not activate LPL appreciably on soluble substrates and activates LPL to variable degrees on different emulsified fats, TRL mimetics, and native TRL substrates. Taken together, though there may be some remaining questions on the residues that are critical for LPL activation, there is a consensus that APOC2 structure and activation of LPL are sensitive to the presence and characteristics of other hydrophobic and amphipathic molecules in solution. Several studies converge on the region of APOC2 known as alpha helix III (residues 83–101 or 61–79) as the region that interacts with LPL. All studies seem to agree that the C-terminus of APOC2 is necessary for LPL activation.

Apolipoprotein C3

APOC3 has 79 amino acids and a mature molecular weight of 8.8 KDa. Transgenic expression of APOC3 in mice causes hypertriglyceridemia (Ito et al. 1990), whereas disruption of the APOC3 gene in mice results in hypotriglyceridemia (Maeda et al. 1994). Based on reports that individuals lacking APOC3 have decreased circulating triglyceride levels and protection from heart disease (Pollin et al. 2008), an anti-sense oligonucleotide (ASO) therapeutic called Volanesorsen (Waylivra) was developed and approved in the EU for the treatment of familial chylomicronemia syndrome (FCS) (Kolovou et al. 2022).

APOC3 has been proposed to impact lipid metabolism through both LPL-dependent and LPL-independent (specifically, inhibiting hepatic uptake of TRLs) pathways. Supporting an LPL-dependent pathway, humans heterozygous for a loss of function APOC3 mutation have higher rates of conversion of VLDL to LDL relative to individuals with two functional copies of APOC3 (Reyes-Soffer et al. 2019). Additionally, in humans with homozygous loss of APOC3 and APOA1, catabolism of VLDL was much faster than in controls with functional APOC3 and APOA1 (Ginsberg et al. 1986). Conversely, Volanesorsen is effective in patients with LPL deficiency, a result which implies that loss of APOC3 enhances removal of TRLs through uptake by the liver (Gaudet et al. 2014). Additional studies support an LPL-independent pathway in mice, in which ASOs against APOC3 were used in mice with targeted deletion of LPL, LDL receptor (LDLR), LDL receptor-related protein (LRP), or HSPG receptors. They found that APOC3 ASOs did not lower plasma triglycerides in mice lacking both LDLr and LRP but could lower plasma triglycerides in mice lacking LPL (Gordts et al. 2016). Thus, there is data in humans and mice to support both LPL-dependent and independent pathways for APOC3 effects on plasma TGs. We further explore *in vitro* data for LPL-dependent pathways here.

In 1972, Brown and Baginsky showed that high concentrations of purified APOC3 could inhibit LPL hydrolysis of triolein (Brown and Baginsky 1972). These results were corroborated in a study published the following year that showed that APOC3 could inhibit LPL when low concentrations of triolein were used (Krauss et al. 1973). In 1992, another *in vitro* study compared the effects of plasma-derived APOC3, synthesized APOC3, and APOC3 peptides on LPL. APOC3 is sialylated, and the authors found that the sialic acid moiety was nonessential for LPL activation based on plasma-derived sialylated APOC3-1 having a similar IC₅₀ value (low micromolar range) compared to synthesized, unsialylated APOC3 (McConathy et al. 1992). Lastly, because their APOC3 fragments *could* inhibit LPL, and full-length APOC3 inhibited LPL more potently, they suggested that there are multiple attachment sites on APOC3 that bind to LPL (McConathy et al. 1992). A later study investigated the mechanism of APOC3 and APOA1 inhibition of LPL (Larsson et al. 2013). The main findings were that APOC3 and APOC1 displace LPL from the substrate particles used in their assay (intralipid and rat lymph chylomicrons). Based on the observation that APOC3 inhibited LPL less in the presence of more substrate, the authors suggested that APOC3 competes with LPL for lipoprotein

binding. These data show that APOC3 does affect LPL hydrolysis of emulsified lipids *in vitro*, but it is still unclear whether this is due to complex formation/direct interaction or a more transient competition between the two proteins for binding to substrate particles. Mechanistic and structural detail is still missing from our knowledge of the interaction between APOC3 and LPL.

Apolipoprotein E

The primary role of APOE in triglyceride homeostasis is its LDL receptor-binding activity, which facilitates uptake of triglyceride-rich lipoproteins and promotes their clearance from circulation (Lalazar and Mahley 1989). After signal peptide processing, APOE has 299 amino acids and a mature molecular weight of 34 kDa. APOE is found in chylomicrons, VLDL, IDL, and HDL and is polymorphic in humans. The main isoforms of human APOE (APOE2, APOE3, and APOE4) have been extensively studied for their roles in cardiovascular disease and Alzheimer's disease. The most common of these alleles is APOE3 (Corbo and Scacchi 1999). APOE has two structural domains. The N-terminal, receptor-binding domain was solved by X-ray crystallography and is comprised of an antiparallel four-helix bundle (Wilson et al. 1991). It is connected by a linker to the C-terminal domain, which contains a lipid binding region (Aggerbeck et al. 1988).

The effect of APOE on LPL activity is contentious, with different groups coming to opposite conclusions. In 1980, one group reported that APOE activates LPL using an assay that contained human plasma and used detergent-emulsified triolein as the LPL substrate. Supplementing their activity assay mixture with an anti-ApoE antibody diminished LPL activity, and adding extra APOE (5–50 µg/mL, within the physiological range) to the assay increased LPL activity (Yamada and Murase 1980). In contrast, a study published in 1996 reported that APOE (up to 60 µg/mL, within the physiological range) inhibits LPL activity on triolein (emulsified with cholesteryl oleate to mimic a chylomicron) and that chemical modification of the APOE arginine residues with 1,2-cyclohexadione completely restored LPL activity (Rensen and van Berkel 1996). A study from 2022 revisited the question of ApoE-mediated LPL inhibition with an assay designed to mimic the native plasma environment (Whitacre et al. 2022). The assay used VLDL isolated from humans incubated with purified protein, followed by separation of the reaction products by size exclusion chromatography. The proteins detected on the VLDL particles (identified by MS and ELISA) were correlated with the degree of VLDL processing. The authors found that VLDL processing by LPL was inversely correlated with ApoE concentration and that this effect was isoform-dependent, with APOE4 inhibiting LPL the most and APOE2 inhibiting LPL the least (Whitacre et al. 2022). Overall, APOE is important for the receptor-mediated uptake of the triglyceride-rich lipoproteins that LPL processes. There is more evidence in the literature pointing to APOE as a suppressor of LPL than as an enhancer.

LPL and Receptors

LDLR, LRP, and VLDLR

LPL not only reduces lipoprotein size and density but also mediates the binding and uptake of lipoproteins and lipoprotein remnants by lipoprotein receptors in the liver. Lipoprotein particles are taken up by members of the structurally related LDL receptor (LDLR) family. Three receptors contribute to lipoprotein remnant uptake: LDLR, LDL receptor-related protein (LRP), and VLDL receptor (VLDLR). LDLr family members share four major structural modules: cysteine-rich complement-type repeats, EGFR-like repeats, a transmembrane domain, and a cytoplasmic domain. In addition to LPL, APOE is also an important signal for the uptake of remnant lipoproteins (Zhang et al. 1992). APOE works in combination with LPL as a recognition signal for lipoprotein remnant uptake into cells.

A role for LPL in remnant uptake was first postulated in 1975, when it was shown that LPL was associated with lipoprotein remnants in circulation and that LPL and the remnants were taken up by the liver at the same rate (Felts et al. 1975). These data led the authors to suggest that LPL served as a marker for remnant uptake by the liver. Later studies confirmed that in both pre- and post-heparin plasma, LPL was associated with VLDL, provided that LPL activity was inhibited so that the VLDL was not converted to LDL (Zambon et al. 1996). Additional studies showed that LPL enhances the binding of chylomicron remnants and VLDL to liver cells and that LPL and APOE could be cross-linked to LRP (Beisiegel et al. 1991; Chappell et al. 1992). Furthermore, LPL could bind to both VLDL particles and purified LRP, and this binding depended on LPL's C-terminal domain (Nykjaer et al. 1993; Nielsen et al. 1997).

Studies also addressed the role of LPL in enhancing uptake of other lipoproteins by the VLDLR and LDLR. The VLDLR mediates uptake of chylomicron remnants into cells, and LPL enhances this uptake (Niemeier et al. 1996). In vitro studies showed that purified LPL bound with high affinity to the VLDLR and that the C-terminal domain of LPL mediated this interaction (Argaves et al. 1995). Addition of LPL doubled LDL uptake by cells, a phenomenon that was not seen when cells lacked the LDLR (Loeffler et al. 2007).

Summary

LPL is a highly efficient lipase that is regulated by an extensive collection of partners. Before secretion, LPL relies on LMF1 and SEL1L to be folded and appropriately exported from the ER. During transit through the cell in SDC1-containing vesicles, LPL may be stabilized as a helix through its interactions with heparin. After secretion, LPL is diluted and encounters other HSPGs and albumin in the crowded interstitial space. LPL must interact with GPIHBP1, located on the exterior

of the endothelial cells, to reach its final destination and site of activity, the capillary lumen. When in the capillary, LPL is still in the presence of albumin and many other interacting proteins and small molecules. ANGPTL3, 4, and 8 are secreted by various tissues to suppress LPL activity where and when fatty acids are not needed. When LPL encounters its substrate (triglyceride-rich lipoproteins), its interactions with apolipoproteins located on the lipoprotein surface offer yet another opportunity for regulation. Apolipoprotein C2 is needed for full LPL efficacy. If LPL clears all these checkpoints, it rapidly catabolizes the triglycerides in the lipoprotein core, increasing the density of the lipoprotein particle and thus changing its properties, and simultaneously providing fatty acids to proximal tissue. This enzymatic reaction is so critical to our biology that loss-of-function mutations in LPL or its critical partners (such as GPIHBP1) can result in plasma that is so oversaturated with lipids that it appears opaque and white—"lactescent" or lipemic plasma, in clinical terms—which is a threat to the pancreas. Even with less severe deficits, chronic excess of triglycerides in circulation is a risk factor for cardiovascular disease, which continues to be the leading cause of death globally. Since its discovery 70 years ago, the study of this enzyme has and will continue to yield insight into enzyme regulation, lipid metabolism, and cardiovascular health.

References

- Aggerbeck LP, Wetterau JR, Weisgraber KH, Wu CS, Lindgren FT (1988) Human apolipoprotein E3 in aqueous solution. II. Properties of the amino- and carboxyl-terminal domains. *J Biol Chem* 263(13):6249–6258
- Allan CM, Larsson M, Jung RS, Ploug M, Bensadoun A, Beigneux AP, Fong LG, Young SG (2017) Mobility of "HSPG-bound" LPL explains how LPL is able to reach GPIHBP1 on capillaries. *J Lipid Res* 58(1):216–225. <https://doi.org/10.1194/jlr.M072520>
- Anderson NG, Fawcett B (1950) An antichylomicronemic substance produced by heparin injection. *Proc Soc Exp Biol Med* 74(4):768–771. <https://doi.org/10.3181/00379727-74-18042>
- Ando Y, Shimizugawa T, Takeshita S, Ono M, Shimamura M, Koishi R, Furukawa H (2003) A decreased expression of angiopoietin-like 3 is protective against atherosclerosis in apoE-deficient mice. *J Lipid Res* 44(6):1216–1223. <https://doi.org/10.1194/jlr.M300031-JLR200>
- Anfinsen CB, Boyle E, Brown RK (1952) The role of heparin in lipoprotein metabolism. *Science* 115(2996):583–586. <https://doi.org/10.1126/science.115.2996.583>
- Argraives KM, Battey FD, MacCalman CD, McCrae KR, Gafvels M, Kozarsky KF, Chappell DA, Strauss JF 3rd, Strickland DK (1995) The very low density lipoprotein receptor mediates the cellular catabolism of lipoprotein lipase and urokinase-plasminogen activator inhibitor type I complexes. *J Biol Chem* 270(44):26550–26557. <https://doi.org/10.1074/jbc.270.44.26550>
- Arora R, Nimonkar AV, Baird D, Wang C, Chiu CH, Horton PA, Hanrahan S, Cubbon R, Weldon S, Tschantz WR, Mueller S, Brunner R, Lehr P, Meier P, Ottl J, Voznesensky A, Pandey P, Smith TM, Stojanovic A, Flyer A, Benson TE, Romanowski MJ, Trauger JW (2019) Structure of lipoprotein lipase in complex with GPIHBP1. *Proc Natl Acad Sci U S A* 116(21):10360–10365. <https://doi.org/10.1073/pnas.1820171116>
- Babilonia-Rosa M, Neher SB (2014) Purification, cellular levels, and functional domains of LMF1. *Biochem Biophys Res Commun*. <https://doi.org/10.1016/j.bbrc.2014.05.136>
- Balasubramaniam D, Schroeder O, Russell AM, Fitchett JR, Austin AK, Beyer TP, Chen YQ, Day JW, Ehsani M, Heng AR, Zhen EY, Davies J, Glaesner W, Jones BE, Siegel RW, Qian

- YW, Konrad RJ (2022) An anti-ANGPTL3/8 antibody decreases circulating triglycerides by binding to a LPL-inhibitory leucine zipper-like motif. *J Lipid Res* 63(5):100198. <https://doi.org/10.1016/j.jlr.2022.100198>
- Bateman A, Sandford R (1999) The PLAT domain: a new piece in the PKD1 puzzle. *Curr Biol* 9(16):R588–R590. [https://doi.org/10.1016/s0960-9822\(99\)80380-7](https://doi.org/10.1016/s0960-9822(99)80380-7)
- Beigneux AP, Davies BS, Gin P, Weinstein MM, Farber E, Qiao X, Peale F, Bunting S, Walzem RL, Wong JS, Blazer WS, Ding ZM, Melford K, Wongsiriroj N, Shu X, de Sauvage F, Ryan RO, Fong LG, Bensadoun A, Young SG (2007) Glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 plays a critical role in the lipolytic processing of chylomicrons. *Cell Metab* 5(4):279–291. <https://doi.org/10.1016/j.cmet.2007.02.002>
- Beigneux AP, Allan CM, Sandoval NP, Cho GW, Heizer PJ, Jung RS, Stanhope KL, Havel PJ, Birrane G, Meiyappan M, Gill JE, Murakami M, Miyashita K, Nakajima K, Ploug M, Fong LG, Young SG (2019) Lipoprotein lipase is active as a monomer. *Proc Natl Acad Sci U S A* 116(13):6319–6328. <https://doi.org/10.1073/pnas.1900983116>
- Beisiegel U, Weber W, Bengtsson-Olivecrona G (1991) Lipoprotein lipase enhances the binding of chylomicrons to low density lipoprotein receptor-related protein. *Proc Natl Acad Sci U S A* 88(19):8342–8346. <https://doi.org/10.1073/pnas.88.19.8342>
- Ben-Avram CM, Ben-Zeev O, Lee TD, Haaga K, Shively JE, Goers J, Pedersen ME, Reeve JR Jr, Schotz MC (1986) Homology of lipoprotein lipase to pancreatic lipase. *Proc Natl Acad Sci U S A* 83(12):4185–4189. <https://doi.org/10.1073/pnas.83.12.4185>
- Bengtsson G, Olivecrona T (1979) Binding of deoxycholate to lipoprotein lipase. *Biochim Biophys Acta* 575(3):471–474
- Bengtsson-Olivecrona G, Olivecrona T, Jornvall H (1986) Lipoprotein lipases from cow, guinea-pig and man. Structural characterization and identification of protease-sensitive internal regions. *Eur J Biochem* 161(2):281–288. <https://doi.org/10.1111/j.1432-1033.1986.tb10444.x>
- Ben-Zeev O, Stahnke G, Liu G, Davis RC, Doolittle MH (1994) Lipoprotein lipase and hepatic lipase: the role of asparagine-linked glycosylation in the expression of a functional enzyme. *J Lipid Res* 35(9):1511–1523
- Berbee JF, van der Hoogt CC, Sundararaman D, Havekes LM, Rensen PC (2005) Severe hypertriglyceridemia in human APOC1 transgenic mice is caused by apoC-I-induced inhibition of LPL. *J Lipid Res* 46(2):297–306. <https://doi.org/10.1194/jlr.M400301-JLR200>
- Bergmark BA, Marston NA, Bramson CR, Curto M, Ramos V, Jevne A, Kuder JF, Park JG, Murphy SA, Verma S, Wojakowski W, Terra SG, Sabatine MS, Wiviott SD, Investigators T-T (2022) Effect of vupanors on non-high-density lipoprotein cholesterol levels in statin-treated patients with elevated cholesterol: TRANSLATE-TIMI 70. *Circulation* 145(18):1377–1386. <https://doi.org/10.1161/CIRCULATIONAHA.122.059266>
- Bini S, Pecce V, Di Costanzo A, Polito L, Ghadiri A, Minicocci I, Tambaro F, Covino S, Arca M, D'Erasmo L (2022) The fibrinogen-like domain of ANGPTL3 facilitates lipolysis in 3T3-L1 Cells by activating the intracellular Erk pathway. *Biomolecules* 12(4). <https://doi.org/10.3390/biom12040585>
- Birrane G, Beigneux AP, Dwyer B, Strack-Logue B, Kristensen KK, Franccone OL, Fong LG, Mertens HDT, Pan CQ, Ploug M, Young SG, Meiyappan M (2019) Structure of the lipoprotein lipase-GPIHBP1 complex that mediates plasma triglyceride hydrolysis. *Proc Natl Acad Sci U S A* 116(5):1723–1732. <https://doi.org/10.1073/pnas.1817984116>
- Biterova E, Esmaeili M, Alanen HI, Saarinen M, Ruddock LW (2018) Structures of Angptl3 and Angptl4, modulators of triglyceride levels and coronary artery disease. *Sci Rep* 8(1):6752. <https://doi.org/10.1038/s41598-018-25237-7>
- Biunno I, Appierto V, Cattaneo M, Leone BE, Balzano G, Socci C, Saccone S, Letizia A, Della Valle G, Sgarbetta V (1997) Isolation of a pancreas-specific gene located on human chromosome 14q31: expression analysis in human pancreatic ductal carcinomas. *Genomics* 46(2):284–286. <https://doi.org/10.1006/geno.1997.5018>

- Boyle E, Bragdon JH, Brown RK (1952) Role of heparin in in vitro production of alpha1 lipoproteins in human plasma. *Proc Soc Exp Biol Med* 81(2):475–477. <https://doi.org/10.3181/00379727-81-19915>
- Brown WV, Baginsky ML (1972) Inhibition of lipoprotein lipase by an apoprotein of human very low density lipoprotein. *Biochem Biophys Res Commun* 46(2):375–382
- Brown RK, Boyle E, Anfinsen CB (1953) The enzymatic transformation of lipoproteins. *J Biol Chem* 204(1):423–434
- Camenisch G, Pisabarro MT, Sherman D, Kowalski J, Nagel M, Hass P, Xie MH, Gurney A, Bodary S, Liang XH, Clark K, Beresini M, Ferrara N, Gerber HP (2002) ANGPTL3 stimulates endothelial cell adhesion and migration via integrin alpha v beta 3 and induces blood vessel formation in vivo. *J Biol Chem* 277(19):17281–17290. <https://doi.org/10.1074/jbc.M109768200>
- Chappell DA, Fry GL, Waknitz MA, Iverius PH, Williams SE, Strickland DK (1992) The low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor binds and mediates catabolism of bovine milk lipoprotein lipase. *J Biol Chem* 267(36):25764–25767
- Chebotaeva NA, Harding SE, Winzor DJ (2001) Ultracentrifugal studies of the effect of molecular crowding by trimethylamine N-oxide on the self-association of muscle glycogen phosphorylase b. *Eur J Biochem* 268(3):506–513. <https://doi.org/10.1046/j.1432-1327.2001.01838.x>
- Chen YQ, Pottanat TG, Siegel RW, Ehsani M, Qian YW, Zhen EY, Regmi A, Roell WC, Guo H, Luo MJ, Gimeno RE, Van't Hooft F, Konrad RJ (2020) Angiotensin-like protein 8 differentially regulates ANGPTL3 and ANGPTL4 during postprandial partitioning of fatty acids. *J Lipid Res* 61(8):1203–1220. <https://doi.org/10.1194/jlr.RA120000781>
- Chen YQ, Pottanat TG, Zhen EY, Siegel RW, Ehsani M, Qian YW, Konrad RJ (2021) ApoA5 lowers triglyceride levels via suppression of ANGPTL3/8-mediated LPL inhibition. *J Lipid Res* 62:100068. <https://doi.org/10.1016/j.jlr.2021.100068>
- Cheng CF, Oosta GM, Bensadoun A, Rosenberg RD (1981) Binding of lipoprotein lipase to endothelial cells in culture. *J Biol Chem* 256(24):12893–12898
- Chi X, Britt EC, Shows HW, Hjelmaas AJ, Shetty SK, Cushing EM, Li W, Dou A, Zhang R, Davies BSJ (2017) ANGPTL8 promotes the ability of ANGPTL3 to bind and inhibit lipoprotein lipase. *Mol Metab* 6(10):1137–1149. <https://doi.org/10.1016/j.molmet.2017.06.014>
- Cohn JS, Tremblay M, Boulet L, Jacques H, Davignon J, Roy M, Bernier L (2003) Plasma concentration and lipoprotein distribution of ApoC-I is dependent on ApoE genotype rather than the Hpa I ApoC-I promoter polymorphism. *Atherosclerosis* 169(1):63–70. [https://doi.org/10.1016/s0021-9150\(03\)00139-4](https://doi.org/10.1016/s0021-9150(03)00139-4)
- Cohn JS, Patterson BW, Uffelman KD, Davignon J, Steiner G (2004) Rate of production of plasma and very-low-density lipoprotein (VLDL) apolipoprotein C-III is strongly related to the concentration and level of production of VLDL triglyceride in male subjects with different body weights and levels of insulin sensitivity. *J Clin Endocrinol Metab* 89(8):3949–3955. <https://doi.org/10.1210/jc.2003-032056>
- Conklin D, Gilbertson D, Taft DW, Maurer MF, Whitmore TE, Smith DL, Walker KM, Chen LH, Wattler S, Nehls M, Lewis KB (1999) Identification of a mammalian angiotensin-related protein expressed specifically in liver. *Genomics* 62(3):477–482. <https://doi.org/10.1006/geno.1999.6041>
- Corbo RM, Scacchi R (1999) Apolipoprotein E (APOE) allele distribution in the world. Is APOE*4 a 'thrifty' allele? *Ann Hum Genet* 63 (Pt 4):301–310. <https://doi.org/10.1046/j.1469-1809.1999.6340301.x>
- Davies BS, Beigneux AP, Barnes RH 2nd, Tu Y, Gin P, Weinstein MM, Nobumori C, Nyren R, Goldberg I, Olivecrona G, Bensadoun A, Young SG, Fong LG (2010) GPIHBP1 is responsible for the entry of lipoprotein lipase into capillaries. *Cell Metab* 12(1):42–52. <https://doi.org/10.1016/j.cmet.2010.04.016>
- Davis RC, Wong H, Nikazy J, Wang K, Schotz MC (1992) Chimeras of hepatic lipase and lipoprotein lipase. Domain localization of enzyme-specific properties. *J Biol Chem* 267(30):21499–21504

- Deng M, Kutrolli E, Sadewasser A, Michel S, Joibari MM, Jaschinski F, Olivecrona G, Nilsson SK, Kersten S (2022) ANGPTL4 silencing via antisense oligonucleotides reduces plasma triglycerides and glucose in mice without causing lymphadenopathy. *J Lipid Res* 63(7):100237. <https://doi.org/10.1016/j.jlr.2022.100237>
- Desai U, Lee EC, Chung K, Gao C, Gay J, Key B, Hansen G, Machajewski D, Platt KA, Sands AT, Schneider M, Van Sligtenhorst I, Suwanichkul A, Vogel P, Wilganowski N, Wingert J, Zambrowicz BP, Landes G, Powell DR (2007) Lipid-lowering effects of anti-angiopoietin-like 4 antibody recapitulate the lipid phenotype found in angiopoietin-like 4 knockout mice. *Proc Natl Acad Sci U S A* 104(28):11766–11771. <https://doi.org/10.1073/pnas.0705041104>
- Dewey FE, Gusarova V, O'Dushlaine C, Gottesman O, Trejos J, Hunt C, Van Hout CV, Habegger L, Buckler D, Lai KM, Leader JB, Murray MF, Ritchie MD, Kirchner HL, Ledbetter DH, Penn J, Lopez A, Borecki IB, Overton JD, Reid JG, Carey DJ, Murphy AJ, Yancopoulos GD, Baras A, Gromada J, Shuldiner AR (2016) Inactivating variants in ANGPTL4 and risk of coronary artery disease. *N Engl J Med* 374(12):1123–1133. <https://doi.org/10.1056/NEJMoa1510926>
- Dewey FE, Gusarova V, Dunbar RL, O'Dushlaine C, Schurmann C, Gottesman O, McCarthy S, Van Hout CV, Bruse S, Dansky HM, Leader JB, Murray MF, Ritchie MD, Kirchner HL, Habegger L, Lopez A, Penn J, Zhao A, Shao W, Stahl N, Murphy AJ, Hamon S, Bouzelmat A, Zhang R, Shumel B, Pordy R, Gipe D, Herman GA, Sheu WHH, Lee IT, Liang KW, Guo X, Rotter JJ, Chen YI, Kraus WE, Shah SH, Damrauer S, Small A, Rader DJ, Wulff AB, Nordestgaard BG, Tybjaerg-Hansen A, van den Hoek AM, Princen HMG, Ledbetter DH, Carey DJ, Overton JD, Reid JG, Sasiela WJ, Banerjee P, Shuldiner AR, Borecki IB, Teslovich TM, Yancopoulos GD, Mellis SJ, Gromada J, Baras A (2017) Genetic and pharmacologic inactivation of ANGPTL3 and cardiovascular disease. *N Engl J Med* 377(3):211–221. <https://doi.org/10.1056/NEJMoa1612790>
- Dichek HL, Parrott C, Ronan R, Brunzell JD, Brewer HB Jr, Santamarina-Fojo S (1993) Functional characterization of a chimeric lipase genetically engineered from human lipoprotein lipase and human hepatic lipase. *J Lipid Res* 34(8):1393–1340
- Dijk W, Kersten S (2016) Regulation of lipid metabolism by angiopoietin-like proteins. *Curr Opin Lipidol* 27(3):249–256. <https://doi.org/10.1097/MOL.0000000000000290>
- Dijk W, Ruppert PMM, Oost LJ, Kersten S (2018) Angiopoietin-like 4 promotes the intracellular cleavage of lipoprotein lipase by PCSK3/furin in adipocytes. *J Biol Chem*. <https://doi.org/10.1074/jbc.RA118.002426>
- Dugi KA, Dichek HL, Santamarina-Fojo S (1995) Human hepatic and lipoprotein lipase: the loop covering the catalytic site mediates lipase substrate specificity. *J Biol Chem* 270(43):25396–25401. <https://doi.org/10.1074/jbc.270.43.25396>
- Emmerich J, Beg OU, Peterson J, Previato L, Brunzell JD, Brewer HB Jr, Santamarina-Fojo S (1992) Human lipoprotein lipase. Analysis of the catalytic triad by site-directed mutagenesis of Ser-132, Asp-156, and His-241. *J Biol Chem* 267(6):4161–4165
- Enerback S, Semb H, Bengtsson-Olivecrona G, Carlsson P, Hermansson ML, Olivecrona T, Bjursell G (1987) Molecular cloning and sequence analysis of cDNA encoding lipoprotein lipase of guinea pig. *Gene* 58(1):1–12. [https://doi.org/10.1016/0378-1119\(87\)90023-0](https://doi.org/10.1016/0378-1119(87)90023-0)
- Felts JM, Itakura H, Crane RT (1975) The mechanism of assimilation of constituents of chylomicrons, very low density lipoproteins and remnants - a new theory. *Biochem Biophys Res Commun* 66(4):1467–1475. [https://doi.org/10.1016/0006-291x\(75\)90524-0](https://doi.org/10.1016/0006-291x(75)90524-0)
- Fiamoncini J, Yiorkas AM, Gedrich K, Rundel M, Alsters SI, Roeselers G, van den Broek TJ, Clavel T, Lagkouvardos I, Wopereis S, Frost G, van Ommen B, Blakemore AI, Daniel H (2017) Determinants of postprandial plasma bile acid kinetics in human volunteers. *Am J Physiol Gastrointest Liver Physiol* 313(4):G300–G312. <https://doi.org/10.1152/ajpgi.00157.2017>
- Fielding CJ (1969) Purification of lipoprotein lipase from rat post-heparin plasma. *Biochim Biophys Acta* 178(3):499–507. [https://doi.org/10.1016/0005-2744\(69\)90219-8](https://doi.org/10.1016/0005-2744(69)90219-8)
- Fitzgerald ML, Wang Z, Park PW, Murphy G, Bernfield M (2000) Shedding of syndecan-1 and -4 ectodomains is regulated by multiple signaling pathways and mediated by a TIMP-3-sensitive metalloproteinase. *J Cell Biol* 148(4):811–824. <https://doi.org/10.1083/jcb.148.4.811>

- Forte TM, Ryan RO (2015) Apolipoprotein A5: extracellular and intracellular roles in triglyceride metabolism. *Curr Drug Targets* 16(12):1274–1280. <https://doi.org/10.2174/1389450116666150531161138>
- Francisco AB, Singh R, Li S, Vani AK, Yang L, Munroe RJ, Diaferia G, Cardano M, Biunno I, Qi L, Schimenti JC, Long Q (2010) Deficiency of suppressor enhancer Lin12 1 like (SEL1L) in mice leads to systemic endoplasmic reticulum stress and embryonic lethality. *J Biol Chem* 285(18):13694–13703. <https://doi.org/10.1074/jbc.M109.085340>
- Frommherz L, Bub A, Hummel E, Rist MJ, Roth A, Watzl B, Kulling SE (2016) Age-related changes of plasma bile acid concentrations in healthy adults—results from the cross-sectional KarMeN study. *PLoS One* 11(4):e0153959. <https://doi.org/10.1371/journal.pone.0153959>
- Fruchart-Najib J, Bauge E, Niculescu LS, Pham T, Thomas B, Rommens C, Majd Z, Brewer B, Pennacchio LA, Fruchart JC (2004) Mechanism of triglyceride lowering in mice expressing human apolipoprotein A5. *Biochem Biophys Res Commun* 319(2):397–404. <https://doi.org/10.1016/j.bbrc.2004.05.003>
- Fuier EV, Gafencu AV (2019) Apolipoprotein C1: its pleiotropic effects in lipid metabolism and beyond. *Int J Mol Sci* 20(23). <https://doi.org/10.3390/ijms20235939>
- Garfinkel AS, Kempner ES, Ben-Zeev O, Nikazy J, James SJ, Schotz MC (1983) Lipoprotein lipase: size of the functional unit determined by radiation inactivation. *J Lipid Res* 24(6):775–780
- Gaudet D, Brisson D, Tremblay K, Alexander VJ, Singleton W, Hughes SG, Geary RS, Baker BF, Graham MJ, Crooke RM, Witztum JL (2014) Targeting APOC3 in the familial chylomicronemia syndrome. *N Engl J Med* 371(23):2200–2206. <https://doi.org/10.1056/NEJMoa1400284>
- Gautier T, Masson D, de Barros JP, Athias A, Gamber P, Aunis D, Metz-Boutigue MH, Lagrost L (2000) Human apolipoprotein C-I accounts for the ability of plasma high density lipoproteins to inhibit the cholesteryl ester transfer protein activity. *J Biol Chem* 275(48):37504–37509. <https://doi.org/10.1074/jbc.M007210200>
- Ge H, Yang G, Yu X, Pourbahrami T, Li C (2004) Oligomerization state-dependent hyperlipidemic effect of angiopoietin-like protein 4. *J Lipid Res* 45(11):2071–2079. <https://doi.org/10.1194/jlr.M400138-JLR200>
- Ge H, Cha JY, Gopal H, Harp C, Yu X, Repa JJ, Li C (2005) Differential regulation and properties of angiopoietin-like proteins 3 and 4. *J Lipid Res* 46(7):1484–1490. <https://doi.org/10.1194/jlr.M500005-JLR200>
- Ginsberg HN, Le NA, Goldberg IJ, Gibson JC, Rubinstein A, Wang-Iverson P, Norum R, Brown WV (1986) Apolipoprotein B metabolism in subjects with deficiency of apolipoproteins CIII and AI. Evidence that apolipoprotein CIII inhibits catabolism of triglyceride-rich lipoproteins by lipoprotein lipase in vivo. *J Clin Invest* 78(5):1287–1295. <https://doi.org/10.1172/JCI112713>
- Gordts PL, Nock R, Son NH, Ramms B, Lew I, Gonzales JC, Thacker BE, Basu D, Lee RG, Mullick AE, Graham MJ, Goldberg IJ, Crooke RM, Witztum JL, Esko JD (2016) ApoC-III inhibits clearance of triglyceride-rich lipoproteins through LDL family receptors. *J Clin Invest* 126(8):2855–2866. <https://doi.org/10.1172/JCI86610>
- Graham DM, Lyon TP, Gofman JW, Jones HB, Yankley A, Simonton J, White S (1951) Blood lipids and human atherosclerosis. II. The influence of heparin upon lipoprotein metabolism. *Circulation* 4(5):666–673. <https://doi.org/10.1161/01.cir.4.5.666>
- Grochulski P, Bouthillier F, Kazlauskas RJ, Serreqi AN, Schrag JD, Ziomek E, Cygler M (1994) Analogs of reaction intermediates identify a unique substrate binding site in *Candida rugosa* lipase. *Biochemistry* 33(12):3494–3500. <https://doi.org/10.1021/bi00178a005>
- Guardiola M, Ribalta J (2017) Update on APOA5 genetics: toward a better understanding of its physiological impact. *Curr Atheroscler Rep* 19(7):30. <https://doi.org/10.1007/s11883-017-0665-y>
- Gunn KH, Neher SB (2023) Structure of dimeric lipoprotein lipase reveals a pore adjacent to the active site. *Nat Commun* 14(1):2569. <https://doi.org/10.1038/s41467-023-38243-9>
- Gunn KH, Roberts BS, Wang F, Strauss JD, Borgnia MJ, Egelman EH, Neher SB (2020) The structure of helical lipoprotein lipase reveals an unexpected twist in lipase storage. *Proc Natl Acad Sci U S A* 117(19):10254–10264. <https://doi.org/10.1073/pnas.1916555117>

- Gunn KH, Gutsell A, Xu Y, Johnson CV, Liu J, Neher SB (2021) Mechanistic Insight into ANGPTL3 Inhibition of LPL. *J Biol Chem*
- Gutsell AR, Ghodse SV, Bowers AA, Neher SB (2019) Mapping the sites of the lipoprotein lipase (LPL)-angiopoietin-like protein 4 (ANGPTL4) interaction provides mechanistic insight into LPL inhibition. *J Biol Chem* 294(8):2678–2689. <https://doi.org/10.1074/jbc.RA118.005932>
- Hahn PF (1943) Abolishment of alimentary lipemia following injection of heparin. *Science* 98(2531):19–20. <https://doi.org/10.1126/science.98.2531.19>
- Haller JF, Mintah IJ, Shihanian LM, Stevis P, Buckler D, Alexa-Braun CA, Kleiner S, Banfi S, Cohen JC, Hobbs HH, Yancopoulos GD, Murphy AJ, Gusarova V, Gromada J (2017) ANGPTL8 requires ANGPTL3 to inhibit lipoprotein lipase and plasma triglyceride clearance. *J Lipid Res* 58. <https://doi.org/10.1194/JLR.M075689>
- Hata A, Ridinger DN, Sutherland S, Emi M, Shuhua Z, Myers RL, Ren K, Cheng T, Inoue I, Wilson DE et al (1993) Binding of lipoprotein lipase to heparin. Identification of five critical residues in two distinct segments of the amino-terminal domain. *J Biol Chem* 268(12):8447–8457
- Havel RJ, Fielding CJ, Olivecrona T, Shore VG, Fielding PE, Egelrud T (1973) Cofactor activity of protein components of human very low density lipoproteins in the hydrolysis of triglycerides by lipoproteins lipase from different sources. *Biochemistry* 12(9):1828–1833. <https://doi.org/10.1021/bi00733a026>
- Hayne CK, Yumerefendi H, Cao L, Gauer JW, Lafferty MJ, Kuhlman B, Erie DA, Neher SB (2018) We FRET so you don't have to: new models of the lipoprotein lipase dimer. *Biochemistry* 57(2):241–254. <https://doi.org/10.1021/acs.biochem.7b01009>
- Hill JS, Yang D, Nikazy J, Curtiss LK, Sparrow JT, Wong H (1998) Subdomain chimeras of hepatic lipase and lipoprotein lipase. Localization of heparin and cofactor binding. *J Biol Chem* 273(47):30979–30984. <https://doi.org/10.1074/jbc.273.47.30979>
- Hollett C, Meng HC (1956) Purification and characterization of the lipemia-clearing factor of postheparin plasma. *Biochim Biophys Acta* 20(2):421–422. [https://doi.org/10.1016/0006-3002\(56\)90322-5](https://doi.org/10.1016/0006-3002(56)90322-5)
- Hortin GL, Sviridov D, Anderson NL (2008) High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. *Clin Chem* 54(10):1608–1616. <https://doi.org/10.1373/clinchem.2008.108175>
- Hsu LA, Ko YL, Chang CJ, Hu CF, Wu S, Teng MS, Wang CL, Ho WJ, Ko YS, Hsu TS, Lee YS (2006) Genetic variations of apolipoprotein A5 gene is associated with the risk of coronary artery disease among Chinese in Taiwan. *Atherosclerosis* 185(1):143–149. <https://doi.org/10.1016/j.atherosclerosis.2005.05.031>
- Hsu CC, Kanter JE, Kothari V, Bornfeldt KE (2023) Quartet of APOCs and the different roles they play in diabetes. *Arterioscler Thromb Vasc Biol* 43(7):1124–1133. <https://doi.org/10.1161/ATVBAHA.122.318290>
- Huff MW, Hegele RA (2013) Apolipoprotein C-III: going back to the future for a lipid drug target. *Circ Res* 112(11):1405–1408. <https://doi.org/10.1161/CIRCRESAHA.113.301464>
- Ioka RX, Kang MJ, Kamiyama S, Kim DH, Magoori K, Kamataki A, Ito Y, Takei YA, Sasaki M, Suzuki T, Sasano H, Takahashi S, Sakai J, Fujino T, Yamamoto TT (2003) Expression cloning and characterization of a novel glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein, GPI-HBP1. *J Biol Chem* 278(9):7344–7349. <https://doi.org/10.1074/jbc.M211932200>
- Ito Y, Azrolan N, O'Connell A, Walsh A, Breslow JL (1990) Hypertriglyceridemia as a result of human apo CIII gene expression in transgenic mice. *Science* 249(4970):790–793. <https://doi.org/10.1126/science.2167514>
- Jin N, Matter WF, Michael LF, Qian Y, Gheyi T, Cano L, Perez C, Lafuente C, Broughton HB, Espada A (2021) The angiopoietin-like protein 3 and 8 complex interacts with lipoprotein lipase and induces LPL cleavage. *ACS Chem Biol* 16(3):457–462. <https://doi.org/10.1021/acscmbio.0c00954>

- Kane JP, Hardman DA, Paulus HE (1980) Heterogeneity of apolipoprotein B: isolation of a new species from human chylomicrons. *Proc Natl Acad Sci U S A* 77(5):2465–2469. <https://doi.org/10.1073/pnas.77.5.2465>
- Kersten S (2014) Physiological regulation of lipoprotein lipase. *Biochim Biophys Acta* 1841(7):919–933. <https://doi.org/10.1016/j.bbali.2014.03.013>
- Kersten S, Mandard S, Tan NS, Escher P, Metzger D, Chambon P, Gonzalez FJ, Desvergne B, Wahli W (2000) Characterization of the fasting-induced adipose factor FIAF, a novel peroxisome proliferator-activated receptor target gene. *J Biol Chem* 275(37):28488–28493. <https://doi.org/10.1074/jbc.M004029200>
- Khan FI, Lan D, Durrani R, Huan W, Zhao Z, Wang Y (2017) The lid domain in lipases: structural and functional determinant of enzymatic properties. *Front Bioeng Biotechnol* 5:16. <https://doi.org/10.3389/fbioe.2017.00016>
- Kim I, Kim HG, Kim H, Kim HH, Park SK, Uhm CS, Lee ZH, Koh GY (2000) Hepatic expression, synthesis and secretion of a novel fibrinogen/angiopoietin-related protein that prevents endothelial-cell apoptosis. *Biochem J* 346 Pt 3(Pt 3):603–610
- Kingsley LJ, Lill MA (2015) Substrate tunnels in enzymes: structure-function relationships and computational methodology. *Proteins* 83(4):599–611. <https://doi.org/10.1002/prot.24772>
- Kobayashi Y, Nakajima T, Inoue I (2002) Molecular modeling of the dimeric structure of human lipoprotein lipase and functional studies of the carboxyl-terminal domain. *Eur J Biochem* 269(18):4701–4710. <https://doi.org/10.1046/j.1432-1033.2002.03179.x>
- Kolovou G, Kolovou V, Katsiki N (2022) Volanesorsen: a new era in the treatment of severe hypertriglyceridemia. *J Clin Med* 11(4). <https://doi.org/10.3390/jcm11040982>
- Korn ED (1954) Properties of clearing factor obtained from rat heart acetone powder. *Science* 120(3114):399–400. <https://doi.org/10.1126/science.120.3114.399-a>
- Korn ED (1955) Clearing factor, a heparin-activated lipoprotein lipase. I. Isolation and characterization of the enzyme from normal rat heart. *J Biol Chem* 215(1):1–14
- Krauss RM, Herbert PN, Levy RI, Fredrickson DS (1973) Further observations on the activation and inhibition of lipoprotein lipase by apolipoproteins. *Circ Res* 33(4):403–411. <https://doi.org/10.1161/01.res.33.4.403>
- Kristensen KK, Leth-Espensen KZ, Mertens HDT, Birrane G, Meiyappan M, Olivecrona G, Jorgensen TJD, Young SG, Ploug M (2020) Unfolding of monomeric lipoprotein lipase by ANGPTL4: insight into the regulation of plasma triglyceride metabolism. *Proc Natl Acad Sci U S A* 117(8):4337–4346. <https://doi.org/10.1073/pnas.192020117>
- Kumari A, Gronnemoose AL, Kristensen KK, Winther AL, Young SG, Jorgensen TJD, Ploug M (2023) Inverse effects of APOC2 and ANGPTL4 on the conformational dynamics of lid-anchoring structures in lipoprotein lipase. *Proc Natl Acad Sci U S A* 120(18):e2221888120. <https://doi.org/10.1073/pnas.2221888120>
- Lafferty MJ, Bradford KC, Erie DA, Neher SB (2013) Angiopoietin-like protein 4 inhibition of lipoprotein lipase: evidence for reversible complex formation. *J Biol Chem* 288(40):28524–28534. <https://doi.org/10.1074/jbc.M113.497602>
- Lalazar A, Mahley RW (1989) Human apolipoprotein E. Receptor binding activity of truncated variants with carboxyl-terminal deletions. *J Biol Chem* 264(15):8447–8450
- Larsson M, Vorrso E, Talmud P, Lookene A, Olivecrona G (2013) Apolipoproteins C-I and C-III inhibit lipoprotein lipase activity by displacement of the enzyme from lipid droplets. *J Biol Chem* 288(47):33997–34008. <https://doi.org/10.1074/jbc.M113.495366>
- Le NA, Walter MF (2007) The role of hypertriglyceridemia in atherosclerosis. *Curr Atheroscler Rep* 9(2):110–115. <https://doi.org/10.1007/s11883-007-0006-7>
- Lee EC, Desai U, Gololobov G, Hong S, Feng X, Yu XC, Gay J, Wilganowski N, Gao C, Du LL, Chen J, Hu Y, Zhao S, Kirkpatrick L, Schneider M, Zambrowicz BP, Landes G, Powell DR, Sonnenburg WK (2009) Identification of a new functional domain in angiopoietin-like 3 (ANGPTL3) and angiopoietin-like 4 (ANGPTL4) involved in binding and inhibition of lipoprotein lipase (LPL). *J Biol Chem* 284(20):13735–13745. <https://doi.org/10.1074/jbc.M807899200>

- Lei X, Shi F, Basu D, Huq A, Routhier S, Day R, Jin W (2011) Proteolytic processing of angiopoietin-like protein 4 by proprotein convertases modulates its inhibitory effects on lipoprotein lipase activity. *J Biol Chem* 286(18):15747–15756. <https://doi.org/10.1074/jbc.M110.217638>
- Levi V, Gonzalez Flecha FL (2002) Reversible fast-dimerization of bovine serum albumin detected by fluorescence resonance energy transfer. *Biochim Biophys Acta* 1599(1-2):141–148. [https://doi.org/10.1016/s1570-9639\(02\)00414-4](https://doi.org/10.1016/s1570-9639(02)00414-4)
- Loeffler B, Heeren J, Blaesser M, Radner H, Kayser D, Aydin B, Merkel M (2007) Lipoprotein lipase-facilitated uptake of LDL is mediated by the LDL receptor. *J Lipid Res* 48(2):288–298. <https://doi.org/10.1194/jlr.M600292-JLR200>
- Lookene A, Chevreuil O, Ostergaard P, Olivecrona G (1996) Interaction of lipoprotein lipase with heparin fragments and with heparan sulfate: stoichiometry, stabilization, and kinetics. *Biochemistry* 35(37):12155–12163. <https://doi.org/10.1021/bi960008e>
- Lookene A, Groot NB, Kastelein JJ, Olivecrona G, Bruin T (1997) Mutation of tryptophan residues in lipoprotein lipase. Effects on stability, immunoreactivity, and catalytic properties. *J Biol Chem* 272(2):766–772. <https://doi.org/10.1074/jbc.272.2.766>
- Lookene A, Zhang L, Hultin M, Olivecrona G (2004) Rapid subunit exchange in dimeric lipoprotein lipase and properties of the inactive monomer. *J Biol Chem* 279(48):49964–49972. <https://doi.org/10.1074/jbc.M407419200>
- Lookene A, Beckstead JA, Nilsson S, Olivecrona G, Ryan RO (2005) Apolipoprotein A-V-heparin interactions: implications for plasma lipoprotein metabolism. *J Biol Chem* 280(27):25383–25387. <https://doi.org/10.1074/jbc.M501589200>
- Lutz EP, Merkel M, Kako Y, Melford K, Radner H, Breslow JL, Bensadoun A, Goldberg IJ (2001) Heparin-binding defective lipoprotein lipase is unstable and causes abnormalities in lipid delivery to tissues. *J Clin Invest* 107(9):1183–1192. <https://doi.org/10.1172/JCI11774>
- MacRaild CA, Howlett GJ, Gooley PR (2004) The structure and interactions of human apolipoprotein C-II in dodecyl phosphocholine. *Biochemistry* 43(25):8084–8093. <https://doi.org/10.1021/bi0498171>
- Maeda N, Li H, Lee D, Oliver P, Quarfordt SH, Osada J (1994) Targeted disruption of the apolipoprotein C-III gene in mice results in hypotriglyceridemia and protection from postprandial hypertriglyceridemia. *J Biol Chem* 269(38):23610–23616
- Mahley RW (2016) Apolipoprotein E: from cardiovascular disease to neurodegenerative disorders. *J Mol Med (Berl)* 94(7):739–746. <https://doi.org/10.1007/s00109-016-1427-y>
- Mahley RW, Innerarity TL, Rall SC Jr, Weisgraber KH (1984) Plasma lipoproteins: apolipoprotein structure and function. *J Lipid Res* 25(12):1277–1294
- Mao HZ, Ehrhardt N, Bedoya C, Gomez JA, DeZwaan-McCabe D, Mungrue IN, Kaufman RJ, Rutkowski DT, Peterfy M (2014) Lipase maturation factor 1 (lmf1) is induced by endoplasmic reticulum stress through activating transcription factor 6alpha (Atf6alpha) signaling. *J Biol Chem* 289(35):24417–24427. <https://doi.org/10.1074/jbc.M114.588764>
- Masuno H, Blanchette-Mackie EJ, Chernick SS, Scow RO (1990) Synthesis of inactive nonsecretable high mannose-type lipoprotein lipase by cultured brown adipocytes of combined lipase-deficient cld/cld mice. *J Biol Chem* 265(3):1628–1638
- McConathy WJ, Gesquiere JC, Bass H, Tartar A, Fruchart JC, Wang CS (1992) Inhibition of lipoprotein lipase activity by synthetic peptides of apolipoprotein C-III. *J Lipid Res* 33(7):995–1003
- McIlhargey TL, Yang Y, Wong H, Hill JS (2003) Identification of a lipoprotein lipase cofactor-binding site by chemical cross-linking and transfer of apolipoprotein C-II-responsive lipolysis from lipoprotein lipase to hepatic lipase. *J Biol Chem* 278(25):23027–23035. <https://doi.org/10.1074/jbc.M300315200>
- McQueen AE, Kanamaluru D, Yan K, Gray NE, Wu L, Li ML, Chang A, Hasan A, Stifler D, Koliwad SK, Wang JC (2017) The C-terminal fibrinogen-like domain of angiopoietin-like 4 stimulates adipose tissue lipolysis and promotes energy expenditure. *J Biol Chem* 292(39):16122–16134. <https://doi.org/10.1074/jbc.M117.803973>

- Meneghetti MC, Hughes AJ, Rudd TR, Nader HB, Powell AK, Yates EA, Lima MA (2015) Heparan sulfate and heparin interactions with proteins. *J R Soc Interface* 12(110):0589. <https://doi.org/10.1098/rsif.2015.0589>
- Merkel M, Loeffler B, Kluger M, Fabig N, Geppert G, Pennacchio LA, Laatsch A, Heeren J (2005) Apolipoprotein AV accelerates plasma hydrolysis of triglyceride-rich lipoproteins by interaction with proteoglycan-bound lipoprotein lipase. *J Biol Chem* 280(22):21553–21560. <https://doi.org/10.1074/jbc.M411412200>
- Meyers NL, Larsson M, Olivecrona G, Small DM (2015) A pressure-dependent model for the regulation of lipoprotein lipase by apolipoprotein C-II. *J Biol Chem* 290(29):18029–18044. <https://doi.org/10.1074/jbc.M114.629865>
- Miller NE (1990) HDL metabolism and its role in lipid transport. *Eur Heart J* 11(Suppl H):1–3. https://doi.org/10.1093/eurheartj/11.suppl_h.1
- Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) ColabFold: making protein folding accessible to all. *Nat Methods* 19(6):679–682. <https://doi.org/10.1038/s41592-022-01488-1>
- Mueller B, Lilley BN, Ploegh HL (2006) SEL1L, the homologue of yeast Hrd3p, is involved in protein dislocation from the mammalian ER. *J Cell Biol* 175(2):261–270. <https://doi.org/10.1083/jcb.200605196>
- Mueller B, Klemm EJ, Spooner E, Claessen JH, Ploegh HL (2008) SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. *Proc Natl Acad Sci U S A* 105(34):12325–12330. <https://doi.org/10.1073/pnas.0805371105>
- Mysling S, Kristensen KK, Larsson M, Beigneux AP, Gardsvoll H, Fong LG, Bensadoun A, Jorgensen TJ, Young SG, Ploug M (2016a) The acidic domain of the endothelial membrane protein GPIHBP1 stabilizes lipoprotein lipase activity by preventing unfolding of its catalytic domain. *Elife* 5:e12095. <https://doi.org/10.7554/eLife.12095>
- Mysling S, Kristensen KK, Larsson M, Kovrov O, Bensadoun A, Jorgensen TJ, Olivecrona G, Young SG, Ploug M (2016b) The angiopoietin-like protein ANGPTL4 catalyzes unfolding of the hydrolase domain in lipoprotein lipase and the endothelial membrane protein GPIHBP1 counteracts this unfolding. *Elife* 5. <https://doi.org/10.7554/eLife.20958>
- Nielsen MS, Brejning J, Garcia R, Zhang H, Hayden MR, Vilaro S, Gliemann J (1997) Segments in the C-terminal folding domain of lipoprotein lipase important for binding to the low density lipoprotein receptor-related protein and to heparan sulfate proteoglycans. *J Biol Chem* 272(9):5821–5827. <https://doi.org/10.1074/jbc.272.9.5821>
- Niemeier A, Gafvels M, Heeren J, Meyer N, Angelin B, Beisiegel U (1996) VLDL receptor mediates the uptake of human chylomicron remnants in vitro. *J Lipid Res* 37(8):1733–1742
- Nilsson SK, Lookene A, Beckstead JA, Gliemann J, Ryan RO, Olivecrona G (2007) Apolipoprotein A-V interaction with members of the low density lipoprotein receptor gene family. *Biochemistry* 46(12):3896–3904. <https://doi.org/10.1021/bi7000533>
- Nykjaer A, Bengtsson-Olivecrona G, Lookene A, Moestrup SK, Petersen CM, Weber W, Beisiegel U, Gliemann J (1993) The alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein binds lipoprotein lipase and beta-migrating very low density lipoprotein associated with the lipase. *J Biol Chem* 268(20):15048–15055
- O'Brien PJ, Alborn WE, Sloan JH, Ulmer M, Boodhoo A, Knierman MD, Schultze AE, Konrad RJ (2005) The novel apolipoprotein A5 is present in human serum, is associated with VLDL, HDL, and chylomicrons, and circulates at very low concentrations compared with other apolipoproteins. *Clin Chem* 51(2):351–359. <https://doi.org/10.1373/clinchem.2004.040824>
- Oka K, Wang-Iverson P, Paterniti JR Jr, Brown WV (1989a) Interaction of lipoprotein lipase with heparin. *Ann NY Acad Sci* 556:173–180. <https://doi.org/10.1111/j.1749-6632.1989.tb22501.x>
- Oka K, Yuan JG, Senda M, Masibay AS, Qasba PK, Masuno H, Scow RO, Paterniti JR Jr, Brown WV (1989b) Expression of lipoprotein lipase gene in combined lipase deficiency. *Biochim Biophys Acta* 1008(3):351–354. [https://doi.org/10.1016/0167-4781\(89\)90027-4](https://doi.org/10.1016/0167-4781(89)90027-4)

- Olivecrona T, Egelrud T, Iverius PH, Lindahl U (1971) Evidence for an ionic binding of lipoprotein lipase to heparin. *Biochem Biophys Res Commun* 43(3):524–529. [https://doi.org/10.1016/0006-291x\(71\)90645-0](https://doi.org/10.1016/0006-291x(71)90645-0)
- Olivecrona T, Bengtsson-Olivecrona G, Chernick SS, Scow RO (1986) Effect of combined lipase deficiency (cld/cld) on hepatic and lipoprotein lipase activities in liver and plasma of newborn mice. *Biochim Biophys Acta* 876(2):243–248
- Ollila OH, Lamberg A, Lehtivaara M, Koivuniemi A, Vattulainen I (2012) Interfacial tension and surface pressure of high density lipoprotein, low density lipoprotein, and related lipid droplets. *Biophys J* 103(6):1236–1244. <https://doi.org/10.1016/j.bpj.2012.08.023>
- Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J et al (1992) The alpha/beta hydrolase fold. *Protein Eng* 5(3):197–211. <https://doi.org/10.1093/protein/5.3.197>
- Ono M, Shimizugawa T, Shimamura M, Yoshida K, Noji-Sakikawa C, Ando Y, Koishi R, Furukawa H (2003) Protein region important for regulation of lipid metabolism in angiopoietin-like 3 (ANGPTL3): ANGPTL3 is cleaved and activated in vivo. *J Biol Chem* 278(43):41804–41809. <https://doi.org/10.1074/jbc.M302861200>
- Osborne JC Jr, Bengtsson-Olivecrona G, Lee NS, Olivecrona T (1985) Studies on inactivation of lipoprotein lipase: role of the dimer to monomer dissociation. *Biochemistry* 24(20):5606–5611. <https://doi.org/10.1021/bi00341a048>
- Paterniti JR Jr, Brown WV, Ginsberg HN, Artzt K (1983) Combined lipase deficiency (cld): a lethal mutation on chromosome 17 of the mouse. *Science* 221(4606):167–169
- Pennacchio LA, Olivier M, Hubacek JA, Cohen JC, Cox DR, Fruchart JC, Krauss RM, Rubin EM (2001) An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing. *Science* 294(5540):169–173. <https://doi.org/10.1126/science.1064852>
- Peterfy M, Ben-Zeev O, Mao HZ, Weissglas-Volkov D, Aouizerat BE, Pullinger CR, Frost PH, Kane JP, Malloy MJ, Reue K, Pajukanta P, Doolittle MH (2007) Mutations in LMF1 cause combined lipase deficiency and severe hypertriglyceridemia. *Nat Genet* 39(12):1483–1487. <https://doi.org/10.1038/ng.2007.24>
- Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci* 30(1):70–82. <https://doi.org/10.1002/pro.3943>
- Pollin TI, Damcott CM, Shen H, Ott SH, Shelton J, Horenstein RB, Post W, McLenithan JC, Bielak LF, Peyser PA, Mitchell BD, Miller M, O’Connell JR, Shuldiner AR (2008) A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. *Science* 322(5908):1702–1705. <https://doi.org/10.1126/science.1161524>
- Pradines-Figuères A, Vannier C, Ailhaud G (1990) Lipoprotein lipase stored in adipocytes and muscle cells is a cryptic enzyme. *J Lipid Res* 31(8):1467–1476
- Quagliarini F, Wang Y, Kozlitina J, Grishin NV, Hyde R, Boerwinkle E, Valenzuela DM, Murphy AJ, Cohen JC, Hobbs HH (2012) Atypical angiopoietin-like protein that regulates ANGPTL3. *Proc Natl Acad Sci U S A* 109(48):19751–19756. <https://doi.org/10.1073/pnas.1217552109>
- Rapraeger AC (2001) Molecular interactions of syndecans during development. *Semin Cell Dev Biol* 12(2):107–116. <https://doi.org/10.1006/scdb.2000.0239>
- Razzaghi H, Day BW, McClure RJ, Kamboh MI (2001) Structure-function analysis of D9N and N291S mutations in human lipoprotein lipase using molecular modelling. *J Mol Graph Model* 19(6):487–494, 587–490. [https://doi.org/10.1016/s1093-3263\(00\)00096-6](https://doi.org/10.1016/s1093-3263(00)00096-6)
- Reddy MN, Maraganore JM, Meredith SC, Heinrikson RL, Kezdy FJ (1986) Isolation of an active-site peptide of lipoprotein lipase from bovine milk and determination of its amino acid sequence. *J Biol Chem* 261(21):9678–9683
- Reichl D (1969) Stabilization and partial purification of lipoprotein lipase. *Biochem J* 114(4):71P. <https://doi.org/10.1042/bj1140071p>
- Reimund M, Larsson M, Kovrov O, Kasvandik S, Olivecrona G, Lookene A (2015) Evidence for two distinct binding sites for lipoprotein lipase on glycosylphosphatidylinositol-anchored high

- density lipoprotein-binding protein 1 (GPIHBP1). *J Biol Chem* 290(22):13919–13934. <https://doi.org/10.1074/jbc.M114.634626>
- Reimund M, Kovrov O, Olivecrona G, Lookene A (2017) Lipoprotein lipase activity and interactions studied in human plasma by isothermal titration calorimetry. *J Lipid Res* 58(1):279–288. <https://doi.org/10.1194/jlr.D071787>
- Reizes O, Goldberger O, Smith AC, Xu Z, Bernfield M, Bickel PE (2006) Insulin promotes shedding of syndecan ectodomains from 3T3-L1 adipocytes: a proposed mechanism for stabilization of extracellular lipoprotein lipase. *Biochemistry* 45(18):5703–5711. <https://doi.org/10.1021/bi052263h>
- Rensen PC, van Berkel TJ (1996) Apolipoprotein E effectively inhibits lipoprotein lipase-mediated lipolysis of chylomicron-like triglyceride-rich lipid emulsions in vitro and in vivo. *J Biol Chem* 271(25):14791–14799. <https://doi.org/10.1074/jbc.271.25.14791>
- Reyes-Soffer G, Sztalryd C, Horenstein RB, Holleran S, Matveyenko A, Thomas T, Nandakumar R, Ngai C, Karmally W, Ginsberg HN, Ramakrishnan R, Pollin TI (2019) Effects of APOC3 heterozygous deficiency on plasma lipid and lipoprotein metabolism. *Arterioscler Thromb Vasc Biol* 39(1):63–72. <https://doi.org/10.1161/ATVBAHA.118.311476>
- Risti R, Gunn KH, Hiis-Hommuk K, Seeba NN, Karimi H, Villo L, Vendelin M, Neher SB, Lookene A (2023) Combined action of albumin and heparin regulates lipoprotein lipase oligomerization, stability, and ligand interactions. *PLoS One* 18(4):e0283358. <https://doi.org/10.1371/journal.pone.0283358>
- Roberts BS, Babilonia-Rosa MA, Broadwell LJ, Wu MJ, Neher SB (2018) Lipase maturation factor 1 affects redox homeostasis in the endoplasmic reticulum. *EMBO J* 37(19). <https://doi.org/10.15252/embj.201797379>
- Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen A, Hobbs HH, Cohen JC (2007) Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. *Nat Genet* 39(4):513–516. <https://doi.org/10.1038/ng1984>
- Sarrazin S, Lamanna WC, Esko JD (2011) Heparan sulfate proteoglycans. *Cold Spring Harb Perspect Biol* 3(7). <https://doi.org/10.1101/cshperspect.a004952>
- Segrest JP, Jones MK, De Loof H, Brouillette CG, Venkatachalapathi YV, Anantharamaiah GM (1992) The amphipathic helix in the exchangeable apolipoproteins: a review of secondary structure and function. *J Lipid Res* 33(2):141–166
- Sehayek E, Eisenberg S (1991) Mechanisms of inhibition by apolipoprotein C of apolipoprotein E-dependent cellular metabolism of human triglyceride-rich lipoproteins through the low density lipoprotein receptor pathway. *J Biol Chem* 266(27):18259–18267
- Sendak RA, Bensadoun A (1998) Identification of a heparin-binding domain in the distal carboxyl-terminal region of lipoprotein lipase by site-directed mutagenesis. *J Lipid Res* 39(6):1310–1315
- Sha H, Sun S, Francisco AB, Ehrhardt N, Xue Z, Liu L, Lawrence P, Mattijssen F, Guber RD, Panhwar MS, Brenna JT, Shi H, Xue B, Kersten S, Bensadoun A, Peterfy M, Long Q, Qi L (2014) The ER-associated degradation adaptor protein Sel1L regulates LPL secretion and lipid metabolism. *Cell Metab* 20(3):458–470. <https://doi.org/10.1016/j.cmet.2014.06.015>
- Shan L, Yu XC, Liu Z, Hu Y, Sturgis LT, Miranda ML, Liu Q (2009) The angiopoietin-like proteins ANGPTL3 and ANGPTL4 inhibit lipoprotein lipase activity through distinct mechanisms. *J Biol Chem* 284(3):1419–1424. <https://doi.org/10.1074/jbc.M808477200>
- Shen Y, Lookene A, Nilsson S, Olivecrona G (2002) Functional analyses of human apolipoprotein CII by site-directed mutagenesis: identification of residues important for activation of lipoprotein lipase. *J Biol Chem* 277(6):4334–4342. <https://doi.org/10.1074/jbc.M105421200>
- Shetty SK, Walzem RL, Davies BSJ (2020) A novel NanoBiT-based assay monitors the interaction between lipoprotein lipase and GPIHBP1 in real time. *J Lipid Res* 61(4):546–559. <https://doi.org/10.1194/jlr.D119000388>
- Shimada K, Gill PJ, Silbert JE, Douglas WH, Fanburg BL (1981) Involvement of cell surface heparin sulfate in the binding of lipoprotein lipase to cultured bovine endothelial cells. *J Clin Invest* 68(4):995–1002. <https://doi.org/10.1172/jci110354>

- Shimizu-gawa T, Ono M, Shimamura M, Yoshida K, Ando Y, Koishi R, Ueda K, Inaba T, Minekura H, Kohama T, Furukawa H (2002) ANGPTL3 decreases very low density lipoprotein triglyceride clearance by inhibition of lipoprotein lipase. *J Biol Chem* 277(37):33742–33748. <https://doi.org/10.1074/jbc.M203215200>
- Song W, Beigneux AP, Winther AL, Kristensen KK, Gronnemoose AL, Yang Y, Tu Y, Munguia P, Morales J, Jung H, de Jong PJ, Jung CJ, Miyashita K, Kimura T, Nakajima K, Murakami M, Birrane G, Jiang H, Tontonoz P, Ploug M, Fong LG, Young SG (2022) Electrostatic sheathing of lipoprotein lipase is essential for its movement across capillary endothelial cells. *J Clin Invest* 132(5). <https://doi.org/10.1172/JCI157500>
- Sonnenburg WK, Yu D, Lee EC, Xiong W, Gololobov G, Key B, Gay J, Wilganowski N, Hu Y, Zhao S, Schneider M, Ding ZM, Zambrowicz BP, Landes G, Powell DR, Desai U (2009) GPIHBP1 stabilizes lipoprotein lipase and prevents its inhibition by angiopoietin-like 3 and angiopoietin-like 4. *J Lipid Res* 50(12):2421–2429. <https://doi.org/10.1194/jlr.M900145-JLR200>
- Sosnowska B, Adach W, Surma S, Rosenson RS, Banach M (2022) Evinacumab, an ANGPTL3 Inhibitor, in the treatment of dyslipidemia. *J Clin Med* 12(1). <https://doi.org/10.3390/jcm12010168>
- Stitzel NO, Khera AV, Wang X, Bierhals AJ, Vourakis AC, Sperry AE, Natarajan P, Klarin D, Emdin CA, Zekavat SM, Nomura A, Erdmann J, Schunkert H, Samani NJ, Kraus WE, Shah SH, Yu B, Boerwinkle E, Rader DJ, Gupta N, Frossard PM, Rasheed A, Danesh J, Lander ES, Gabriel S, Saleheen D, Musunuru K, Kathiresan S, PROMIS and Myocardial Infarction Genetics Consortium Investigators (2017) ANGPTL3 Deficiency and Protection Against Coronary Artery Disease. *J Am Coll Cardiol* 69(16):2054–2063. <https://doi.org/10.1016/j.jacc.2017.02.030>
- Sukonina V, Lookene A, Olivecrona T, Olivecrona G (2006) Angiopoietin-like protein 4 converts lipoprotein lipase to inactive monomers and modulates lipase activity in adipose tissue. *Proc Natl Acad Sci U S A* 103(46):17450–17455
- Sundberg EL, Deng Y, Burd CG (2019) Syndecan-1 mediates sorting of soluble lipoprotein lipase with sphingomyelin-rich membrane in the golgi apparatus. *Dev Cell* 51(3):387–398.e384. <https://doi.org/10.1016/j.devcel.2019.08.014>
- Sviridov D, Dasseux A, Reimund M, Pryor M, Drake SK, Jarin Z, Wolska A, Pastor RW, Remaley AT (2023) Short hydrocarbon stapled ApoC2-mimetic peptides activate lipoprotein lipase and lower plasma triglycerides in mice. *Front Cardiovasc Med* 10:1223920. <https://doi.org/10.3389/fcvm.2023.1223920>
- Sylvers-Davie KL, Segura-Roman A, Salvi AM, Schache KJ, Davies BSJ (2021) Angiopoietin-like 3 inhibition of endothelial lipase is not modulated by angiopoietin-like 8. *J Lipid Res* 62:100112. <https://doi.org/10.1016/j.jlr.2021.100112>
- Talmud PJ, Smart M, Presswood E, Cooper JA, Nicaud V, Drenos F, Palmen J, Marmot MG, Boekholdt SM, Wareham NJ, Khaw KT, Kumari M, Humphries SE, EARSII Consortium; HIFMECH Consortium (2008) ANGPTL4 E40K and T266M: effects on plasma triglyceride and HDL levels, postprandial responses, and CHD risk. *Arterioscler Thromb Vasc Biol* 28(12):2319–2325. <https://doi.org/10.1161/ATVBAHA.108.176917>
- van der Vliet HN, Schaap FG, Levels JH, Ottenhoff R, Looije N, Wesseling JG, Groen AK, Chamuleau RA (2002) Adenoviral overexpression of apolipoprotein A-V reduces serum levels of triglycerides and cholesterol in mice. *Biochem Biophys Res Commun* 295(5):1156–1159. [https://doi.org/10.1016/s0006-291x\(02\)00808-2](https://doi.org/10.1016/s0006-291x(02)00808-2)
- van Tilbeurgh H, Sarda L, Verger R, Cambillau C (1992) Structure of the pancreatic lipase-procolipase complex. *Nature* 359(6391):159–162. <https://doi.org/10.1038/359159a0>
- van Tilbeurgh H, Egloff MP, Martinez C, Rugani N, Verger R, Cambillau C (1993) Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography. *Nature* 362(6423):814–820. <https://doi.org/10.1038/362814a0>
- van Tilbeurgh H, Roussel A, Lalouel JM, Cambillau C (1994) Lipoprotein lipase. Molecular model based on the pancreatic lipase x-ray structure: consequences for heparin binding and catalysis. *J Biol Chem* 269(6):4626–4633

- Vilella E, Joven J, Fernandez M, Vilaro S, Brunzell JD, Olivecrona T, Bengtsson-Olivecrona G (1993) Lipoprotein lipase in human plasma is mainly inactive and associated with cholesterol-rich lipoproteins. *J Lipid Res* 34(9):1555–1564
- Wang Y, Quagliarini F, Gusarova V, Gromada J, Valenzuela DM, Cohen JC, Hobbs HH (2013) Mice lacking ANGPTL8 (Betatrophin) manifest disrupted triglyceride metabolism without impaired glucose homeostasis. *Proc Natl Acad Sci U S A* 110(40):16109–16114. <https://doi.org/10.1073/pnas.1315292110>
- Whitacre BE, Howles P, Street S, Morris J, Swertfeger D, Davidson WS (2022) Apolipoprotein E content of VLDL limits LPL-mediated triglyceride hydrolysis. *J Lipid Res* 63(1):100157. <https://doi.org/10.1016/j.jlr.2021.100157>
- Wilson C, Wardell MR, Weisgraber KH, Mahley RW, Agard DA (1991) Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science* 252(5014):1817–1822. <https://doi.org/10.1126/science.2063194>
- Wolska A, Lo L, Sviridov DO, Pourmousa M, Pryor M, Ghosh SS, Kakkar R, Davidson M, Wilson S, Pastor RW, Goldberg IJ, Basu D, Drake SK, Cougnoux A, Wu MJ, Neher SB, Freeman LA, Tang J, Amar M, Devalaraja M, Remaley AT (2020) A dual apolipoprotein C-II mimetic-apolipoprotein C-III antagonist peptide lowers plasma triglycerides. *Sci Transl Med* 12(528). <https://doi.org/10.1126/scitranslmed.aaw7905>
- Wu X, Siggel M, Ovchinnikov S, Mi W, Svetlov V, Nudler E, Liao M, Hummer G, Rapoport TA (2020) Structural basis of ER-associated protein degradation mediated by the Hrd1 ubiquitin ligase complex. *Science* 368(6489). <https://doi.org/10.1126/science.aaz2449>
- Wu SA, Kersten S, Qi L (2021) Lipoprotein lipase and its regulators: an unfolding story. *Trends Endocrinol Metab* 32(1):48–61. <https://doi.org/10.1016/j.tem.2020.11.005>
- Wu SA, Shen C, Wei X, Zhang X, Wang S, Chen X, Torres M, Lu Y, Lin LL, Wang HH, Hunter AH, Fang D, Sun S, Ivanova MI, Lin Y, Qi L (2023) The mechanisms to dispose of misfolded proteins in the endoplasmic reticulum of adipocytes. *Nat Commun* 14(1):3132. <https://doi.org/10.1038/s41467-023-38690-4>
- Yamada N, Murase T (1980) Modulation, by apolipoprotein E, of lipoprotein lipase activity. *Biochem Biophys Res Commun* 94(2):710–715. [https://doi.org/10.1016/0006-291x\(80\)91290-5](https://doi.org/10.1016/0006-291x(80)91290-5)
- Yau MH, Wang Y, Lam KS, Zhang J, Wu D, Xu A (2009) A highly conserved motif within the NH2-terminal coiled-coil domain of angiopoietin-like protein 4 confers its inhibitory effects on lipoprotein lipase by disrupting the enzyme dimerization. *J Biol Chem* 284(18):11942–11952. <https://doi.org/10.1074/jbc.M809802200>
- Yin W, Romeo S, Chang S, Grishin NV, Hobbs HH, Cohen JC (2009) Genetic variation in ANGPTL4 provides insights into protein processing and function. *J Biol Chem* 284(19):13213–13222. <https://doi.org/10.1074/jbc.M900553200>
- Yoon JC, Chickering TW, Rosen ED, Dussault B, Qin Y, Soukas A, Friedman JM, Holmes WE, Spiegelman BM (2000) Peroxisome proliferator-activated receptor gamma target gene encoding a novel angiopoietin-related protein associated with adipose differentiation. *Mol Cell Biol* 20(14):5343–5349. <https://doi.org/10.1128/MCB.20.14.5343-5349.2000>
- Yoshida K, Shimizugawa T, Ono M, Furukawa H (2002) Angiopoietin-like protein 4 is a potent hyperlipidemia-inducing factor in mice and inhibitor of lipoprotein lipase. *J Lipid Res* 43(11):1770–1772
- Zambon A, Schmidt I, Beisiegel U, Brunzell JD (1996) Dimeric lipoprotein lipase is bound to triglyceride-rich plasma lipoproteins. *J Lipid Res* 37(11):2394–2404
- Zdunek J, Martinez GV, Schleucher J, Lycksell PO, Yin Y, Nilsson S, Shen Y, Olivecrona G, Wijmenga S (2003) Global structure and dynamics of human apolipoprotein CII in complex with micelles: evidence for increased mobility of the helix involved in the activation of lipoprotein lipase. *Biochemistry* 42(7):1872–1889. <https://doi.org/10.1021/bi0267184>
- Zhang R (2012) Lipasin, a novel nutritionally-regulated liver-enriched factor that regulates serum triglyceride levels. *Biochem Biophys Res Commun* 424(4):786–792. <https://doi.org/10.1016/j.bbrc.2012.07.038>

- Zhang R (2016) The ANGPTL3-4-8 model, a molecular mechanism for triglyceride trafficking. *Open Biol* 6(4):150272. <https://doi.org/10.1098/rsob.150272>
- Zhang SH, Reddick RL, Piedrahita JA, Maeda N (1992) Spontaneous hypercholesterolemia and arterial lesions in mice lacking apolipoprotein E. *Science* 258(5081):468–471. <https://doi.org/10.1126/science.1411543>

Chapter 9

Viral Genomic DNA Packaging Machinery



Dorothy E. D. P. Hawkins, Owen C. Godwin, and Alfred A. Antson

Abstract Tailed double-stranded DNA bacteriophage employs a protein terminase motor to package their genome into a preformed protein shell—a system shared with eukaryotic dsDNA viruses such as *herpesviruses*. DNA packaging motor proteins represent excellent targets for antiviral therapy, with Letermovir, which binds *Cytomegalovirus* terminase, already licensed as an effective prophylaxis. In the realm of bacterial viruses, these DNA packaging motors comprise three protein constituents: the portal protein, small terminase and large terminase. The portal protein guards the passage of DNA into the preformed protein shell and acts as a protein interaction hub throughout viral assembly. Small terminase recognises the viral DNA and recruits large terminase, which in turn pumps DNA in an ATP-dependent manner. Large terminase also cleaves DNA at the termination of packaging. Multiple high-resolution structures of each component have been resolved for different phages, but it is only more recently that the field has moved towards cryo-EM reconstructions of protein complexes. In conjunction with highly informative single-particle studies of packaging kinetics, these structures have begun to inspire models for the packaging process and its place among other DNA machines.

Keywords DNA translocation · Large terminase · Small terminase · Portal protein · Bacteriophage · ATPase · Nuclease

Dorothy E. D. P. Hawkins (✉)

York Structural Biology Laboratory, Department of Chemistry, University of York, York, UK
e-mail: dorothy.hawkins@crick.ac.uk

O. C. Godwin · A. A. Antson (✉)

York Structural Biology Laboratory, Department of Chemistry, University of York, York, UK

Structural Biology, The Francis Crick Institute, London, UK

e-mail: owen.godwin@york.ac.uk; fred.antson@york.ac.uk

List of Abbreviations

ASCE	Additional strand, conserved glutamate
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
ATP- γ -S	Adenosine-5'-o-(3-thio-triphosphate), tetralithium salt
Bp	Base pair
CMV	<i>Cytomegalovirus</i>
CTD	Terminal domain
Cryo-EM	Cryo-electron microscopy
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
EPR	Electron paramagnetic resonance
FRET	Fluorescence resonance energy transfer
HTH	Helix-turn-helix
HSV1	<i>Herpes Simplex Virus 1</i>
MD	Molecular dynamics
NTD	N-terminal domain
RNA	Ribonucleic acid
PDB	Protein Data Bank
SPR	Surface plasmon resonance

Introduction to Viral DNA Packaging

Encapsulation of genetic material represents an essential process for viral assembly and thus viability. For prokaryotic dsDNA viruses, this process may be fulfilled by one of three distinct DNA packing mechanisms. Bacteriophage with very small genomes (typically less than 20 kDa) may utilise the nucleation of capsid proteins around the genomic DNA, in an energy-independent manner (Burroughs et al. 2007). However, above this 20 kb threshold, (and indeed for most smaller viruses too), bacteriophage genomes encode an ATPase for DNA packaging, which belongs to either FtsK like or terminase subcategories. FtsK-based motors are utilised exclusively by phage with internal lipid membranes such as *Corticoviruses* and *Tectiviruses*, as well as all *Nucleocytoviricota* eukaryotic viruses such as *Poxvirus* and *Mimivirus* (Burroughs et al. 2007; Chelikani et al. 2014). The FtsK-like motors are also encoded by large eukaryotic transposons known as *Polintons* which may represent the evolutionary origin of viruses with this packaging system (Krupovic and Koonin 2015). Terminase-type systems are instead utilised by *Herpesviridae* viruses and tailed bacteriophages (the *Caudovirales*) which are thought to represent up to 96 % of all known phage (Ackermann 2007). The terminase-based motor system has been subject to extensive study and will thus represent the focus of this chapter.

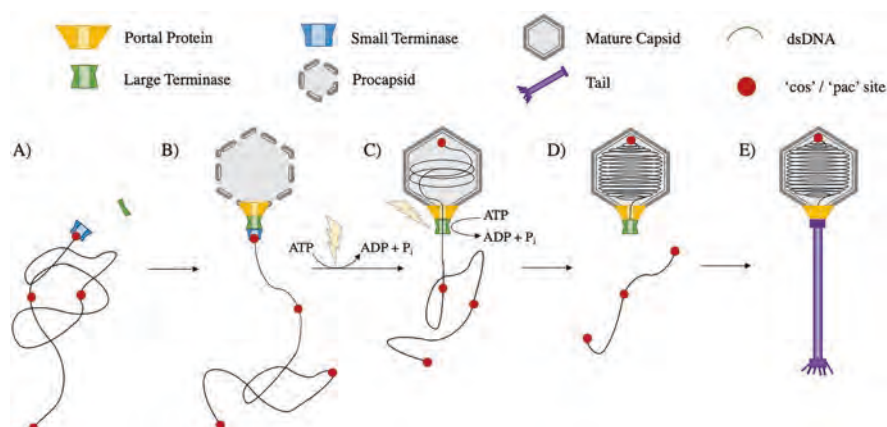


Fig. 9.1 dsDNA bacteriophage assembly. (a) Small terminase recruits large terminase to the viral genomic DNA. (b) Large terminase docks onto the portal vertex of the preformed protein prohead. (c) Large terminase powers packaging of DNA into the prohead using ATP hydrolysis. (d) The prohead expands into a mature capsid, and DNA is cleaved after one genome length has been packaged. (e) The terminase complex dissociates and is replaced by neck and tail proteins to form a mature virion

In terminase packaging systems, a dodecameric portal protein at a unique pentameric vertex of the prohead represents the site for prohead assembly, DNA packaging and later tail attachment. DNA packaging triggers the maturation of proheads into expanded, DNA-filled capsids which show thinner shells, sharper vertices and a larger diameter despite displaying the same T-number and possessing the same number of viral coat proteins (Murialdo and Becker 1978; Earnshaw and Casjens 1980; Tu et al. 2001). Additional proteins often contribute to capsid stability or host recognition. The DNA packaging terminase complex consists of a ring of large terminase subunits (Hendrix 1998) which translocates the genome into the prohead in an ATP-dependent manner (Daudén et al. 2013; Migliori et al. 2014; Mao et al. 2016; Dai et al. 2021; Reyes-Aldrete et al. 2021). Small terminase subunits are required for viral DNA recognition and likely form a second ring beneath. After packaging is completed, the terminase cleaves the DNA, dissociates, and the complex is replaced by neck proteins to prevent the release of genetic material. This is followed by attachment of a pre-assembled tail or tail proteins (Camacho et al. 1979; Rossmann et al. 2004; Lander et al. 2009) to complete virus assembly (Fig. 9.1).

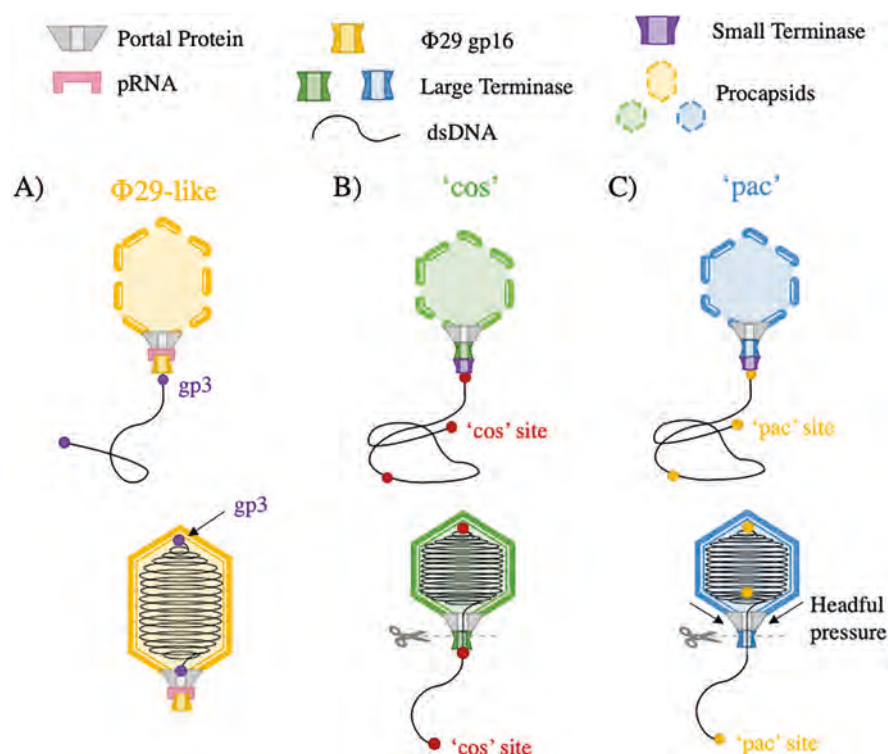


Fig. 9.2 DNA packaging strategies by ds DNA phage. (a) $\Phi 29$ like phages package protein capped genome lengths, (b) cos phage package cos site capped genome units, (c) pac phage package in excess of one complete genome length

Genome Packaging Strategies

Although the mechanism of DNA packaging is generally conserved among the tailed dsDNA bacteriophages, further classifications can be highlighted based on processing of the viral DNA (Fig. 9.2). Most bacteriophage replicate DNA as concatemers containing multiple consecutive copies of the genome (Black 1989). Packaging occurs in a series of unidirectional packaging events along the concatemer, where cleavage of the DNA substrate at termination produces a fresh dsDNA end for a subsequent initiation event. Packaging series typically consists of two to five full genome packaging events but may elongate depending on conditions (Adams et al. 1983).

Replication of the phage $\Phi 29$ DNA unusually produces unit genome length copies, where each 5' end is covalently bound to viral protein gp3 (Guo et al. 1987). The DNA protein-complex is then packaged by the large terminase motor, which is here referred to simply as 'ATPase', where a pentameric RNA molecule bridges the large terminase-portal interaction (Fig. 9.2a) (Reid et al. 1994; Ding et al. 2011).

This prohead RNA (pRNA) is necessary for both packaging initiation and substrate selectivity (Zhao et al. 2015b; Peixuan et al. 1987).

Pac viruses, such as P22, P1, SPP1 and T4, initiate packaging from a specific *pac* (packaging) site within the genome (Tye et al. 1974; Sternberg and Coulby 1990; Tavares et al. 1996; Lin and Black 1998). For SPP1, two separate small terminase oligomers are thought to bind DNA at sequences *pac* L (100 base pairs) and *pac* R (30 base pairs) (Chai et al. 1995). This loops DNA between binding sites, exposing a third sequence, *pac* C, to cleavage by large terminase. Pac virus termination occurs via a ‘head full’ mechanism (Jackson et al. 1978), whereby pressure changes within the capsid are relayed to the terminase. This stimulates large terminase nuclease activity required for DNA cleavage. In practice, a small amount of redundant DNA is always packaged with SPP1 virions harbouring approximately 103% of the genome length (Oliveira et al. 2013) (Fig. 9.2c).

Cos viruses, such as λ , HK97 and P2, instead package a single genome-length DNA (Ziermann and Calendar 1990; Juhala et al. 2000). Upon bacterial infection, the linear genome is injected into the cell, where complimentary sticky ends are sealed by a host ligase. This forms a circular viral genome, which serves as a template for rolling circle replication where the site of adhesion is known as the *cos* site. This sequence is recognised by small terminase, which recruits large terminase to instigate a staggered, sequence-specific cut, producing complimentary overhangs with either a 5’ or 3’ extension (Feiss and Catalano 2007).

For λ phage, the core element within the 200 base pair *cos* sequence is *cos* N, where large terminase produces nicks separated by 12 base pair overhangs (Feiss et al. 1983a, b; Hohn 1983). Downstream element *cos* B is also required for efficient initiation of packaging. Packaging termination occurs at the next *cos* site, where large terminase once again cleaves the *cos* site and efficient termination is reliant on upstream element *cos* (Fig. 9.2b) Q (Cue and Feiss 1993). The minimal DNA-binding site for HK97 has been pinpointed to 15 to 29 base pairs downstream of the *cos* site, presumably positioning large terminase in place for cleavage (Chechik et al. 2023). Termination for both λ and HK97 is also proposed to encompass a headful element, as the efficiency of cleavage is highly dependent on the length of DNA packaged (Feiss et al. 1977; Hawkins et al. 2023).

In common with dsDNA bacteriophage, human *Herpes* virus genome replication also occurs in a rolling concatemeric fashion. DNA sequences *pac* 1 and *pac* 2 must be recognised and cleaved for subsequent packaging of genome length units into preformed proheads (Adelman et al. 2001). The delineation of DNA packaging activities (i.e. DNA recognition, cleavage and ATP-powered translocation) does not strictly adhere to small terminase and large terminase assignments. Indeed, for HSV1, a total of 7 gene products have been implicated in DNA packaging (Heming et al. 2017). Most prominently pUL15 demonstrates ATPase activity critical for packaging, in common with large terminases, but adopts a 50% larger structure including additional domains. The pUL15 complex is also coordinated by two extra proteins PUL28 which PUL33 (Yang et al. 2020). Cytomegalovirus proteins pUL56 and pUL89 have each been implicated in packaging. pUL56 displays ATPase activity and also recognises *pac* sites 1 and 2 (Bogner et al. 1998; Heming et al. 2017),

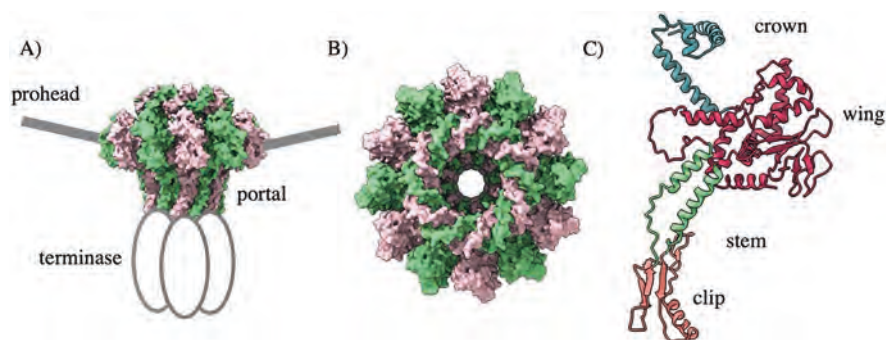


Fig. 9.3 The T4 portal protein BDB 3JA7. (a) The portal is positioned at the unique prohead vertex and coordinates binding of the motor. (b) Portal proteins from dodecameric rings. (c) The portal protein domain architecture

while both proteins demonstrate nuclease activity (Scheffczik et al. 2002). Mutational and structural studies of pUL89 have revealed residues critical for nuclease activity and DNA binding (Theiß et al. 2019).

The Portal Protein

The portal protein exists at the unique prohead vertex and acts as a door in to the prohead (Fig. 9.3a). Despite dramatic divergence of sequence and size, portal proteins always assemble as dodecameric rings (Fig. 9.3b) (Valpuesta et al. 1992; Agirrezabala et al. 2005; Doan and Dokland 2007) with a common homology comprised of a clip, stem, wing and crown domain (Fig. 9.3c). The wing domain shows the widest structural variation and facilitates contact with the coat proteins. Meanwhile, the crown interacts with packaged DNA, and the stem lines the internal channel, which is negatively charged for the smooth passage of DNA. The clip domain makes contact with the packaging machinery during assembly, and later the adaptor proteins for tail attachment. An additional barrel domain is found above the crown within the portal proteins of many *podoviruses* (Tang et al. 2011). Portal incorporation is critical for native capsid formation in phages T4, SPP1 and $\Phi 29$ (Guo et al. 1991; Dröge et al. 2000).

Portal proteins act as interaction hubs throughout the packaging process, coordinating the binding and release of large terminase and later tail proteins. The portal also interacts with the surrounding prohead and remains bound throughout the capsid maturation pathway. Such a variety of binding partners, and the inherent symmetry mismatches of the system, demands plasticity and flexibility of the portal protein. Indeed, asymmetric reconstructions of portal protein structures in situ show deviation from C_{12} symmetry. For instance, a recent structure of the PaP3 portal displays a ‘corkscrew’ architecture: employing helical rather than rotational

quaternary structure (Hou et al. 2022). Meanwhile, an asymmetric reconstruction of the T4 unique portal-containing vertex reveals significant structural variation within the portal protein N-terminus. This permits similar interactions with the capsid's coat protein subunits, which in turn display only minimal conformational deviation from each other (Fang et al. 2020). In situ structures of portals in Φ 29, P23-45 and P68 show poorly resolved N-terminal whiskers (Bayfield et al. 2019; Hrebík et al. 2019; Xu et al. 2019) which could implicate the N terminus more generally in this symmetry breaking function. The clip domain of the HK97 portal also shows deviation from C_{12} symmetry when in complex with large terminase (Hawkins et al. 2023).

The plasticity of portal proteins is utilised for signal transduction. In *pac* viruses, packaging termination is instigated in response to the increasing internal pressure as DNA fills the capsid. This pressure change is thought to be relayed to the large terminase through the portal protein (Lokareddy et al. 2017). For P22, comparative structures of prohead and mature phage portals reveal an increase in portal density and reduction in portal volume with the progression of DNA packaging. DNA is also spooled around the mature phage portal structure in a configuration that computational modelling deems incompatible with the procapsid portal form (Lokareddy et al. 2017). In corroboration of this suspected signal transducing function, several discrete mutations within the P22 portal core produce an overpacking phenotype, where portals remain in a more 'procapsid like' form. This in turn prevents termination (Bedwell and Prevelige 2017). The T7 portal protein also undergoes a striking structural transformation between its procapsid and mature forms. Within the mature head of T7, the portal is compressed outwards from the capsid centre, reducing its radial length by 12 Å and exposing the clip domain. This is facilitated by a 25° rigid body rotation of the wing domain (Chen et al. 2020).

The role of the portal as a DNA pressure sensor is similarly apparent in other ds DNA phage which do not employ a headful termination strategy. Mutations within the λ phage portal protein core have been shown to impede termination (Wieczorek et al. 2002). Furthermore, it has been proposed that HK97 relies on portal-mediated pressure changes for the release of its large terminase (Hawkins et al. 2023). Meanwhile, the mature Φ 29 portal protein shows a 16 Å reduction of the diameter of the clip domain and a comparable expansion at the top of the wing domain relative to the procapsid form (Xu et al. 2019).

The portal protein also plays a key role in preventing DNA leakage after packaging has been terminated. P22 mature head portal shows a dramatic decrease in the channel diameter from the procapsid form, from 40 Å to 25 Å (Olia et al. 2011; Lokareddy et al. 2017; Dedeo et al. 2019). This facilitated by a 10° increase in the angle of tunnel helices towards the central axis. A similar phenomenon is apparent for *Thermophages* P23-45, where the in situ prohead structure displays a constricted channel relative to the crystal structure (Bayfield et al. 2020). Here, channel loop conformations are inverted, reducing the diameter to just 14 Å and subsequently altering the nature of the channel interface from hydrophilic to hydrophobic. This is accompanied by positional adjustment of the crown domain (Bayfield et al. 2020). In addition to these broad structural rearrangements, portal proteins also display mechanisms to prevent DNA slippage during packaging. Pairs of charged residues

within the clip region (K200 and K209 for Φ 29 and K331 with K342 for SPP1) (Chaban et al. 2015) are thought to act as a DNA clamp. Single-particle studies on T4 also indicate that the portal prevents DNA release during motor slipping or stalling (Fuller et al. 2007a).

The Small Terminase

Most dsDNA phage employs small terminase proteins for recognition of the viral genome. The structure of numerous small terminases have been determined to reveal a central oligomerisation domain joining the N-terminal DNA-binding domain and C-terminal large terminase-binding domain (Hamada et al. 1986a, b; Casjens et al. 1987; Rao and Black 1988). Whether small terminase remains bound to large terminase throughout the packaging process remains unclear.

Historically, the nature of the small terminase DNA interaction at packaging initiation has proved a subject of debate. One school of thought favours the passage of DNA through the small terminase central channel, while another predicts DNA wrapping around the outside of the oligomeric ring. SPR and EPR data for Sf6 small terminase support this wrapping hypothesis, as binding between small terminase and DNA is shown to be weak but cooperative (Büttner et al. 2012). Indeed, this small terminase can bind the genome of related phage SPP1, and so specificity is thus thought to be determined more by DNA shape (i.e. intrinsically bent) rather than sequence. This is supported by the structures of p74-26 (Fig. 9.4a) and T4 like small terminase oligomers (Suna et al. 2012) which each show DNA binding HTH motifs arranged radially (Hayes et al. 2020). However, it has been argued that such bending of DNA is energetically unfavourable and a high-resolution structure of the P22 small terminase oligomeric ring features a sufficiently wide lumen for hydrated B-DNA (Fig. 9.4b) (Roy et al. 2011).

Recently, the first structure for a small terminase bound to a DNA substrate has been determined (Chechik et al. 2023). Here, the HK97 small terminase shows a nonameric ring encircling DNA, with the disordered N- and C-terminus of two

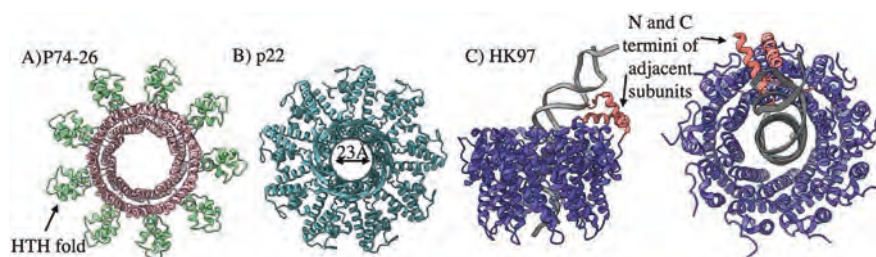


Fig. 9.4 Small terminase architecture. (a) P74-26 small terminase. PDB 6V1I (b) p22 small terminase. PDB 3P9A (c) The HK97 small terminase in complex with DNA. Two unstructured regions fold into helices on DNA binding. PDB 8POP

adjacent subunits folded into helices which form a DNA binding substructure (Fig. 9.4c). Two separate arginine residues, which are positioned within the major and minor DNA grooves, appear to stabilise the DNA in a bent conformation. The small terminase is thus locked in place in a sequence-specific manner, presumably allowing the ring structure to slide along the DNA substrate freely until it reaches the binding site. This is supported by the recent model of λ small terminase as a ‘sliding clamp’, limiting back slipping of large terminase (Rawson et al. 2023). While there is wide variation among other small terminase oligomeric states and reported structures, the N-terminus is rarely well resolved and always includes at least one positively charged residue. This suggests that folding of unstructured regions in response to specific DNA binding may be conserved. However, it is of note that the isolated DNA binding domain of some small terminases can be expressed, while others are entirely unknown, and that some dsDNA phage do not employ small terminases at all.

The Large Terminase

Translocation of DNA into the capsid is powered by the large terminase consisting of an N-terminal ATPase domain (Morita et al. 1993) adjoined to a C-terminal endonuclease domain by a flexible linker (Fig. 9.5). The terminase functions akin to other ring-shaped, oligomeric translocases: utilising ATP hydrolysis cycles to

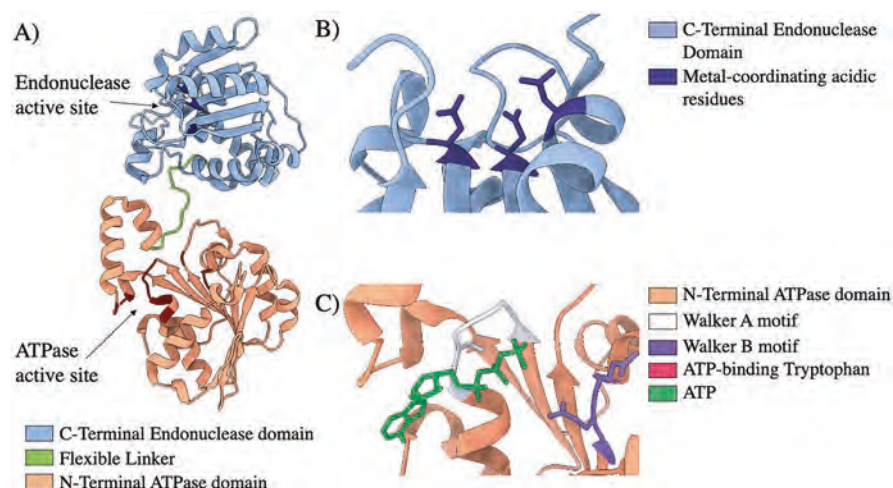


Fig. 9.5 Large Terminase Structure (a) Ribbon diagram of bacteriophage Sf6 large terminase, PDB code 4IFE. (b) The nuclease active site of Sf6, large terminase PDB code 4IFE. (c) The ATPase active site of the Sf6 large terminase coordinating ATP, PDB code 4IFE

translocate biological monomers through a central pore (Lyubimov et al. 2011). Once a complete genome is packaged, the large terminase nuclease domain is responsible for DNA cleavage.

Terminases belong to the additional strand, conserved glutamate (ASCE) subset of P-loop NTPases, which employ conserved Walker A and B motifs for ATP binding and hydrolysis (Iyer et al. 2004a; Ogura et al. 2004; delToro et al. 2019). The ASCE family utilises a second conserved acidic residue within the Walker B domain and a β -strand inserted between the Walker A and B domains (Leipe et al. 2003; Iyer et al. 2004b). The β - and γ -phosphates of ATP are coordinated by the Walker A domain, while the Walker B motifs coordinates Mg^{2+} (Walker et al. 1982). This facilitates the conserved catalytic carboxylate to activate water for ATP hydrolysis, which in turn instigates conformational change to translocate the substrate (Kenniston et al. 2003; Hanson and Whiteheart 2005). Each hydrolysis event instigates a subsequent ATP hydrolysis event in the neighbouring subunit (Moffitt et al. 2009; Mao et al. 2016).

The C-terminal domain of large terminases strongly resembles each other (Sun et al. 2008; Smits et al. 2009; Nadal et al. 2010; Roy and Cingolani 2012; Zhao et al. 2015a; Xu et al. 2017b), adopting the RNase H-like fold and mechanism. This fold is shared by the C-terminus domain of HSV1 packaging protein UL15 (Adelman et al. 2001) and CMCV packaging protein UL89 (Nadal et al. 2010). The nuclease domain adopts a two-metal catalysis mechanism where each metal is coordinated by active site carboxylate groups (Nowotny et al. 2005). Metal A activates a coordinated water for nucleophilic attack, and metal B stabilises the transition state oxyanion leaving group. The structure of the *Thermus thermophilus* bacteriophage G20C nuclease domain indicates a Ruv-C type mechanism which requires Mn^{2+} , Mg^{2+} or Co^{2+} for functionality (Xu et al. 2017b). While this mechanism is widely considered to be adopted by all large terminase, two structures have inspired conflicting ideas. The Sf6 nuclease structure suggests the two metals ions are brought unusually close (2.42 Å) during catalysis (Zhao et al. 2015a). In stark contrast, the two magnesium ions identified in the recent *Pseudomonas* phage E217 large terminase structure appear overly separated for two metal catalyses. This suggests rearrangement of the active site is necessary for nuclease activity (Lokareddy et al. 2022).

Translocation Kinetics of the DNA Motor

The large terminase packaging motor inspires particular mechanistic intrigue, as a result of the extreme repulsive forces it must overcome. Within the prohead, DNA is packaged to a crystalline density, which represents a highly unfavourable

energetic state. Not only must the electrostatic repulsion of confining the negatively charged DNA molecule be accounted for but also the necessity for DNA bending and dehydration (Riemer and Bloomfield 1978; Purohit et al. 2005; Petrov and Harvey 2008). It is thus perhaps unsurprising that viral terminases represent the most powerful biological machines studied, reaching forces of up to 100 pN (Chemla et al. 2005; Fuller et al. 2007b; Rickgauer et al. 2008). In addition, the motor must remain stable throughout packaging, in spite of the symmetry mismatch between the portal and large terminase oligomers.

Much mechanistic understanding of viral DNA packaging motors has been drawn from single particle optical tweezer studies. Here, a DNA substrate is tethered to a microbead held in place by a laser beam (Ashkin et al. 1986; Fuller et al. 2006; Moffitt et al. 2008). When a small external force is applied, i.e. when the DNA is packaged, the force can be measured by the laser beam which applies an equal force in the opposite direction.

Single-particle studies of the $\Phi 29$ motor indicated that packaging proceeds through two alternating modes: the dwell phase and burst phase. During the dwell phase, ATP binds cooperatively to each subunit around the ATPase ring. During the burst phase, DNA translocation into the prohead occurs in four subsequent 2.5 bp steps corresponding to four ATP hydrolysis events (Fig. 9.6) (Moffitt et al. 2009). As packaging of the genome approaches completion, pausing and slipping of the motor occur more frequently, and the burst phase step size decreases to approximately 2.3 bp per hydrolysis event, while dwells lengthen (Liu et al. 2014).

Single-molecule studies have also been used to probe T4 and λ DNA packaging machines. For each phage, the rate of translocation, reaching an average of 700 bp/s for T4 (Fuller et al. 2007a), is proportional to the size of the genome. This means the total genome packaging time corresponds to roughly 2–3 minutes for each. Reduced packaging velocity, as internal pressure increased towards packaging completion, was also ubiquitous. However, in the case of bacteriophage λ , an early pressure peak followed by a drop was also observed after approximately 30% of the genome had been packaged. This pressure is considered responsible for prohead expansion,

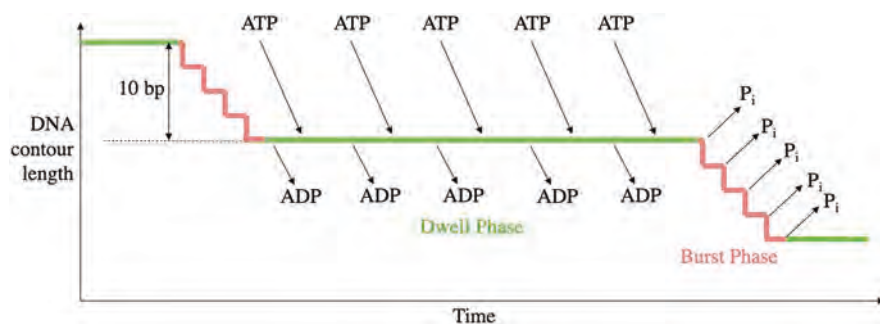


Fig. 9.6 Kinetic cycle of $\Phi 29$. Timeline of the burst/dwell cycle, adapted from Chistol et al. (2012). Five sequential ATP hydrolysis events produce just four translocation steps of 2.5 bp, followed by sequential ADP release and ATP binding

where the mature capsid displays a twofold increase in volume, relieving some internal pressure (Fuller et al. 2007b). For T4, packaging appears especially versatile, with reduced slipping and packaging into a mature capsid occurring with equal velocity to the immature prohead (Zhang et al. 2011). Single-molecule fluorescence studies have also revealed that multiple DNA substrates may be packaged into a single prohead in sequential packaging events (Vafabakhsh et al. 2014).

Coordination of the DNA Motor

Structural events within the motor require tight coordination in order to package DNA efficiently into the prohead. The first point of regulation is often the small terminase, which selectively recruits viral DNA to the large terminase motor, in preference to host DNA. Next, cycles of ATP hydrolysis by large terminase subunits must be coupled to DNA translocation. This requires not only intra-subunit coordination but also inter-subunit communication for the motor to function as a single entity. Lastly, timely termination of packaging requires strict regulation of the large terminase endonuclease domain. Endonuclease activity must remain inactive throughout packaging and be stimulated only after a full genome length has been packaged.

Small terminases have been shown to stimulate large terminase packaging activity (Al-Zahrani et al. 2009). For instance, binding of homologous T4 small terminase (gp16) to large terminase (gp17) is proposed to stimulate a conformational change which repositions residues within the large terminase ATPase catalytic pocket (Baumann and Black 2003; Al-Zahrani et al. 2009). Numerous sites of interaction between the two proteins have been identified, suggesting that multiple weak interactions could facilitate the rapid multimeric assembly and disassembly required for packaging. For bacteriophage P22, it has been deemed critical for small terminase to its cognate viral DNA in order to stimulate large terminase. This depicts an elegant way of discriminating against wasteful packaging of host DNA (Roy et al. 2012). Interestingly, the bacteriophage P74-26 small terminase has been shown to stimulate the ATPase activity of large terminase while simultaneously inhibiting the endonuclease activity (Hayes et al. 2020), presumably preventing the motor from premature termination.

Coordination of the ATP active site and DNA binding site within the large terminase monomer has largely been informed by crystallography. ATP bound large terminase subunits are broadly thought to adopt a ‘closed’ or ‘tense’ conformation with high DNA affinity. Hydrolysis of ATP causes movement between the two domains, coordinated by the linker or lid domain, as well as rearrangement of the DNA binding site. Now, in the ‘open’ conformation, large terminase releases DNA, facilitating its passage into the prohead. This model is supported by comparative crystal structures of monomeric terminases in apo, ADP-bound, and ATP analogue-bound conformations. In particular, significant conformational changes have been observed between ATP analogue-bound and apo structures of the P74-26 large

terminase, where an approximate 13° rotation occurs between the ATPase and lid subdomains (Hilbert et al. 2015).

Structures of T4, P74-26 and Sf6 large terminase monomers each reveal an additional conserved arginine in the Walker A motif. Molecular dynamics simulations indicate that this arginine may act as a ‘toggle’, switching coordination from the active site to a glutamate within the ATPase lid (or linker domain) on substrate release (Ortiz et al. 2019). A crystal structure of the Sf6 large terminase with bound ATP γ S shows critical arginine residue R24 coordinating the gamma phosphate. On ATP hydrolysis and P_i release, R24 instead interacts with and E187 in the lid domain, as shown in the ADP bound structure. This conformational change is proposed to be propagated to the bound DNA, which is subsequently ‘pushed’ into the prohead (Zhao et al. 2013).

In addition to this conserved ‘sensor’ arginine, MD simulations have indicated a single glutamate switch residue as critical to the transition between the ATP bound DNA tight binding state and the ADP bound state. In each of the four large terminases investigated, Sf6, ϕ 29, ascc- ϕ 28 and P74-26, a polar or charged switch residue appears to ‘fix’ the catalytic glutamate to point away from the active site rendering it inactive. The switch residue in turn relays structural rearrangement at the DNA binding site to promote binding when in the presence of ATP (Pajak et al. 2021a).

Subunit cooperativity is also critical for the motor function, with purified large terminase monomers displaying limited spontaneous ATPase activity, incompatible with demands of DNA packaging (Tafoya et al. 2018). The additional stimulation is thought to be provided by a transacting arginine finger, from a neighbouring subunit, which coordinates the ATP γ -phosphate (Fig. 9.7). Transacting arginine fingers have been identified for Φ 29, P74-26 and D6E, each of which is indispensable for

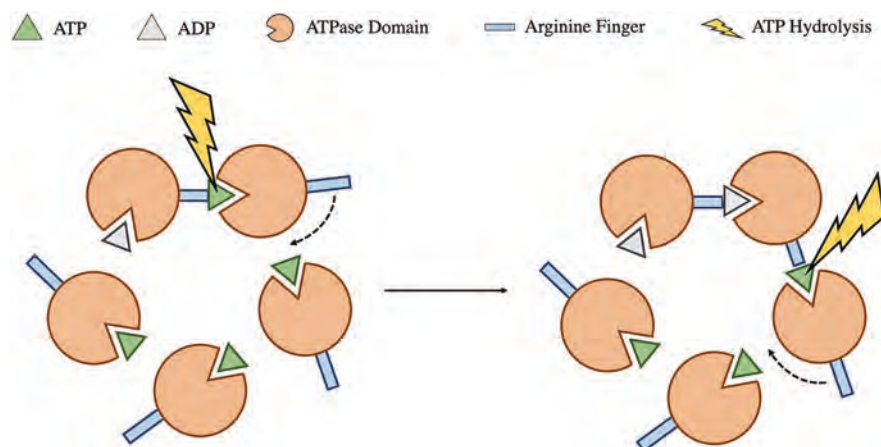


Fig. 9.7 Coordination of ATP hydrolysis. On ATP hydrolysis, conformational change within a single large terminase subunit repositions the transacting arginine finger into the adjacent ATP active site catalysing a subsequent hydrolysis event. This occurs sequentially around the ring

ATP hydrolysis (Mao et al. 2016; Hilbert et al. 2017; Xu et al. 2017a). The conformational change associated with ATP hydrolysis in one subunit is proposed to insert this residue into the active site of the adjacent subunit, producing a burst of sequential ATP hydrolysis events around the ring as catalogued by the $\Phi 29$ dwell burst translocation cycle. It is of note that, for $\Phi 29$ at least, the first hydrolysis event must occur without a trigger and also that only four of these five hydrolysis events produce translocation steps, indicating a unique role for a single subunit (Mao et al. 2016). The D6E arginine finger position is conserved among the FtsK/HerA superfamily, where the same mechanism is widely employed for coordination (Xu et al. 2017a). Unusually, extensive mutational studies on HK97 large terminase, in conjunction with the high-resolution crystal structure and a functional packaging assay, eliminate all potential transacting arginine finger candidates. Instead, K57 is proposed to act in cis, providing the additional charge for ATP coordination (Fung et al. 2022).

Regulation of the large terminase nuclease domain is critical during packaging to prevent early termination. Stimulation is thought to be modulated by the N-terminal ATPase domain. Indeed, in the absence of the ATPase domain, the isolated P22 endonuclease domain shows reduced or obliterated activity. An additional extended β -sheet and an auxiliary β -hairpin are seen in terminase nuclease families relative to the classical RNaseH-type fold. Since this β -hairpin is positioned to block dsDNA access to the nuclease active site, it has been proposed to play the role of the nuclease activity ‘switch’ (Smits et al. 2009; Roy and Cingolani 2012). For *cos* viruses such as HK97, small terminase is expected to play a role in positioning large terminase for DNA cleavage (Hawkins et al. 2023; Chechik et al. 2023). In *pac* viruses, however, the headful pressure from within the prohead is likely propagated through the portal protein to stimulate packaging termination, as discussed earlier.

Models for DNA Translocation from Cryo-EM Structures

Over the last decade, several structures of complete dsDNA viral packaging systems have been resolved by cryo-electron microscopy, in spite of the inherent challenges posed by the flexibility, asymmetry and transient association of these complexes. These reconstructions include active packaging systems for $\Phi 29$, T4 and HK97 (Fig. 9.8), and isolated packaging motors, in the absence of DNA or proheads, for T7 and HSV1. Such structures, in combination with the wealth of complimentary data from single-particle studies, molecular dynamics simulations and structures of individual protein constituents, have informed new understanding of DNA packaging mechanisms.

A high-resolution cryo-EM reconstruction of the intact $\Phi 29$ packaging motor, comprising the capsid, pRNA, ATPase and DNA, shows the five large terminase ATPases arranged in a ‘cracked’ helical conformation, stabilised by ATP γ S (Fig. 9.8a) (Woodson et al. 2021). This contrasts previously determined ‘planar’ assemblies of the $\Phi 29$ ATPase (Koti et al. 2008; Morais et al. 2008), as well as an

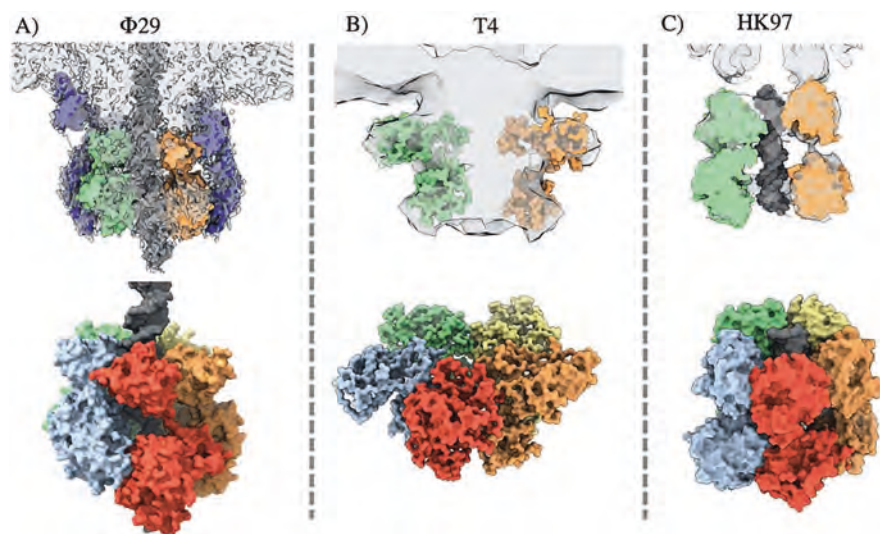


Fig. 9.8 Comparison of complete packaging complexes of ds DNA phage. Cryo-EM reconstructions are shown for (a) $\Phi 29$ EMDB 22441 fitted PDB 7JQQ, (b) T4 EMDB 1572 fitted PDB 3EZK, (c) HK97 EMDB 16653 fitted PDB 6Z6D

ADP-bound planar structure of the highly related phage $\text{asc}\phi 28$ large terminase (Pajak et al. 2021b). The necessary transition required between these two states has been modelled by molecular dynamics simulations and inspired a translocation model that agrees with the burst-dwell cycle (Pajak et al. 2021b; Woodson et al. 2021).

In the ATP-bound cracked helical conformation, all subunits tightly grip the DNA. When Subunit 1 hydrolyses ATP it releases DNA, and Subunit 2 is moved up into the Subunit 1 plane, moving approximately 2.5 bp of DNA into the prohead. Subunit 1 residue K105 is now poised to trigger ATP hydrolysis in Subunit 2, which in turn loses grip of the DNA and brings Subunit 3 into the Subunit 1 plane. Thus, four ATP hydrolysis steps occur, resulting in four 2.5 bp DNA translocation steps. When all five subunits are present in a planar ring, the pentamer has twice the buried surface area between subunits as the cracked helical state and thus represents a more stable state. Subunit 5 must then hydrolyse ATP to release the DNA substrate and prime S1 for nucleotide exchange and the beginning of the dwell phase. Here, ADP-ATP exchange again occurs sequentially, allowing each subunit in turn to move down the helix and make contact with the DNA (Pajak et al. 2021b; Woodson et al. 2021). This model is summarised in Fig. 9.9. DNA binding is energetically favourable, compensating for the less stable protein conformation, with decreased interaction interfaces between subunits. ATP hydrolysis is thus required to climb out of the energy minima in the transition to the planar state once again (Woodson et al. 2021).

Movement between the ATPase and nuclease domains is coordinated by a linker, folded into a 3-helix bundle, which also makes contact with the adjacent subunit.

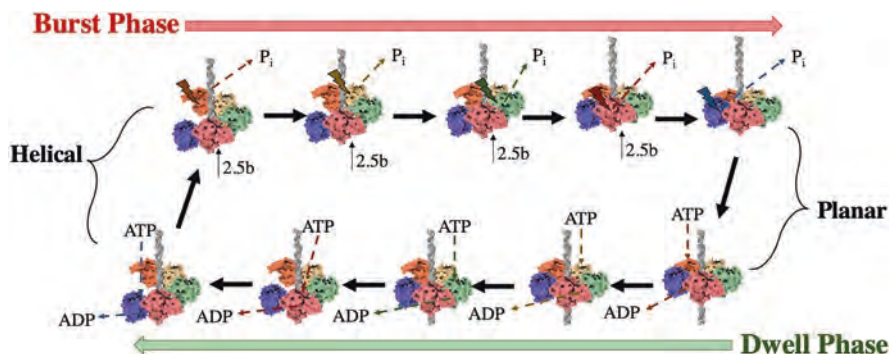


Fig. 9.9 Schematic of the $\Phi 29$ packaging mechanism. The ATPase domains of the large terminase pentameric ring sequentially transition from a cracked helix to planar arrangement in agreement with the dwell-burst kinetic cycle

Variations in twist and pitch of the helices allow subunits to adopt the different orientations required for translocation. The ATPase active site exists at the subunit periphery, where ATP appears to be sandwiched between R146 and K105 from the neighbouring (transacting) subunit. While all five subunits appear to contact DNA, K56 in S2–S5 is positioned to the track the 5′–3′ DNA strand. Meanwhile in S1, K56 is closer to the 3′–5′ strand, perhaps helping to distinguish a unique role for this subunit in initiating each burst cycle (Woodson et al. 2021).

For the bacteriophage T4, a crystal structure of the apo form of the large terminase and cryo-EM structure of a fully assembled packaging motor has also been derived (Fig. 9.8b) (Sun et al. 2008). The monomer depicts a ‘tense’, more compact state relative to a ‘relaxed’ extended state seen in the low-resolution cryo-EM reconstruction. The nuclease domain is thought to bind DNA in the relaxed state, moving it towards the ATPase domain. This was proposed to induce a conformational change which flips an arginine finger from the same subunit (cis) into the active site, catalysing ATP hydrolysis. In turn, the nuclease domain rotates aligning charge pairs producing a 2 bp movement of DNA via electrostatic attraction. This represents the tense state. Release of ATP hydrolysis products triggers the relaxation of the large terminase and positions DNA ready to bind the next subunit, so that at any point, just one of the five subunits adopts the tense state (Suna et al. 2012).

In place of the burst-dwell sequence adopted by the $\Phi 29$ motor, the T4 motor is proposed to operate in a continuous burst cycle. In this model, ATP hydrolysis in one large terminase subunit could occur simultaneously to ATP loading in the adjacent subunit, thus removing the ‘dwell’ phase and facilitating faster packaging (Rao et al. 2023). This could account for the eightfold increased speed of the T4 motor relative to $\Phi 29$ (Zhang et al. 2011). A more flexibly coordinated motor may also be responsible for the frequent slipping and restarting of the T4 motor during single-molecule studies and its ability to tolerate up to three inactive subunits (Kottadiel et al. 2012). Such a mechanism is also supported by enhanced DNA gripping capacity and prevention of slipping of ATP bound subunits (Ordyan et al. 2018).

More recently, a cryo-EM structure of the HK97 packaging motor has been determined in a pseudo termination state (Hawkins et al. 2023). The motor assembly is highly asymmetric, showing a ring of five large terminase monomers tilted against the portal axis at 10° . DNA runs through the channel and appears cleaved at the portal entrance in agreement with packaging termination. Interestingly, new biophysical and structural data for λ phage terminase instead suggests that DNA cleavage occurs in a tetrameric conformation, which welcomes in an additional subunit for packaging (Prokhorov et al. 2022).

Docking five monomers of HK97 large terminase (Fung et al. 2022) reveals variable degrees of extension between N- and C-terminal domains of individual subunits. More extended subunits make clear contact with DNA via both domains, as well as the portal clip domain, and are assumed ATP bound. Meanwhile more compact subunits only make contact with DNA via the C-terminal domain and do not appear portal bound. This suggests ATP-powered large terminase contraction may mediate portal contact in addition to DNA translocation. Such asymmetry also likely requires strict coordination between subunits, but interestingly the assembly does not adopt the cracked helix conformation of its $\Phi 29$ counterpart.

Two high-resolution structures of the HSV1 ds DNA packaging motor have also been determined by cryo-EM, depicting a ring of the large terminase homologue pUL15 subunits in (i) an apo and (ii) an ATP analogue-bound conformation (Yang et al. 2020). Regulatory proteins pUL28 and pUL33, which have no bacteriophage homologs, also feature in the structure, with each of the three proteins interdigitating into one unit of the compact hexameric conformation (Yang et al. 2020). The pUL15 endonuclease domain is positioned away from the centre, suggesting that headful packaging pressure must induce domain rearrangement for the nuclease to engage the DNA at packaging termination.

The hexameric conformation presented represents the major form when the three proteins are co-expressed and also mimics the oligomeric state favoured by other DNA translocases. This produces a central pUL15 ATPase channel, lined with conserved basic amino acids, whose diameter ($39 \text{ \AA}$) near matches the portal lumen ($36 \text{ \AA}$) (Wang et al. 2020). However, it is of note that these structures do not include the viral DNA or capsid and that a pentameric structure can theoretically be arranged from the pUL33-pUL28-pUL15 unit in which the central channel diameter varies between $19 \text{ \AA}$ and $24 \text{ \AA}$. In addition, essential packaging accessory proteins for both Epstein–Barr virus and Kaposi’s sarcoma-associated herpesvirus each self-associate into pentameric rings with positively charged DNA-binding central channels. This implicates association with the pentameric terminase machinery (Didychuk et al. 2021).

Despite variations, each of the dsDNA packaging structures shows common features which may prevail throughout dsDNA packaging machines. In each system, DNA propagation occurs via contraction of DNA-bound subunits, with cycling DNA binding and release coordinated by the ATP hydrolysis cycle. This is in keeping with other translocases such as the FtsK and TraB (Jing et al. 2016). For $\Phi 29$ and HK97, the large terminase nuclease domain is adjacent to the portal, although for $\Phi 29$, this interaction is mediated by pRNA. This orientation is also supported by

a low-resolution cryo-EM structure of the T7 large terminase-portal complex (Daudén et al. 2013) and FRET data on the T4 packaging motor (Dixit et al. 2013), although the orientation assigned by the low-resolution cryo-EM structure was opposing (Sun et al. 2008). As more and higher-resolution structures undoubtedly emerge, it will be of great interest as to the extent to which DNA packaging machines are conserved between phage and eukaryotic double-stranded DNA viruses.

Conclusion

In this chapter, we catalogue the current understanding of the molecular mechanisms of terminase-based DNA packaging systems of dsDNA viruses. Each utilises a portal protein, which exhibits conformational differences between prohead and mature capsid forms, with deviations from the C_{12} symmetry. This demonstrates the portal protein's plasticity, which is crucial for interactions with multiple binding partners and for propagation of the 'headful' pressure signal from within the prohead to large terminase. Portal protein also acts as gatekeeper to prevent DNA leakage before tail attachment. The mechanism by which small terminase recognises viral DNA and regulates large terminase remains more elusive. However, the recent cryo-EM structure of HK97 small terminase has shed light on a novel DNA-binding mechanism involving the folding of disordered protein regions on binding to the *cos* site. It is of great interest whether this mechanism is conserved among other *cos* phage and equally if it translates to *pac* viruses.

For large terminase too, cryo-EM structures have begun to tie together a wealth of data from crystallography, molecular dynamics, and single-particle studies. Five subunits encircle the DNA substrate, and coordination around the ATPase ring often occurs via a transacting arginine finger which coordinates the neighbouring active site. For $\Phi 29$, a high-resolution cryo-EM structure, in conjunction with the dwell-burst kinetic cycle derived from single-molecule data, has inspired a mechano-chemical translocation mechanism, describing sequential conformational changes in each ATPase subunit. Together these result in the transition of the pentameric ATPase from a helical to planar configuration and the propagation of DNA into the prohead. While lower-resolution data from other HK97 and T4 packaging motors do not strictly support this model, they each support a subunit contraction driven model. In this way, DNA tethered to an extended ATP-coordinated subunit is propelled into the prohead on ATP hydrolysis, and subsequently released, allowing the process to iterate within an adjacent subunit. This mechanism aligns with other translocases such as FtsK and TraB.

References

Ackermann HW (2007) 5500 Phages examined in the electron microscope. *Arch Virol* 152:227–243

- Adams MB, Hayden M, Casjens S (1983) On the sequential packaging of bacteriophage P22 DNA. *J Virol* 46:673–677. <https://doi.org/10.1128/jvi.46.2.673-677.1983>
- Adelman K, Salmon B, Baines JD (2001) Herpes simplex virus DNA packaging sequences adopt novel structures that are specifically recognized by a component of the cleavage and packaging machinery. *Proc Natl Acad Sci U S A* 98:3086–3091. <https://doi.org/10.1073/pnas.061555698>
- Agirrezabala X, Martín-Benito J, Valle M et al (2005) Structure of the connector of bacteriophage T7 at 8 Å resolution: Structural homologies of a basic component of a DNA translocating machinery. *J Mol Biol* 347:895–902. <https://doi.org/10.1016/j.jmb.2005.02.005>
- Al-Zahrani AS, Kondabagil K, Gao S et al (2009) The small terminase, gp16, of bacteriophage T4 is a regulator of the DNA packaging motor. *J Biol Chem* 284:24490–24500. <https://doi.org/10.1074/jbc.M109.025007>
- Ashkin A, Dziedzic JM, Bjorkholm JE, Chu S (1986) Observation of a single-beam gradient force optical trap for dielectric particles. *Opt Lett*, pp 288–290
- Baumann RG, Black LW (2003) Isolation and characterization of T4 bacteriophage gp17 terminase, a large subunit multimer with enhanced ATPase activity. *J Biol Chem* 278:4618–4627. <https://doi.org/10.1074/jbc.M208574200>
- Bayfield OW, Klimuk E, Winkler DC et al (2019) Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids. *Proc Natl Acad Sci U S A* 116:3556–3561. <https://doi.org/10.1073/pnas.1813204116>
- Bayfield OW, Steven AC, Antson AA (2020) Cryo-EM structure in situ reveals a molecular switch that safeguards virus against genome loss. *Elife* 14(9):e55517. <https://doi.org/10.7554/eLife.55517>
- Bedwell GJ, Prevelige PE (2017) Targeted mutagenesis of the P22 portal protein reveals the mechanism of signal transmission during DNA packaging. *Virology* 505:127–138. <https://doi.org/10.1016/j.virol.2017.02.019>
- Black LW (1989) DNA packaging in dsDNA bacteriophages. *Annu Rev Microbiol* 43:267–292
- Bogner E, Radsak K, Stinski MF (1998) The gene product of human cytomegalovirus open reading frame UL56 binds the pac motif and has specific nuclease activity. *J Virol* 72:2259–2264. <https://doi.org/10.1128/jvi.72.3.2259-2264.1998>
- Burroughs A, Iyer L, Aravind L (2007) Comparative genomics and evolutionary trajectories of viral ATP dependent DNA-packaging systems. *Genome Dyn* 3:48–65
- Büttner CR, Chechik M, Ortiz-Lombardía M et al (2012) Structural basis for DNA recognition and loading into a viral packaging motor. *Proc Natl Acad Sci U S A* 109:811–816. <https://doi.org/10.1073/pnas.1110270109>
- Camacho A, Jiménez F, Viñuela E, Salas M (1979) Order of assembly of the lower collar and the tail proteins of *Bacillus subtilis* bacteriophage phi 29. *J Virol* 29:540–545. <https://doi.org/10.1128/jvi.29.2.540-545.1979>
- Casjens S, Huang WM, Hayden M, Parr R (1987) Initiation of bacteriophage P22 DNA packaging series. Analysis of a mutant that alters the DNA target specificity of the packaging apparatus. *J Mol Biol* 194:411–422. [https://doi.org/10.1016/0022-2836\(87\)90671-1](https://doi.org/10.1016/0022-2836(87)90671-1)
- Chaban Y, Lurz R, Brasilès S et al (2015) Structural rearrangements in the phage head-to-tail interface during assembly and infection. *Proc Natl Acad Sci U S A* 112:7009–7014. <https://doi.org/10.1073/pnas.1504039112>
- Chai S, Lurz R, Alonso JC (1995) The small subunit of the terminase enzyme of *Bacillus subtilis* bacteriophage SPP1 forms a specialized nucleoprotein complex with the packaging initiation region. *J Mol Biol* 252:386–398. <https://doi.org/10.1006/jmbi.1995.0505>
- Chechik M, Greive SJ, Antson AA, Jenkins HT (2023) Structure of HK97 small terminase:DNA complex unveils a novel DNA binding mechanism by a circular protein. *bioRxiv*. <https://doi.org/10.1101/2023.07.17.549218>
- Chelikani V, Ranjan T, Kondabagil K (2014) Revisiting the genome packaging in viruses with lessons from the “giants”. *Virology* 466–467:15–26
- Chemla YR, Aathavan K, Michaelis J et al (2005) Mechanism of force generation of a viral DNA packaging motor. *Cell* 122:683–692. <https://doi.org/10.1016/j.cell.2005.06.024>

- Chen W, Xiao H, Wang X et al (2020) Structural changes of a bacteriophage upon DNA packaging and maturation. *Protein Cell* 11:374–379
- Chistol G, Liu S, Hetherington CL et al (2012) High degree of coordination and division of labor among subunits in a homomeric ring ATPase. *Cell* 151:1017–1028. <https://doi.org/10.1016/j.cell.2012.10.031>
- Cue D, Feiss M (1993) A site required for termination of packaging of the phage λ chromosome. *Proc Natl Acad Sci U S A* 90:9290–9294. <https://doi.org/10.1073/pnas.90.20.9290>
- Dai L, Singh D, Lu S et al (2021) A viral genome packaging ring-ATPase is a flexibly coordinated pentamer. *Nat Commun* 12(1):6548. <https://doi.org/10.1038/s41467-021-26800-z>
- Daudén MI, Martín-Benito J, Sánchez-Ferrero JC et al (2013) Large terminase conformational change induced by connector binding in bacteriophage T7. *J Biol Chem* 288:16998–17007. <https://doi.org/10.1074/jbc.M112.448951>
- Dedeo CL, Cingolani G, Teschke CM (2019) Portal protein: the orchestrator of capsid assembly for the dsDNA tailed bacteriophages and herpesviruses. *Annu Rev Virol* 6:141–160. <https://doi.org/10.1146/annurev-virology-092818-015819>
- delToro D, Ortiz D, Ordyan M et al (2019) Functional dissection of a viral DNA packaging machine's Walker B Motif. *J Mol Biol* 431:4455–4474. <https://doi.org/10.1016/j.jmb.2019.08.012>
- Didychuk AL, Gates SN, Gardner MR et al (2021) A pentameric protein ring with novel architecture is required for herpesviral packaging. *Elife* 10. <https://doi.org/10.7554/eLife.62261>
- Ding F, Lu C, Zhao W et al (2011) Structure and assembly of the essential RNA ring component of a viral DNA packaging motor. *Proc Natl Acad Sci U S A* 108:7357–7362. <https://doi.org/10.1073/pnas.1016690108>
- Dixit AB, Ray K, Thomas JA, Black LW (2013) The C-terminal domain of the bacteriophage T4 terminase docks on the prohead portal clip region during DNA packaging. *Virology* 446:293–302. <https://doi.org/10.1016/j.virol.2013.07.011>
- Doan DNP, Dokland T (2007) The gpQ portal protein of bacteriophage P2 forms dodecameric connectors in crystals. *J Struct Biol* 157:432–436. <https://doi.org/10.1016/j.jsb.2006.08.009>
- Dröge A, Santos MA, Stiege AC et al (2000) Shape and DNA packaging activity of bacteriophage SPP1 procapsid: protein components and interactions during assembly. *J Mol Biol* 296:117–132. <https://doi.org/10.1006/jmbi.1999.3450>
- Earnshaw WC, Casjens SR (1980) DNA packaging by the double-stranded DNA bacteriophages. *Cell* 21:319–331
- Fang Q, Tang WC, Tao P et al (2020) Structural morphing in a symmetry-mismatched viral vertex. *Nat Commun* 11:1703. <https://doi.org/10.1038/s41467-020-15575-4>
- Feiss M, Catalano CE (2007) Bacteriophage lambda terminase and the mechanism of viral DNA packaging. In: *Viral genome packaging machines: genetics, structure, and mechanism*. Kluwer Academic/Plenum Publishers, New York, pp 2133–2141
- Feiss M, Fisher RA, Crayton MA, Egner C (1977) Packaging of the bacteriophage λ chromosome: effect of chromosome length. *Virology* 77:281–293. [https://doi.org/10.1016/0042-6822\(77\)90425-1](https://doi.org/10.1016/0042-6822(77)90425-1)
- Feiss M, Kobayashi I, Widner W (1983a) Separate sites for binding and nicking of bacteriophage λ DNA by terminase. *Proc Natl Acad Sci U S A* 80:955–959. <https://doi.org/10.1073/pnas.80.4.955>
- Feiss M, Widner W, Miller G et al (1983b) Structure of the bacteriophage lambda cohesive end site: location of the sites of terminase binding (cosB) and nicking (cosN). *Gene* 24:207–218. [https://doi.org/10.1016/0378-1119\(83\)90081-1](https://doi.org/10.1016/0378-1119(83)90081-1)
- Fuller DN, Gemmen GJ, Rickgauer JP et al (2006) A general method for manipulating DNA sequences from any organism with optical tweezers. *Nucleic Acids Res* 34(2):15. <https://doi.org/10.1093/nar/gnj016>
- Fuller DN, Raymer DM, Kottadiel VI et al (2007a) Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. *Proc Natl Acad Sci U S A* 104:16868–16873. <https://doi.org/10.1073/pnas.0704008104>

- Fuller DN, Raymer DM, Rickgauer JP et al (2007b) Measurements of single DNA molecule packaging dynamics in bacteriophage λ reveal high forces, high motor processivity, and capsid transformations. *J Mol Biol* 373:1113–1122. <https://doi.org/10.1016/j.jmb.2007.09.011>
- Fung HKH, Grimes S, Huet A et al (2022) Structural basis of DNA packaging by a ring-type ATPase from an archetypal viral system. *Nucleic Acids Res* 50:8719–8732. <https://doi.org/10.1093/nar/gkac647>
- Guo P, Peterson C, Anderson D (1987) Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage ϕ 29. *J Mol Biol* 197:229–236. [https://doi.org/10.1016/0022-2836\(87\)90121-5](https://doi.org/10.1016/0022-2836(87)90121-5)
- Guo P, Erickso S, Xu W et al (1991) Regulation of the phage ϕ 29 prohead shape and size by the portal vertex. *Virology* 183:366–373. [https://doi.org/10.1016/0042-6822\(91\)90149-6](https://doi.org/10.1016/0042-6822(91)90149-6)
- Hamada K, Fujisawa H, Minagawa T (1986a) A defined in vitro system for packaging of bacteriophage T3 DNA. *Virology* 151:119–123. [https://doi.org/10.1016/0042-6822\(86\)90109-1](https://doi.org/10.1016/0042-6822(86)90109-1)
- Hamada K, Fujisawa H, Minagawa T (1986b) Overproduction and purification of the products of bacteriophage T3 genes 18 and 19, two genes involved in DNA packaging. *Virology* 151:110–118. [https://doi.org/10.1016/0042-6822\(86\)90108-X](https://doi.org/10.1016/0042-6822(86)90108-X)
- Hanson PI, Whiteheart SW (2005) AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* 6:519–529
- Hawkins DEDP, Bayfield OW, Fung HKH et al (2023) Insights into a viral motor: the structure of the HK97 packaging termination assembly. *Nucleic Acids Res* 51:7025–7035. <https://doi.org/10.1093/nar/gkad480>
- Hayes JA, Hilbert BJ, Gaubitz C et al (2020) A thermophilic phage uses a small terminase protein with a fixed helix-turn-helix geometry. *J Biol Chem* 295:3783–3793. <https://doi.org/10.1074/jbc.RA119.012224>
- Heming JD, Conway JF, Homa FL (2017) Herpesvirus capsid assembly and DNA packaging. *Adv Anat Embryol Cell Biol* 223:119–142
- Hendrix RW (1998) Bacteriophage DNA packaging: RNA gears in a DNA transport machine. *Cell* 94:147–150
- Hilbert BJ, Hayes JA, Stone NP et al (2015) Structure and mechanism of the ATPase that powers viral genome packaging. *Proc Natl Acad Sci U S A* 112(29):E3792–E3799. <https://doi.org/10.1073/pnas.1506951112>
- Hilbert BJ, Hayes JA, Stone NP et al (2017) The large terminase DNA packaging motor grips DNA with its ATPase domain for cleavage by the flexible nuclease domain. *Nucleic Acids Res* 45:3591–3605. <https://doi.org/10.1093/nar/gkw1356>
- Hohn B (1983) DNA sequences necessary for packaging of bacteriophage λ DNA. *Proc Natl Acad Sci U S A* 80:7456–7460. <https://doi.org/10.1073/pnas.80.24.7456>
- Hou DC-F, Swanson NA, Li F et al (2022) Cryo-EM structure of a kinetically trapped dodecameric portal protein from the pseudomonas-phage PaP3. *J Mol Biol* 434(9):167537. <https://doi.org/10.1016/j.jmb.2022.167537>
- Hřebík D, Štveráková D, Škubník K et al (2019) Structure and genome ejection mechanism of *Staphylococcus aureus* phage P68. *Sci Adv* 5(10):eaaw7414. <https://doi.org/10.1126/sciadv.aaw7414>
- Iyer LM, Leippe DD, Koonin EV, Aravind L (2004a) Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* 146(1–2):11–31
- Iyer LM, Makarova KS, Koonin EV, Aravind L (2004b) Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. *Nucleic Acids Res* 32:5260–5279
- Jackson EN, Jackson DA, Deans RJ (1978) EcoRI analysis of bacteriophage P22 DNA packaging. *J Mol Biol* 118:365–388. [https://doi.org/10.1016/0022-2836\(78\)90234-6](https://doi.org/10.1016/0022-2836(78)90234-6)
- Jing P, Burris B, Zhang R (2016) Forces from the portal govern the late-stage DNA transport in a viral DNA packaging nanomotor. *Biophys J* 111:162–177. <https://doi.org/10.1016/j.bpj.2016.05.040>

- Juhala RJ, Ford ME, Duda RL et al (2000) Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *J Mol Biol* 299:27–51. <https://doi.org/10.1006/jmbi.2000.3729>
- Kenniston JA, Baker TA, Fernandez JM, Sauer RT (2003) Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* 114:511–520. [https://doi.org/10.1016/S0092-8674\(03\)00612-3](https://doi.org/10.1016/S0092-8674(03)00612-3)
- Koti JS, Morais MC, Rajagopal R et al (2008) DNA packaging motor assembly intermediate of bacteriophage ϕ 29. *J Mol Biol* 381:1114–1132. <https://doi.org/10.1016/j.jmb.2008.04.034>
- Kottadiel VI, Rao VB, Chemla YR (2012) The dynamic pause-unpackaging state, an off-translocation recovery state of a DNA packaging motor from bacteriophage T4. *Proc Natl Acad Sci U S A* 109:20000–20005. <https://doi.org/10.1073/pnas.1209214109>
- Krupovic M, Koonin EV (2015) Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. *Nat Rev Microbiol* 13:105–115
- Lander GC, Khayat R, Li R et al (2009) The P22 tail machine at subnanometer resolution reveals the architecture of an infection conduit. *Structure* 17:789–799. <https://doi.org/10.1016/j.str.2009.04.006>
- Leipe DD, Koonin EV, Aravind L (2003) Evolution and classification of P-loop kinases and related proteins. *J Mol Biol* 333:781–815. <https://doi.org/10.1016/j.jmb.2003.08.040>
- Lin H, Black LW (1998) DNA requirements in vivo for phage T4 packaging. *Virology* 242:118–127. <https://doi.org/10.1006/viro.1997.9019>
- Liu S, Chistol G, Hetherington CL et al (2014) A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills. *Cell* 157:702–713. <https://doi.org/10.1016/j.cell.2014.02.034>
- Lokareddy RK, Sankhala RS, Roy A et al (2017) Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral capsid maturation. *Nat Commun* 8:14310. <https://doi.org/10.1038/ncomms14310>
- Lokareddy RK, Hou CFD, Doll SG et al (2022) Terminase subunits from the pseudomonas-phage E217. *J Mol Biol* 434(20):167799. <https://doi.org/10.1016/j.jmb.2022.167799>
- Lyubimov AY, Strycharska M, Berger JM (2011) The nuts and bolts of ring-translocase structure and mechanism. *Curr Opin Struct Biol* 21:240–248
- Mao H, Saha M, Reyes-Aldrete E et al (2016) Structural and molecular basis for coordination in a viral DNA packaging motor. *Cell Rep* 14:2017–2029. <https://doi.org/10.1016/j.celrep.2016.01.058>
- Migliori AD, Keller N, Alam TI et al (2014) Evidence for an electrostatic mechanism of force generation by the bacteriophage T4 DNA packaging motor. *Nat Commun* 5:6548. <https://doi.org/10.1038/ncomms5173>
- Moffitt JR, Chemla YR, Smith SB, Bustamante C (2008) Recent advances in optical tweezers. *Annu Rev Biochem* 77:205–2028
- Moffitt JR, Chemla YR, Aathavan K et al (2009) Intersubunit coordination in a homomeric ring ATPase. *Nature* 457:446–450. <https://doi.org/10.1038/nature07637>
- Morais MC, Koti JS, Bowman VD et al (2008) Defining molecular and domain boundaries in the bacteriophage ϕ 29 DNA packaging motor. *Structure* 16:1267–1274. <https://doi.org/10.1016/j.str.2008.05.010>
- Morita M, Tasaka M, Fujisawa H (1993) DNA packaging ATPase of bacteriophage T3. *Virology* 193:748–752. <https://doi.org/10.1006/viro.1993.1183>
- Murialdo H, Becker A (1978) Head morphogenesis of complex double-stranded deoxyribonucleic acid bacteriophages. *Microbiol Rev* 42:529–576
- Nadal M, Mas PJ, Blanco AG et al (2010) Structure and inhibition of herpesvirus DNA packaging terminase nuclease domain. *Proc Natl Acad Sci U S A* 107:16078–16083. <https://doi.org/10.1073/pnas.1007144107>
- Nowotny M, Gaidamakov SA, Crouch RJ, Yang W (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* 121:1005–1016. <https://doi.org/10.1016/j.cell.2005.04.024>

- Ogura T, Whiteheart SW, Wilkinson AJ (2004) Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. *J Struct Biol* 146(1–2):106–112
- Olia AS, Prevelige PE, Johnson JE, Cingolani G (2011) Three-dimensional structure of a viral genome-delivery portal vertex. *Nat Struct Mol Biol* 18:597–603. <https://doi.org/10.1038/nsmb.2023>
- Oliveira L, Tavares P, Alonso JC (2013) Headful DNA packaging: Bacteriophage SPPI as a model system. *Virus Res* 173:247–259
- Ordyan M, Alam I, Mahalingam M et al (2018) Nucleotide-dependent DNA gripping and an end-clamp mechanism regulate the bacteriophage T4 viral packaging motor. *Nat Commun* 9. <https://doi.org/10.1038/s41467-018-07834-2>
- Ortiz D, DelToro D, Ordyan M et al (2019) Evidence that a catalytic glutamate and an “Arginine Toggle” act in concert to mediate ATP hydrolysis and mechanochemical coupling in a viral DNA packaging motor. *Nucleic Acids Res* 47:1404–1415. <https://doi.org/10.1093/nar/gky1217>
- Pajak J, Atz R, Hilbert BJ et al (2021a) Viral packaging ATPases utilize a glutamate switch to couple ATPase activity and DNA translocation. *Proc Natl Acad Sci U S A* 118(17):e2024928118. <https://doi.org/10.1073/pnas.2024928118>
- Pajak J, Dill E, Reyes-Aldrete E et al (2021b) Atomistic basis of force generation, translocation, and coordination in a viral genome packaging motor. *Nucleic Acids Res* 49:6474–6488. <https://doi.org/10.1093/nar/gkab372>
- Peixuan G, Erickson S, Anderson D (1987) A small viral RNA is required for in vitro packaging of bacteriophage ϕ 29 DNA. *Science* 236:690–694. <https://doi.org/10.1126/science.3107124>
- Petrov AS, Harvey SC (2008) Packaging double-helical DNA into viral capsids: Structures, forces, and energetics. *Biophys J* 95:497–502
- Prokhorov NS, Davis CR, Maruthi K et al (2022) Biophysical and structural characterization of a viral genome packaging motor. *bioRxiv*. <https://doi.org/10.1101/2022.09.25.509378>
- Purohit PK, Inamdar MM, Grayson PD et al (2005) Forces during bacteriophage DNA packaging and ejection. *Biophys J* 88. <https://doi.org/10.1529/biophysj.104.047134>
- Rao VB, Black LW (1988) Cloning, overexpression and purification of the terminase proteins gp16 and gp17 of bacteriophage T4. Construction of a defined in-vitro DNA packaging system using purified terminase proteins. *J Mol Biol* 200:475–488. [https://doi.org/10.1016/0022-2836\(88\)90537-2](https://doi.org/10.1016/0022-2836(88)90537-2)
- Rao VB, Fokine A, Fang Q, Shao Q (2023) Bacteriophage T4 head: structure, assembly, and genome packaging. *Viruses* 15(2):527
- Rawson B, Ordyan M, Yang Q et al (2023) Regulation of phage lambda packaging motor-DNA interactions: nucleotide independent and dependent gripping and friction. *bioRxiv*. <https://doi.org/10.1101/2022.09.24.509349>
- Reid RJD, Bodley JW, Anderson D (1994) Characterization of the prohead-pRNA interaction of bacteriophage ϕ 29. *J Biol Chem* 269:5157–5162. [https://doi.org/10.1016/s0021-9258\(17\)37669-x](https://doi.org/10.1016/s0021-9258(17)37669-x)
- Reyes-Aldrete E, Dill EA, Bussetta C et al (2021) Biochemical and biophysical characterization of the dsDNA packaging motor from the *Lactococcus lactis* bacteriophage AscCphi28. *Viruses* 13(1):15. <https://doi.org/10.3390/v13010015>
- Rickgauer JP, Fuller DN, Grimes S et al (2008) Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage ϕ 29. *Biophys J* 94:159–167. <https://doi.org/10.1529/biophysj.107.104612>
- Riemer SC, Bloomfield VA (1978) Packaging of DNA in bacteriophage heads: some considerations on energetics. *Biopolymers* 17:785–794. <https://doi.org/10.1002/bip.1978.360170317>
- Rossmann MG, Mesyanzhinov VV, Arisaka F, Leiman PG (2004) The bacteriophage T4 DNA injection machine. *Curr Opin Struct Biol* 14:171–180
- Roy A, Cingolani G (2012) Structure of P22 headful packaging nuclease. *J Biol Chem* 287:28196–28205. <https://doi.org/10.1074/jbc.M112.349894>

- Roy A, Bhardwaj A, Cingolani G (2011) Crystallization of the nonameric small terminase subunit of bacteriophage P22. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 67:104–110. <https://doi.org/10.1107/S174430911004697X>
- Roy A, Bhardwaj A, Datta P et al (2012) Small terminase couples viral DNA binding to genome-packaging ATPase activity. *Structure* 20:1403–1413. <https://doi.org/10.1016/j.str.2012.05.014>
- Scheffczik H, Savva CGW, Holzenburg A et al (2002) The terminase subunits pUL56 and pUL89 of human cytomegalovirus are DNA-metabolizing proteins with toroidal structure. *Nucleic Acids Res* 30:1695–1703
- Smits C, Chechik M, Kovalevskiy OV et al (2009) Structural basis for the nuclease activity of a bacteriophage large terminase. *EMBO Rep* 10:592–598. <https://doi.org/10.1038/embor.2009.53>
- Sternberg N, Coulby J (1990) Cleavage of the bacteriophage P1 packaging site (pac) is regulated by adenine methylation. *Proc Natl Acad Sci U S A* 87:8070–8074. <https://doi.org/10.1073/pnas.87.20.8070>
- Sun S, Kondabagil K, Draper B et al (2008) The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell* 135:1251–1262. <https://doi.org/10.1016/j.cell.2008.11.015>
- Suna S, Gao S, Kondabagil K et al (2012) Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. *Proc Natl Acad Sci U S A* 109:817–822. <https://doi.org/10.1073/pnas.1110224109>
- Tafoya S, Liu S, Castillo JP et al (2018) Molecular switch-like regulation enables global subunit coordination in a viral ring ATPase. *Proc Natl Acad Sci U S A* 115:7961–7966. <https://doi.org/10.1073/pnas.1802736115>
- Tang J, Lander GC, Olia A et al (2011) Peering down the barrel of a bacteriophage portal: the genome packaging and release valve in P22. *Structure* 19:496–502. <https://doi.org/10.1016/j.str.2011.02.010>
- Tavares P, Lurz R, Stiege A et al (1996) Sequential headful packaging and fate of the cleaved DNA ends in bacteriophage SPP1. *J Mol Biol* 264:954–967. <https://doi.org/10.1006/jmbi.1996.0689>
- TheiB J, Sung MW, Holzenburg A, Bogner E (2019) Full-length human cytomegalovirus terminase pUL89 adopts a two-domain structure specific for DNA packaging. *PLoS Pathog* 15(12):e1008175. <https://doi.org/10.1371/journal.ppat.1008175>
- Tu AHT, Voelker LRL, Shen X, Dybvig K (2001) Complete nucleotide sequence of the mycoplasma virus P1 genome. *Plasmid* 45:122–126. <https://doi.org/10.1006/plas.2000.1501>
- Tye BK, Huberman JA, Botstein D (1974) Non-random circular permutation of phage P22 DNA. *J Mol Biol* 85:501–528. [https://doi.org/10.1016/0022-2836\(74\)90312-X](https://doi.org/10.1016/0022-2836(74)90312-X)
- Vafabakhsh R, Kondabagil K, Earnest T et al (2014) Single-molecule packaging initiation in real time by a viral DNA packaging machine from bacteriophage T4. *Proc Natl Acad Sci U S A* 111:15096–15101. <https://doi.org/10.1073/pnas.1407235111>
- Valpuesta J, Fujisawa H, Marco S et al (1992) Three-dimensional structure of T3 connector purified from overexpressing bacteria. *J Mol Biol* 224:103–112. [https://doi.org/10.1016/0022-2836\(92\)90579-9](https://doi.org/10.1016/0022-2836(92)90579-9)
- Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 1:945–951. <https://doi.org/10.1002/j.1460-2075.1982.tb01276.x>
- Wang N, Chen W, Zhu L et al (2020) Structures of the portal vertex reveal essential protein-protein interactions for Herpesvirus assembly and maturation. *Protein Cell* 11(5):366–373
- Wieczorek DJ, Didion L, Feiss M (2002) Alterations of the portal protein, gpB, of bacteriophage λ suppress mutations in cosQ, the site required for termination of DNA packaging. *Genetics* 161:21–31. <https://doi.org/10.1093/genetics/161.1.21>
- Woodson M, Pajak J, Mahler BP et al (2021) A viral genome packaging motor transitions between cyclic and helical symmetry to translocate dsDNA. *Sci Adv* 7. <https://doi.org/10.1126/sciadv.abc1955>

- Xu RG, Jenkins HT, Antson AA, Greive SJ (2017a) Structure of the large terminase from a hyperthermophilic virus reveals a unique mechanism for oligomerization and ATP hydrolysis. *Nucleic Acids Res* 45:13029–13042. <https://doi.org/10.1093/nar/gkx947>
- Xu RG, Jenkins HT, Chechik M et al (2017b) Viral genome packaging terminase cleaves DNA using the canonical RuvC-like two-metal catalysis mechanism. *Nucleic Acids Res* 45:3580–3590. <https://doi.org/10.1093/nar/gkw1354>
- Xu J, Wang D, Gui M, Xiang Y (2019) Structural assembly of the tailed bacteriophage ϕ 29. *Nat Commun* 10:2366. <https://doi.org/10.1038/s41467-019-10272-3>
- Yang Y, Yang P, Wang N et al (2020) Architecture of the herpesvirus genome-packaging complex and implications for DNA translocation. *Protein Cell* 11:339–351. <https://doi.org/10.1007/s13238-020-00710-0>
- Zhang Z, Kottadiel VI, Vafabakhsh R et al (2011) A promiscuous DNA packaging machine from bacteriophage T4. *PLoS Biol* 9(2):e1000592. <https://doi.org/10.1371/journal.pbio.1000592>
- Zhao H, Christensen TE, Kamau YN, Tang L (2013) Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage. *Proc Natl Acad Sci U S A* 110:8075–8080. <https://doi.org/10.1073/pnas.1301133110>
- Zhao H, Lin Z, Lynn AY et al (2015a) Two distinct modes of metal ion binding in the nuclease active site of a viral DNA-packaging terminase: Insight into the two-metal-ion catalytic mechanism. *Nucleic Acids Res* 43:11003–11016. <https://doi.org/10.1093/nar/gkv1018>
- Zhao W, Jardine PJ, Grimes S (2015b) An RNA domain imparts specificity and selectivity to a viral DNA packaging motor. *J Virol* 89:12457–12466. <https://doi.org/10.1128/jvi.01895-15>
- Ziermann R, Calendar R (1990) Characterization of the cos sites of bacteriophages P2 and P4. *Gene* 96:9–15. [https://doi.org/10.1016/0378-1119\(90\)90334-N](https://doi.org/10.1016/0378-1119(90)90334-N)

Chapter 10

Human Transient Receptor Potential Ankyrin 1 Channel: Structure, Function, and Physiology



Viktorie Vlachova, Ivan Barvik, and Lucie Zimova

Abstract The transient receptor potential ion channel TRPA1 is a Ca^{2+} -permeable nonselective cation channel widely expressed in sensory neurons, but also in many nonneuronal tissues typically possessing barrier functions, such as the skin, joint synoviocytes, cornea, and the respiratory and intestinal tracts. Here, the primary role of TRPA1 is to detect potential danger stimuli that may threaten the tissue homeostasis and the health of the organism. The ability to directly recognize signals of different modalities, including chemical irritants, extreme temperatures, or osmotic changes resides in the characteristic properties of the ion channel protein complex. Recent advances in cryo-electron microscopy have provided an important framework for understanding the molecular basis of TRPA1 function and have suggested novel directions in the search for its pharmacological regulation. This chapter summarizes the current knowledge of human TRPA1 from a structural and functional perspective and discusses the complex allosteric mechanisms of activation and modulation that play important roles under physiological or pathophysiological conditions. In this context, major challenges for future research on TRPA1 are outlined.

Keywords TRPA1 channel · Gating · Sensory transduction · Nociception · Chemosensation · Thermosensation

V. Vlachova (✉) · L. Zimova (✉)

Department of Cellular Neurophysiology, Institute of Physiology of the Czech Academy of Sciences, Prague, Czech Republic
e-mail: viktorie.vlachova@fgu.cas.cz; lucie.zimova@fgu.cas.cz

I. Barvik (✉)

Division of Biomolecular Physics, Institute of Physics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic
e-mail: ibarvik@karlov.mff.cuni.cz

Introduction

The transient receptor potential ankyrin 1 (TRPA1) is an ion channel that functions as a molecular sensor or as part of various sensory molecular complexes involved in the detection of environmental and endogenous stimuli. It was first revealed as p120 protein in human fibroblasts and liposarcoma cells, and its role has been proposed in signal transduction and growth control (Jaquemar et al. 1999). As an ion channel, TRPA1 was first identified two decades ago, when it was revealed to be abundantly expressed in a specific subset of primary sensory neurons in mouse dorsal root ganglia and shown to respond to low temperature, below 17 °C (Story et al. 2003). Subsequently, the channel has been found to respond not only to cold, but also to a very broad spectrum of exogenous pungent and irritant chemicals, including allyl isothiocyanate, acrolein, allicin, and cinnamaldehyde (Bandell et al. 2004; Jordt et al. 2004; Bautista et al. 2005, 2006; Kwan et al. 2006). Important endogenous activators have been recognized to include hyperoxia and oxidative stress products (Takahashi et al. 2011; Andersson et al. 2008), changes in calcium homeostasis (Doerner et al. 2007; Zurborg et al. 2007; Wang et al. 2008b), lipids and their metabolites (Bandell et al. 2004), and low pH (de la Roche et al. 2013). In addition to nociceptive somatosensory neurons, TRPA1 is expressed in a wide range of non-neuronal cell types, such as endothelium of cerebral arteries (Earley et al. 2009), enterochromaffin and mast cells (Andrade et al. 2008; Nozawa et al. 2009; Oh et al. 2013), smooth muscle cells (Nassini et al. 2012; Mukhopadhyay et al. 2011), cardiomyocytes (Andrei et al. 2016), hair cells (Corey et al. 2004), astrocytes (Shigetomi et al. 2011), oligodendrocytes (Hamilton et al. 2016), Schwann cells (De Logu et al. 2017), odontoblasts (El Karim et al. 2011; Bernal et al. 2021), keratinocytes, melanocytes (Atoyan et al. 2009), chondrocytes (Nummenmaa et al. 2016), synoviocytes (Kochukov et al. 2006), and pancreatic beta cells (Cao et al. 2012), among others. In many of these cell types, TRPA1 generally functions as a multimodal molecular sensor of cellular threats, capable of responding to a variety of physical and chemical stimuli of exogenous or endogenous origin and enabling an appropriate defensive response of the organism (Fig. 10.1). Thus, the spectrum of roles in which TRPA1 is involved is broad, and knowledge of its function in terms of molecular organization is essential for a thorough understanding of these processes.

This chapter is structured as follows. In the first part, the main findings on the physiological roles of human TRPA1 that have been established so far based on experiments and clinical studies are presented. From the structural point of view, the ability of the channel to detect stimuli of different modalities and to convert them into pore opening lies in the specific arrangement of the channel molecular complex. Three fundamental structural characteristics underlie the physiological functions of TRPA1 as an ion channel: (a) a pore-forming region that ensures efficient and selective ion transfer through the pore and its opening and closing (i.e., gating); (b) the channel specifically binds activating ligands; and (c) it allosterically couples the binding to the pore opening. These three basic properties have already been explored in more detail and are described in the second part of this chapter. However, what distinguishes TRPA1 from conventional ligand-activated channels is that, in addition to ligand binding (d), it possesses several other regions that change conformation in

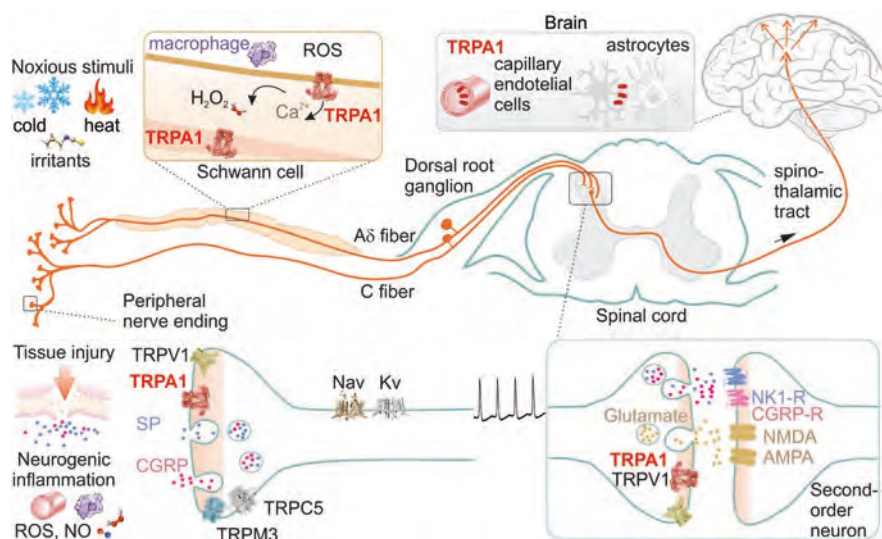


Fig. 10.1 Involvement of the transient receptor potential ankyrin 1 (TRPA1) channel in pain and neurogenic inflammation. TRPA1 is expressed in free nerve endings of primary afferent sensory neurons (unmyelinated C fibers and myelinated Aδ fibers). These neurons are the gateway by which sensory information about noxious stimuli is transduced and transmitted from the peripheral tissues to the spinal cord and brain. Additional thermosensitive TRP channels TRPV1, TRPM3, and TRPC5 are expressed together with TRPA1 in primary sensory neurons. In Schwann cells, macrophage-dependent oxidative stress (ROS; reactive oxygen species) induces TRPA1-mediated Ca²⁺ influx and peroxide (H₂O₂) release, which, in turn, activates neuronal TRPA1 (de Araujo et al. 2020). Primary afferent fibers project into the dorsal horn of the spinal cord; C fibers project superficially to laminae I and II, Aδ fibers project to laminae I and V. Sodium (Na_v) and potassium (K_v) channels are key in setting the resting membrane potential and the threshold for excitation of the neurons. In the spinal cord, the fibers make synaptic contacts with second-order neurons, which subsequently convey the information to the centers in the brain via glutamate receptors NMDA and AMPA. The TRPA1 activation elicits the release of inflammatory agents, including reactive oxygen species (ROS), nitric oxide (NO), and neuropeptides calcitonin-gene related peptide (CGRP) and substance P (SP) from peripheral fibers to produce neurogenic inflammation via CGRP receptor (CGRP-R), neurokinin receptor 1 (NK₁-R), and TRPA1. These mechanisms contribute to the generation of acute pain. In brain capillary endothelial cells, TRPA1 channels act as neuronal activity sensors that initiate microvascular vasodilatory responses to redirect blood to regions of metabolic demand (Thakore et al. 2021). The layout of the figure was inspired by an excellent review article by Bamps et al. (2021), to which we refer readers for further information

response to other chemicals or physical stimuli and also regulate pore domain conformation through allosteric coupling. The third part of the chapter focuses on these processes. Considering the channel's presumed function as a physiological cellular sensor, it is evident that the molecular complex must be a finely tuned structure that responds precisely to changes in the external environment and is only able to integrate signals when a threshold is crossed that may be dangerous to a living organism. It turns out that such a sophisticated mechanism has been subject to strong evolutionary pressure, and we can learn much about it from cross-species comparisons of channel structure and function. What is known about these aspects is the subject of the fourth part of this chapter. Since the main functions of TRPA1 are determined by

the context of native tissues, two separate sections are devoted to posttranslational modifications and the proteins identified so far that have been shown to interact structurally and functionally with the channel complex. The final section then addresses the main outstanding questions and future directions for TRPA1 research.

Physiological Functions of TRPA1 and Basic Properties

Physiological Functions and TRPA1-Related Human Disorders

So far, two hereditary channelopathies caused by mutations in the *TRPA1* gene have been identified in humans. The first was revealed in patients with a rare peripheral neuropathy disorder, autosomal dominant familial episodic pain syndrome, caused by the missense mutation N855S (Kremeyer et al. 2010). The syndrome is manifested by recurrent, intense pain in the upper body or lower limbs, triggered by fasting, fatigue, cold exposure, and/or physical stress. The residue N855 is located in a functionally important transmembrane region that is critical for stabilizing the conformations associated with the channel gating (Zima et al. 2015; Naert et al. 2020). Moreover, recent structural/functional studies have demonstrated that the temperature-sensitive properties of human TRPA1 are determined by its redox state and have suggested that the interaction of N855 with the backbone carbonyl of the proximal C-terminal cysteine 1025 may be one of the mechanisms underlying the syndrome (Moparathi et al. 2022). The second identified *TRPA1* channelopathy is caused by a premature stop codon at amino acid 919 located in the outer pore of the channel so that the protein lacks critical elements involved in the ion permeation, gating, and allosteric coupling (Nirenberg et al. 2018; Bali et al. 2023). This variant is associated with the cramp fasciculation syndrome which is a rare peripheral nerve disorder manifested by stiffness, muscle pain, cramps, fatigue, anxiety, hyperreflexia, and paresthesias. The individuals suffer also with other neuronal hyperexcitability-hypersensitivity symptoms including cold hyperalgesia, chronic itch, and asthma, which support the involvement of TRPA1 in these processes. Indeed, there is still growing evidence linking the physiological functions of human TRPA1 not only to various forms of pain (Marchi et al. 2023; Sleczkowska et al. 2022; Kim et al. 2006; Jhun et al. 2018) and temperature perception (Namer et al. 2005; Binder et al. 2011) but also to mechanosensation (Achenbach et al. 2019), respiratory functions (Deering-Rice et al. 2015; Gallo et al. 2017; Naumov et al. 2021), regulation of the cardiovascular system (Wang et al. 2019), itch (Kittaka and Tominaga 2017), and also the homeostatic balance between the immune and nociceptive systems (Naert et al. 2021; Viana 2016; Pereira et al. 2017).

In light of the above, it is not unexpected that there is sustained interest in TRPA1 as a potential target for the treatment of serious human disorders. Presently, the public database ClinicalTrials.gov provided by the US National Library of Medicine (accessed 18.04.2024) lists 19 entries for TRPA1 with various status indicators that focus on the role of the channel in surgical, thermal, and neuropathic pain, itch, dysphagia, cough, intestinal permeability, and chemotherapy-induced neuropathy

(identifiers, e.g., NCT05086289, NCT05177094, NCT03381144, NCT04529096). Effective TRPA1 inhibitors are being pursued by a number of companies including Eli Lilly and Company, Glenmark Pharmaceuticals, Bayer-Evotec Strategic Alliance, and Genentech (Mesch et al. 2023; Chen and Terrett 2020; Terrett et al. 2021; Chan et al. 2021; Balestrini et al. 2021; Hu et al. 2023). For example, a new oral TRPA1 antagonist ISC 17536 (formerly GRC 17536; Glenmark Pharmaceuticals), evaluated in a phase 2 study, shows promising efficacy against painful diabetic peripheral neuropathy in a subset of patients with preserved small nerve fiber function (Jain et al. 2022).

It is not the purpose of this chapter to list all the (patho)physiological functions of TRPA1 that have so far been experimentally demonstrated in animal models, so we refer here to several excellent review articles on this channel published in the last few years (de Araujo et al. 2020; Landini et al. 2022; Viana 2016; Gouin et al. 2017; Koivisto et al. 2018, 2022; Wang et al. 2019; Skerratt 2017; Bamps et al. 2021; Li et al. 2023). Of these, exceptionally comprehensive information about TRPA1 from all relevant aspects is provided in the recent review article by Talavera and co-workers (2020). These publications also detail and/or refer to a large spectrum of chemical structures of electrophilic and non-electrophilic compounds that induce or inhibit TRPA1 channel activity. Here, we detail only those important substances that are part of the already resolved 3D structures and for which there are some insights into how they structurally contribute to channel modulation.

Structural Characteristics

Transmembrane Domain and Cytoplasmic Regions

Until recently, the lack of high-resolution structural data hindered a comprehensive understanding of TRPA1 architecture and the mechanisms of its activation and modulation. The breakthrough came with the application of cryo-electron microscopy, which has revolutionized the field of structural biology. The structure of TRPA1 was elucidated shortly after the publication of the first TRP channel structure, TRPV1, in 2013 (Cao et al. 2013; Liao et al. 2013). In 2015, Paulsen et al. determined the structural details of human TRPA1 and confirmed that TRPA1 shares design principles with the rest of the TRP family and, in the transmembrane domain, with other tetrameric voltage-gated ion channels (Paulsen et al. 2015). Each subunit of TRPA1 is composed of six transmembrane α -helices (S1–S6) arranged in a domain-swapped configuration, where the S1–S4 voltage sensor-like domain (VSLD) of one subunit interacts with the pore domain of an adjacent subunit (Fig. 10.2). The pore domain is formed by the transmembrane helices S5 and S6 and a reentrant loop between them, which contains two pore helices P1 and P2. The central ion permeation pathway has two major constrictions called the upper and lower gates. The transmembrane part accounts for one-fifth of the ion channel mass, with the remainder consisting of intracellularly oriented C and N termini. The distinctive structural element of the intracellular region that TRPA1 shares with TRPC and TRPM channels

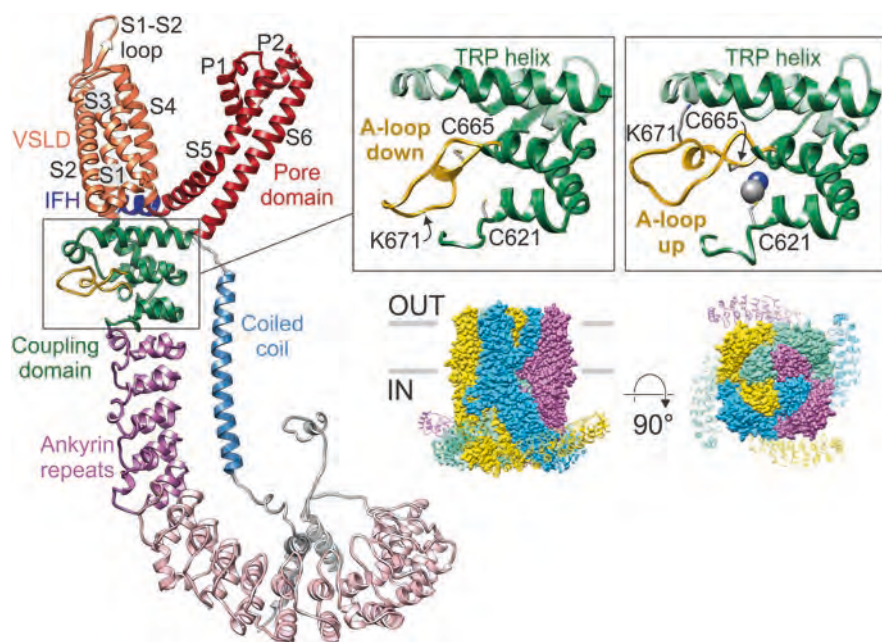


Fig. 10.2 Overview of human TRPA1 structure. Left, ribbon diagram of a single subunit of full-length human TRPA1 generated using AlphaFold prediction server. The structural arrangement elements represented in different colors indicate the voltage sensor-like domain (VSLD) comprised of the transmembrane helices S1–S4, the pore domain formed by the transmembrane helices S5 and S6 and a loop containing two pore helices P1 and P2, the coupling domain containing the reactive cysteines (C621, C665), the TRP helix and the activation loop (A-loop). Right top, close-up view of the coupling domain with the downward and upward configuration of the A-loop. Right bottom, the functional ion channel is a homotetramer with each monomer represented in different colors (left, side view and right, top view)

is a C terminal four-helix coiled-coil domain surrounded by N-terminal ankyrin repeats. While only the last 5 ankyrin repeats (AR12–AR16) of the predicted 16 are well resolved in all previously published structures of human TRPA1, recently published structures of *Drosophila* TRPA1 map all 17 ankyrin repeats stabilized in a 4-leaf propeller-shaped structure (Wang et al. 2023). Just above the ankyrin domain, directly below the membrane domain, is embedded the so-called coupling domain or allosteric nexus. It is formed by the TRP helix, which is implicated in the gating of numerous TRPs, eight short helices and three β -strands that form an antiparallel β -sheet. Importantly, the allosteric nexus contains reactive cysteines (C621, C665) that are crucial for TRPA1 sensitivity to electrophilic irritants (see below).

Pore-Forming Region

Of the more than ten available TRPA1 structures (PDB codes 3J9P, 6PQO-Q, 6V9V-Y, 6X2J, 6WJ5, 7JUP, 7OR0-1), only one (6V9X) is captured with sufficiently broad upper and lower constrictions to be classified as an open state (Zhao et al.

2020). In the open state, the entire transmembrane region is rotated by around 15° with respect to the stationary cytoplasmic ankyrin-repeat domain, and the VSLD is tilted backward from the central axis. The upper gate, formed by the constriction at residues G914 and D915, widens upon activation by upward shift and movement of pore helices P1 and P2 (Fig. 10.3). Two more transitions contribute to the opening of the pore domain: (1) the S4–S5 linker and S5 straighten into a single α -helix; and (2) the flexible π -helical turn in S6 shifts up by one helical turn. All this orchestrates a marked widening of the lower gate formed by residues I957 and V961, increasing its hydrated radius from 0.9 to 2.1 Å.

Compared to TRPV1, the conformational changes that accompany TRPA1 opening are quite pronounced, as evidenced by the salt-bridge between the residues K868 and E854 observed in the structure in the open state, whereas in the closed state, these residues are 16 Å apart (Fig. 10.3). Using site-directed mutagenesis complemented by molecular dynamics simulations, it was shown that this salt bridge is crucial for the stabilization of the open state and that its impairment contributes to the disease-associated phenotype of the N855S mutation (Zima et al. 2015).

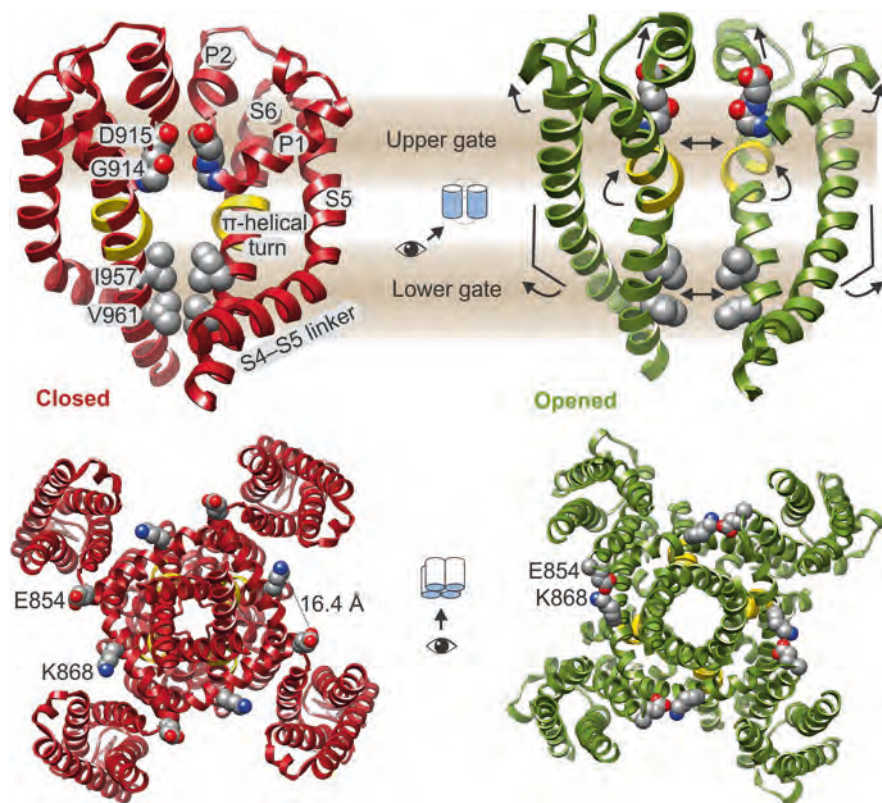


Fig. 10.3 The transmembrane domain of human TRPA1 in two states (PDB 6V9W—red and 6V9X—green). Top, side view of the pore region. Only two subunits are shown for clarity. Residues sealing the upper and lower gate are shown in sphere representation. Bottom, residues E854 and K868 that are 16 Å away in the closed conformation come together to form a salt-bridge when the channel is activated

Molecular Basis of TRPA1 Chemosensation

Activation by Electrophiles and Its Coupling to the Gate

Electrophile Binding Site

The uniqueness of TRPA1 lies in its ability to respond to structurally diverse electrophilic compounds dangerous to the cell. To fulfill this role, TRPA1 must be highly efficient in order to compete with the reactivity of antioxidant enzymes that are abundant in the cytosol. Indeed, the reactivity of conserved C621 has been shown to be at least one order of magnitude higher than that of reactive cysteines in antioxidant defense proteins (Bahia et al. 2016).

Two in-depth structural studies shed light on the molecular mechanism underlying TRPA1 activation by electrophiles and describe the main structural characteristics (Suo et al. 2020; Zhao et al. 2020): First, C621 is located at the base of the bottom half of the clamshell-shaped binding pocket, which provides an ideal chemical environment and promotes its high reactivity. Second, covalent binding of the ligand to C621 causes a structural rearrangement of the upper half of this pocket, the so-called activation loop (A-loop) to an upward configuration, resulting in pocket opening (Fig. 10.2). If the bound ligand is bulky (e.g., JT010), the pocket stabilizes and provides support for the conformation of the allosteric nexus and the transmembrane domain that favors the open state. The nature of the sensitivity of TRPA1 to electrophilic compounds of different shapes and sizes appears to be that the same stabilization of the binding pocket (meaning the same A-loop configuration) that is achieved with one bulky ligand is also achieved by association with two smaller ones. Indeed, weaker densities connected with C665 were observed in the presence of iodoacetamide and benzyl isothiocyanate. Subsequent molecular dynamics simulations proposed that C665 serves as a handle for the opening of the binding pocket (Habgood et al. 2022). The initial association of the ligand with the closed pocket by binding to C665 is followed by its transient opening, which allows rotation of C665, stabilizing the structure of the open pocket and exposing the deeper residing C621 to a permanent, more favorable modification. This hypothesis is supported by the experimental observation that although C665 is important for TRPA1 sensitivity to electrophiles, its occupancy is very low (~3%) compared to that of C621, which is 97% (Bahia et al. 2016). Thus, occupation of the binding pocket by electrophiles leads to a conformational change that stabilizes the A-loop in the upward configuration (Habgood et al. 2022).

Coupling Between the Electrophile-Binding Site and the Pore

There are several suggested explanations how the electrophile-binding site is coupled to the stabilization of the open channel state: First, the upward conformational change of A-loop brings K671 close to the TRP helix and, as a charged residue, increases its dipole moment (Fig. 10.2). Zhao and colleagues proposed that it is the electrostatic repulsion between TRP domains from neighboring subunits that drives

the expansion of the overlying S6 domains, opening the lower gate (Zhao et al. 2020). This is in line with the mutagenesis study that showed that four out of five positively charged TRP domain residues severely reduced the response to electrophiles when replaced by alanine (Samad et al. 2011).

Second, the upward A-loop position is also associated with a conformational change in the β -sheet, which is physically connected to a short α -helix partly buried in the inner membrane leaflet, known as the interfacial helix (IFH). IFH contributes to the stabilization of the TRPA1 channel gate in relation to its dynamic environment (e.g., the presence of phosphatidylinositol 4,5-bisphosphate in the membrane) (Macikova et al. 2019). Comparison of the different solved structures of TRPA1 showed that IFH can occupy a position in close interaction with the S4-S5 linker or a more flexible position nearby (Suo et al. 2020). Interestingly, when the IFH is only loosely positioned, the cavity on the side of the VSLD is filled with phospholipid. Based also on structural studies of the related TRPV1 and TRPV3, displacement of the lipid from the analogous site is predicted to shift the S4-S5 linker away from the central axis, facilitating movement of S5 and S6 and leading to channel opening (Nadezhdin et al. 2021; Nguyen et al. 2021).

Third, the coupling between the upward A-loop configuration and pore opening may also be mediated through the contraction of the membrane proximal ankyrin repeat domain and the C-terminal coiled coil (Bali et al. 2023). This mechanism was proposed by authors who characterized the functional impact of a human *TRPA1* gene variant associated with a rare peripheral nerve disorder, cramp fasciculation syndrome (Nirenberg et al. 2018). This variant introduces a nonsense mutation at R919 located at the beginning of the second pore helix P2, so that the truncated mutant protein lacks P2, S6, and the entire C-terminal tail, including the TRP helix and IFH, which should disconnect communication with the gate. The truncated mutant was not activated by either the electrophile agonist allyl isothiocyanate or the non-electrophilic agonist carvacrol when expressed alone. However, when co-transfected with wild-type TRPA1, it exhibited even higher sensitivity to the electrophilic agonist than the wild-type, indicating that conformational changes in the ankyrin repeat domain and in the coiled coil may contribute to the gating of heteromeric channels containing subunits with a disabled electrophile binding pocket. Alternatively, the electrophile sensitivity of the heteromers could rely on the highly conserved lysine 710 located in the pre-S1 helix, which has been shown to mediate some electrophile-induced activation in the absence of the requisite cysteines 621 and 665 (Hinman et al. 2006).

Modulation by Non-electrophilic Ligands

In addition to electrophilic compounds that interact with cysteines in the coupling domain, TRPA1 activity is influenced by a plethora of non-covalent modulators. The 3D structures published so far show the binding of several such ligands, most of which are bound in the transmembrane domain (Fig. 10.4). The only exception is the inositol hexakisphosphate, a polyanion that is apparently required for

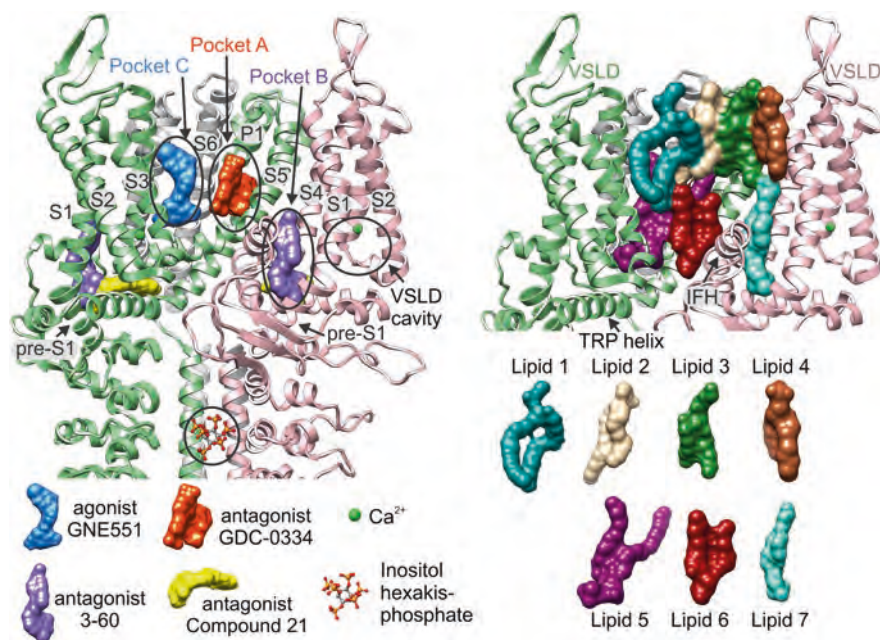


Fig. 10.4 TRPA1 non-electrophilic ligand binding sites. Ribbon diagram of TRPA1 generated using AlphaFold prediction server with each subunit color-coded. Hydrophobic ligands and lipids are shown in a solid surface representation. The polyanion that stabilizes the coiled-coil domain is shown in ball and stick representation

stabilization of the coiled-coil stalk. Within the transmembrane domain, three distinct binding pockets (A–C) can be identified, all of which are accessible from the lipid bilayer environment.

The first described binding site was the pocket A, which carries the inhibitor A-967079 (6V9Y) and also inhibitor GDC-0334 (6WJ5), and is located between S5 and P1 from one subunit and S6 from the adjacent one (Paulsen et al. 2015; Zhao et al. 2020; Balestrini et al. 2021). The inhibitory molecule bound at this site is thought to prevent the opening of the lower gate by locking the S5 helix in its bent conformation and inhibiting the movement of the S5–S6 pore module (Zhao et al. 2020). The lower part of the pocket A contains residues S873 and T874 that are critical for sensitivity of human TRPA1 to menthol (Xiao et al. 2008). Interestingly, in published TRP channel structures, the analogous binding pocket is occupied by both inhibitors (e.g., TRPC5 and Pico145, 6YSN) and agonists (e.g., TRPV2 and cannabidiol, 6U88), suggesting that the molecules that interact here can induce both inhibition and activation based on their shape and chemical nature (Wright et al. 2020; Pumroy et al. 2019).

Recently, a novel class of potent, selective, orally bioavailable TRPA1 small molecule antagonists based on tetrahydrofuran linker has been discovered (Terrett et al. 2021). The cryo-EM structure of TRPA1 with one such antagonist, compound21

(7JUP), revealed a binding site, which we refer here as pocket B. Compound21 is an elongated molecule oriented horizontally and partially wrapped around the TRP domain. One end is in contact with the pre-S1 helix (the cytosolic helix preceding S1) and the other with the bend between the S4 and the S4–S5 linker. There is another solved structure of TRPA1 with a ligand at the pocket B (7OR0, 7OR1). Curiously, the antagonist 3–60 occupies also the space between pre-S1 and TRP domain, but the molecule is oriented vertically and the rest of it is sandwiched between the IFH helix, S1, and the S4–S5 linker. It was suggested that ligands at pocket B antagonize the opening of the TRPA1 channel by acting as a molecular wedge that blocks the straightening of the S4–S5 linker and S5, and thus the transition to the open state (Terrett et al. 2021).

The next recognized binding site on TRPA1 is pocket C, known from several other TRP channel subfamilies and best described in TRPV1 as a vanilloid-binding pocket (Zhao et al. 2021). It is delineated by the S4 and S4–S5 linker of one subunit and S5 and S6 of an adjacent subunit and in case of TRPA1 accommodates agonist GNE551 (6X2J); (Liu et al. 2021). This compound has significantly different activating properties to electrophilic agonists. It does not induce desensitization and induces persistent pain relatively insensitive to antagonist treatment, hence demonstrating biased agonism of the TRPA1 channel.

All of the abovementioned pockets (A–C) face laterally to the membrane environment and are at least partially occupied by membrane lipids in the absence of a specific ligand. Suo and colleagues resolved seven phospholipids per subunit bound to the TRPA1 channel (6PQQ). All resolved lipids surround the pore domain between adjacent VSLD. Four of them are part of the extracellular leaflet of the membrane, three of them cytoplasmic. Note that ligand interaction with pockets A–C is linked with the displacement of such a resident lipid (Suo et al. 2020).

While the S2–S3 linker was not resolved on the first TRPA1 structure (3J9P), Zhao et al. captured this part of the VSLD even with defined side chains due to the stabilization caused by the bound calcium ion (6V9W) (Paulsen et al. 2015; Zhao et al. 2020). Calcium is coordinated by residues E788, Q791, Y799, N805, and E808. It has been shown that double mutation E788A with E808A preserves electrophile-induced activation as wild type, but causes total loss of calcium-dependent potentiation (Zimova et al. 2018). The calcium bound at this site is not unique to TRPA1 but is shared with TRPC5 or TRPM8 (Zhao et al. 2021). Interestingly, there are published structures for both TRPC5 and TRPM8 with small molecule ligands bound inside the VSLD cytosolic cavity (Song et al. 2021; Yin et al. 2018), and so a TRPA1 structure with a ligand-occupied VSLD cavity is expected to emerge in the near future. Moreover, the VSLD cavity is likely not the only site for direct interaction of TRPA1 with calcium, as TRPC3, for example, has two additional sites for calcium at the interface of the coiled coil and ankyrin repeats (7DXB); (Guo et al. 2022). Indeed, TRPA1 was previously predicted to have a calcium interaction site involving residues E1077 and D1080 (Sura et al. 2012), both of which are located at the unresolved distal C-terminus following the coiled coil domain, leaving the structural confirmation pending.

TRPA1 Activation by Physical Stimuli

Activation by Cold and Heat

The activation of TRPA1 in mammals by cold temperatures is still a matter of debate, as studies either support (Story et al. 2003; Sawada et al. 2007; Karashima et al. 2009; del Camino et al. 2010) or deny (Jordt et al. 2004; Knowlton et al. 2010; Cordero-Morales et al. 2011; Chen et al. 2013) the activation of the channel by cold or its involvement in cold perception. Mouse TRPA1 was found to be activated by lowering the temperature to 10 °C in dorsal root ganglion neurons and in transfected human embryonic kidney 293T cells (Bandell et al. 2004; del Camino et al. 2010). Cold increases the channel's open probability, slightly decreases the unitary conductance, and causes a shift of the half-maximal activation voltage closer to physiological membrane potentials (Sawada et al. 2007; Karashima et al. 2009; Chen et al. 2013). However, *Trpa1*-deficient mice were not different from wild-type littermates when examined for exploratory behavioral responses in a two-plate choice test with surface temperatures adjusted to 30 °C and 20 °C (Bautista et al. 2007; Knowlton et al. 2010), indicating an involvement of other cold-activated channels. Mouse TRPA1, on the other hand, mediates a crucial physiological role in the detection of noxious heat, although the channel alone is not sufficient to elicit an appropriate reaction (Hoffmann et al. 2013; Vandewauw et al. 2018). In nociceptive neurons, where TRPA1 expression partially overlaps with the related channels TRPV1 and TRPM3, all three channels are necessary for the defensive response against noxious heat stimuli (Vandewauw et al. 2018).

Human TRPA1 reconstituted into artificial planar lipid bilayers or expressed in human embryonic kidney 293T cells can be activated by 12 °C cooling (Kremeyer et al. 2010; Wang et al. 2013; Moparthi et al. 2014, 2016), but to a lesser extent than mouse TRPA1 (Sinica et al. 2019). Other groups did not observe any cold-induced currents upon cooling in human TRPA1 (Jordt et al. 2004; Cordero-Morales et al. 2011; Chen et al. 2013). A single V875 residue in human TRPA1, corresponding to G878 in mouse, accounts for the differences in cold sensitivity (Chen et al. 2013) and voltage-dependent activation (Sinica et al. 2019) between the two orthologs. In other studies, human and mouse TRPA1 exhibited activation only upon noxious heat stimulation above 50 °C, potentially explaining why some other authors did not observe a response because they did not reach such a high activation threshold (Hynkova et al. 2016; Sinica et al. 2019). Furthermore, the TRPA1-mediated responses to cold and heat are allosterically modulated by depolarizing voltage (Sinica et al. 2019; Karashima et al. 2009; Sinica and Vlachova 2021) and probably depend on membrane phosphatidylinositol-4,5-bisphosphate (Zimova et al. 2020).

The introduction of a portion of the N-terminus of rattlesnake or *Drosophila* TRPA1 into the human orthologue was sufficient to confer heat sensitivity, highlighting the role of the N-terminus in thermal activation (Cordero-Morales et al. 2011). This was recently strengthened by a structural study in which two N-terminal inter-subunit interfaces were identified in *Drosophila* TRPA1 that play a key role in thermal activation (Wang et al. 2023). Human genetic data also indicate that a

single-nucleotide polymorphism in the *TRPA1* gene, which results in a substitution in the N-terminal ankyrin repeat 4 (E179K), is associated with paradoxical heat sensation in neuropathic pain patients (Binder et al. 2011).

Molecular Basis of TRPA1 Mechanosensitivity

TRPA1 is unique among mammalian TRP channels in bearing an extensive cytoplasmic N terminus that consists of a prominent ankyrin repeat domain composed of a tandem of 16–17 ankyrin repeats. Structurally, the ankyrin repeats are characteristic ~33-amino acid motifs, each adopting a helix-turn-helix-extended loop secondary structure topology, stacking together in a close to linear fashion and producing a distinctive modular and elongated architecture (Gaudet 2008). Atomic force microscopy and molecular dynamics simulations confirm the elastic properties of these polyankyrin structures in response to pulling on both ends, suggesting a possible mechanical function of the spring (Sotomayor et al. 2005). It was because of the presence of the N-terminal polyankyrin domain and expression of the channel in hair cells of the inner ear in mice that TRPA1 was originally thought to function as a mechanosensitive channel responsible for hearing (Corey et al. 2004), but this was later not confirmed. On the other hand, TRPA1 has been shown to modulate mechanotransduction in nociceptor terminals and possibly in keratinocytes (Kwan et al. 2006, 2009).

Although there is increasing evidence that TRPA1 is involved in mechanosensory processes, it has long remained controversial as to whether it is the primary, intrinsic transducer of mechanical force. Application of negative pressure to membrane patches did not affect spontaneous channel activity in human embryonic kidney 293T cells expressing mouse TRPA1, indicating that the channel is not a stretch-sensitive primary mechanosensor but rather a secondary membrane receptor involved in downstream mechanosensory signaling cascades (Nikolaev et al. 2019). Human TRPA1 outperforms other known mechanosensitive channels in sensitivity to ultrasound, with the N-terminal tip domain (amino acids 1–25) proving to be critical for this property (Duque et al. 2022). Although the mechanism of ultrasound transduction is unknown, the channels that respond to it are thought to be mechanosensitive. For TRPA1-mediated ultrasound responses, an intact actin cytoskeleton and lipid membrane interactions are essential, suggesting a more indirect mechanism. However, when purified human TRPA1 is incorporated into artificial bilayers, it responds to negative pressure even when its N-terminus is removed, suggesting intrinsic mechanosensitivity (Moparthy and Zygmunt 2020). To reconcile these results, it can be considered that anisotropic forces of the lipid bilayer and their changes can reshape the embedded channel, and if it has a small free-energy difference between the closed and open states, caused, e.g., by truncation or interaction with another protein, the channel pore will expand in the gate and allow the passage of ions (Anishkin et al. 2014). The human TRPA1 mechanosensitivity is dependent on its redox state, so that oxidative stress may shift the channel into a conformation sensitive to mechanical stimuli (Moparthy and Zygmunt 2020).

Interspecies Diversity of TRPA1

Activation of TRPA1 by electrophilic agonists is present in all species, including insects, fish, amphibians, reptiles, birds, and mammals (Laursen et al. 2014). Phylogenetic analysis of TRPA1 orthologs and functional studies point to a role for this channel as an ancient chemical sensor that is conserved across approximately 500 million years of animal evolution (Kang et al. 2010; Arenas et al. 2017). It is supposed that such widespread conservation may be related to the potential toxicity of electrophilic compounds, which threaten all living organisms without exception. These include not only environmental irritant and pungent electrophiles but also endogenous substances commonly found at sites of injury and inflammation, such as by-products of oxidative and nitrative stress, underscoring that the channel is an important component of the cellular defense system (Viana 2016).

Unlike electrophiles, non-covalent modulators usually exhibit a species-specific bimodal action, i.e., activation at lower concentrations and inhibition at higher concentrations. This includes divergent effects of non-covalently acting caffeine (Nagatomo and Kubo 2008), menthol (Xiao et al. 2008) and nicotine (Talavera et al. 2009) on TRPA1 orthologs. Typical non-covalent modulators are monoterpenes comprising a group of naturally occurring organic compounds derived from essential oils that are generally used for analgesic, anesthetic, anti-inflammatory, and antipruritic effects (Takaishi et al. 2014). It is thought that interspecies differences in TRPA1 affinity for these substances may be strongly related to adaptation to the environment and dependence on food intake and availability. Comparisons of different TRPA1 orthologs together with functional characterization of wild-type and chimeric constructs helped to identify binding sites for caffeine (Nagatomo and Kubo 2008) and menthol (Xiao et al. 2008) as well as the synthesized selective inhibitor HC-030031 (Gupta et al. 2016).

Comparison of primary protein sequences in different species is one of the most powerful tools to identify specific sites that have diverged evolutionarily as a result of adaptation to external cues. On the other hand, these substantial interspecies differences complicate common procedures established in the search for potential drugs and do not allow direct projection and translation of the results of laboratory experiments performed on mice to primates and humans.

As can be seen from the comparison of the primary sequence of TRPA1 across all species (Fig. 10.5), the most conserved regions are the sites providing the main functions of the channel, i.e., gating and permeation of cations. In addition, high conservation is also evident on the concave side of the N-terminal ankyrin repeat domain and on the distal part of the C-terminal coiled coil domain. The least conserved are the distal portions of the N- and C-terminal regions, the interfacial helix, and the S1–S2 external loop, which is involved in channel voltage regulation (Marsakova et al. 2017) and N-glycosylation (Egan et al. 2016).

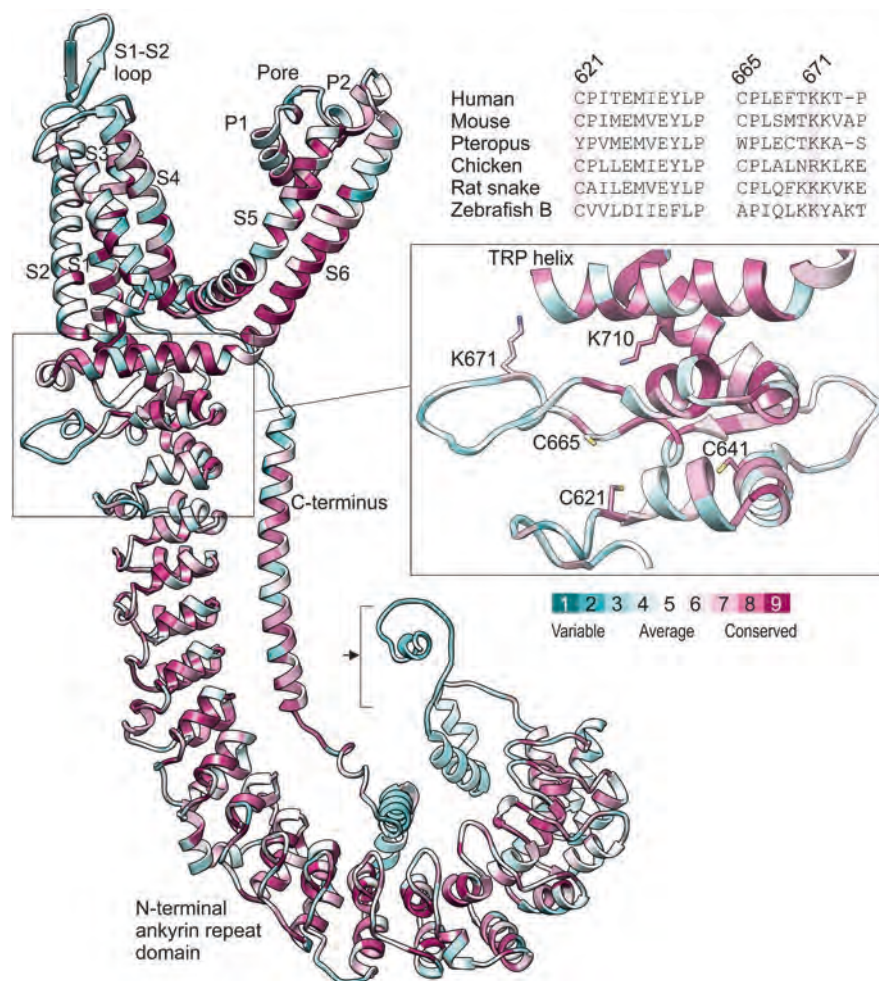


Fig. 10.5 Sequence conservation analysis computed using the ConSurf server (Yariv et al. 2023). Conservation scores are mapped onto the structure of human TRPA1 subunit built using AlphaFold prediction server. The degree of conservation follows a gradient from cyan, least conserved, to magenta, most conserved. Note the N-terminal end domain involved in ultrasound sensitivity (amino acids 1–25 marked with an arrow). In detail (right): The functionally key part of the protein, so-called coupling domain (Suo et al. 2020) or allosteric nexus (Zhao et al. 2020), composed of N-terminal and C-terminal short helices, a dynamic activation loop containing lysine 671, and the TRP helix. Cysteine and lysine residues involved in the electrophile-induced activation of TRPA1 are indicated. Right top, multiple sequence alignment of TRPA1 orthologs. In the process of electrophile-induced activation, covalent modification of cysteine 621 is followed by partial modification of cysteine 665 and repositioning of lysine 671 that is stabilized through its interaction with backbone carbonyl oxygens at the C-terminus of the TRP helix, leading to pore widening. Modification of highly conserved lysine 710, located in the pre-S1 helix, activates the channel through a mechanism independent of cysteine 621 (Zhao et al. 2020; Hinman et al. 2006)

Posttranslational Modifications

Phosphorylation is one of the most important posttranslational mechanisms reversibly regulating TRPA1 in different cellular contexts and under various physiological and pathophysiological conditions. Despite the presumed physiological relevance of the molecular mechanisms that establish and control the subtle balance between phosphorylated and dephosphorylated states of TRPA1, there is still little information about the specific sites underlying these processes (reviewed in Kadkova et al. 2017; Meents et al. 2019; Talavera et al. 2020; Voolstra and Huber 2014).

Under states associated with inflammation, tissue damage, or oxidative stress, endogenous substances are released at the site of injury that directly activate TRPA1 or indirectly regulate its function via G-protein-coupled receptor (GPCR) and phospholipase C-coupled signaling cascades. These substances include extracellular proteases, prostaglandins, bradykinin, serotonin, substance P, calcitonin gene-related peptide, nerve growth factor, tumor necrosis factor α , and many more (Nilius et al. 2012; Talavera et al. 2020; Viana 2016; Zygmunt and Hogestatt 2014). These compounds stimulate protein kinase A (PKA) (Wang et al. 2008a; Brackley et al. 2017; Meents et al. 2017; Anand et al. 2008), protein kinase C (PKC) (Brackley et al. 2017), p38 mitogen-activated protein kinases (Obata et al. 2005), cyclin-dependent kinase 5 (Sulak et al. 2018; Hall et al. 2018; Hynkova et al. 2016), or phospholipase C pathways (Wang et al. 2008a; Dai et al. 2007) to induce phosphorylation of TRPA1. Phosphorylation, in turn, significantly alters the activation properties and/or plasma membrane expression of TRPA1, typically sensitizing the inflammation and sensitizing nociceptive neurons to noxious stimuli. For effective phosphorylation by PKA and PKC, TRPA1 needs the presence of the scaffolding A-kinase anchoring protein 79/150 (AKAP) (Brackley et al. 2017). AKAP directly interacts with TRPA1 (Zhang et al. 2008), spatially constrains phosphorylation, and most likely serves as a molecular hub that contributes to the efficiency and specificity of the cellular signaling network regulating the channel under various physiological or pathophysiological conditions (Zimova et al. 2020; Barvikova et al. 2020). Importantly, the sensitization of TRPA1 through the PKA-AKAP pathway is responsible for persistent mechanical hypersensitivity (Brackley et al. 2017), carboplatin-induced mechanical allodynia, and cold hyperalgesia (Miyano et al. 2019). Most of the kinases described so far acting on TRPA1 upregulate the channel's function. On the other hand, 5' adenosine monophosphate-activated protein kinase (AMPK) downregulates TRPA1 plasma membrane expression in dorsal root ganglion neurons which may represent a potential mechanism underlying painful diabetic peripheral neuropathy and other diseases with a similar pathophysiological profile of metabolic dysfunction (Wang et al. 2018). In addition, several tyrosine residues located at the N-terminus were found to be involved in the inhibition of TRPA1 by Src kinase in SH-SY5Y human neuroblastoma cells (Morgan et al. 2015). Thus, the spectrum of effects by which TRPA1 function can be regulated by phosphorylation in the natural environment is broad.

TRPA1 contains two glycosylation sites ⁷⁴⁷NST and ⁷⁵³NETS on the S1–S2 extracellular linker (Egan et al. 2016; Deering-Rice et al. 2015; Marsakova et al. 2017). Molecular dynamics simulations and mutagenesis studies indicate that the S1–S2 region includes structurally important hydrophobic residues that must be properly positioned relative to the S1–S4 sensor as well as to the pore module to contribute to the stabilization of the channel's open state. This view is also supported by observations that replacing serines S748 and S756 with hydrophobic residues improved responses induced by allyl isothiocyanate (Marsakova et al. 2017).

Protein-Protein Interactions with TRPA1

TRPA1 interacts with a variety of proteins in the course of its life cycle (Fig. 10.6). First, during biogenesis in the endoplasmic reticulum, there are chaperones (EMC, σ 1R). Trafficking to the membrane, where TRPA1 localizes into cholesterol-rich sensory micro-domains, is mediated by the SNARE complex. The interaction of

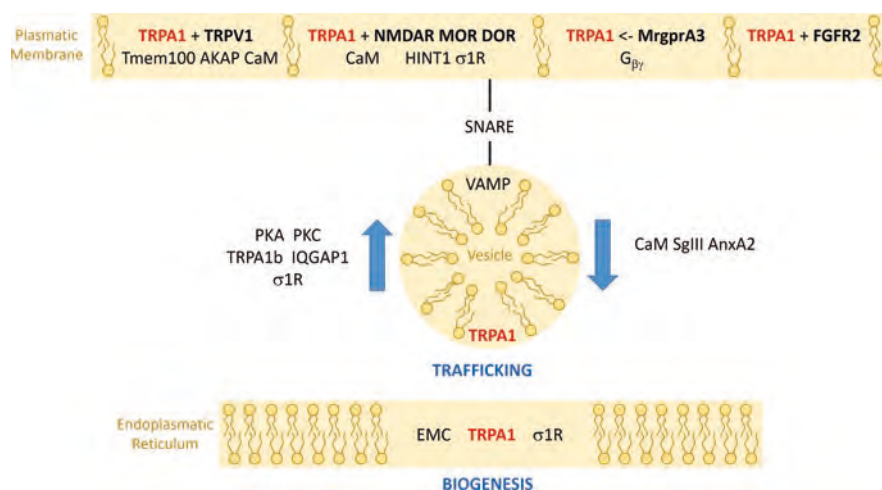


Fig. 10.6 TRPA1-protein interactions during its biogenesis, trafficking, and in the plasmatic membrane. Chaperones endoplasmic reticulum-resident multiprotein complex (EMC) and Sigma-1 receptor (σ 1R) regulate TRPA1 biogenesis. Protein kinases A and C (PKA, PKC), splice variant TRPA1b, IQ motif containing GTPase activating protein 1 (IQGAP1), and Sigma-1 receptor (σ 1R) upregulate TRPA1 trafficking. Calmodulin (CaM), secretogranin III (SgIII), annexin A2 (AnxA2) downregulate TRPA1 trafficking. Adapter proteins CaM, AKAP79/150, or TMEM100 stabilize complexes or super-complexes of TRPA1 with TRPV1, N-methyl-D-aspartate receptor (NMDAR), G-protein-coupled receptors: μ -opioid receptor (MOR) and δ -opioid receptor (DOR), Mas-related G protein-coupled receptor A3 (MrgprA3) and fibroblast growth factor receptor 2 (FGFR2). In the plasma membrane, NMDAR is connected via histidine triad nucleotide-binding protein 1 (HINT1) and σ 1R to MOR

TRPA1 with many other proteins upregulates (PKA, PKC, TRPA1b, IQGAP1, σ 1R) or downregulates (CaM, SgIII, AnxA2) this process. Finally, in the plasma membrane, several adapter proteins (CaM, AKAP79/150, TMEM100) stabilize complexes or super-complexes of TRPA1 with several other membrane proteins: TRPV1, NMDAR, GPCRs (MOR, DOR, MrgrprA3), FGFR2, etc.

Biogenesis

ER-Resident Multiprotein Complex (EMC)

The dynamic ribosome/polysome-translocon complexes, which reside at the endoplasmic reticulum (ER) membrane, produce a major fraction of the human proteome governing the synthesis, translocation, folding, membrane insertion, maturation, N-glycosylation, and disulfide-bond formation of nascent proteins (Gemmer et al. 2023). One of the translocons, the ER-resident multiprotein complex (EMC), can be found in all eukaryotes. It functions as a transmembrane insertase required for the insertion and proper membrane topology of tail-anchored and multi-transmembrane proteins, including GPCRs, transporters, and ion channels—presumably including TRP channels. After all, multiple EMC subunits were identified in a screen of modulators of TRPC6 gain-of-function toxicity, with validated hits EMC2 and EMC6 subunits both necessary for effective TRPC6 expression (Talbot et al. 2019). The recent structures of EMC- Ca_v complexes provide the first view of an EMC-client complex and define EMC sites, where the transmembrane and cytoplasmic parts of ion channels dock (Chen et al. 2023); see Fig. 10.7. Given the conserved structure of the voltage sensor domain (VSD) within the voltage-gated ion channel (VGIC) superfamily, the EMC may participate in the biogenesis of other VGICs through transmembrane dock interactions similar to those defined for Ca_v , or other, yet-to-be-defined, EMC client-binding sites (Chen et al. 2023). Interestingly, many VGICs (including TRPA1) bear disease mutations that are not easily rationalized based on mature channel structure. Some of these mutations may disturb EMC interactions (Chen et al. 2023). Once correctly folded, the transmembrane proteins can be transported to the cell membrane (Sato et al. 2015).

Trafficking

σ 1R

The ability to traffic a variety of functional proteins to the plasma membrane has, among many other functions, the Sigma-1 receptor (σ 1R). It is an endoplasmic reticulum chaperone (Marcotti et al. 2022) that regulates via direct protein-protein interactions the functions of all VGIC superfamilies (i.e., Na^+ , Ca^{2+} , and K^+

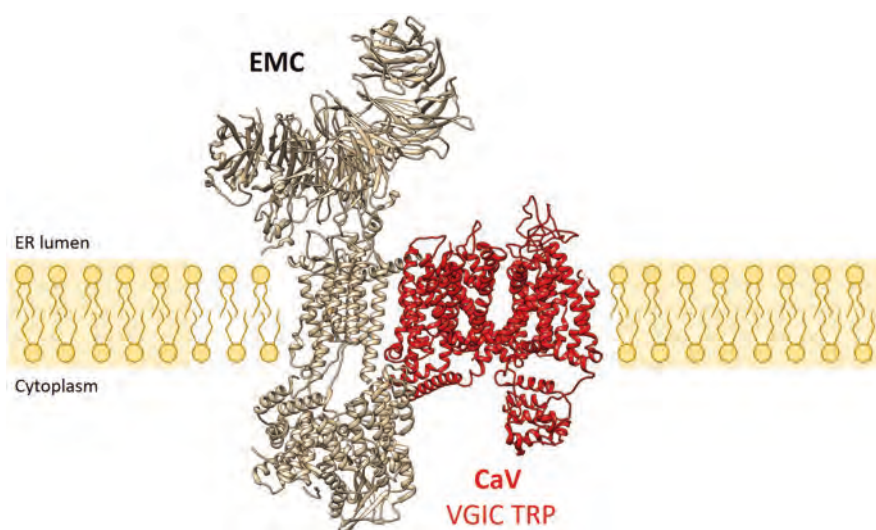


Fig. 10.7 The first EMC-client (Ca_v) complex defines EMC sites, where ion channels dock (Chen et al. 2023)

families) and classes (i.e., ligand-gated ion channels, e.g., NMDA receptors, and GPCRs). In the case of the TRPV1 ion channel, ($\sigma 1R$) stabilizes its complexes with GPCRs μ -opioid receptors (MORs) and δ -opioid receptors (DORs). $\sigma 1R$ can be viewed as an atypical dynamic auxiliary subunit. Interestingly, it has no known homologs in the eukaryotic genome (Soriani and Kourrich 2019). It is an evolutionary isolate with no discernible similarity to any other human protein (Schmidt et al. 2016, 2018). Moreover, $\sigma 1R$ is the only known chaperone protein whose activity is regulated by endogenous and synthetic compounds in a clear agonist-antagonist manner (Soriani and Kourrich 2019). The pharmacological antagonism of $\sigma 1R$ impairs the formation of the $\sigma 1R$ -TRPA1 molecular complex and the trafficking of functional TRPA1 to the plasma membrane (Marcotti et al. 2022).

SNARE

The trafficking of TRP channels to the plasma membrane and the release of calcitonin gene-related peptide (CGRP) from trigeminal ganglion neurons are implicated in some aspects of chronic migraines. These exocytotic processes are inhibited by the cleavage of SNARE multi-protein complexes with botulinum neurotoxins (Belinskaia et al. 2023). More specifically, in the case of sensory neurons expressing TRPA1 and TRPV1, these channels are packaged into CGRP- and VAMP1-containing vesicles and delivered to the plasma membrane using the SNARE complex composed of SNAP-25, Syntaxin 1, and VAMP1 (synaptic vesicle membrane protein), as well as Munc18-1 (Meng et al. 2016); see Fig. 10.8. Since TRPA1

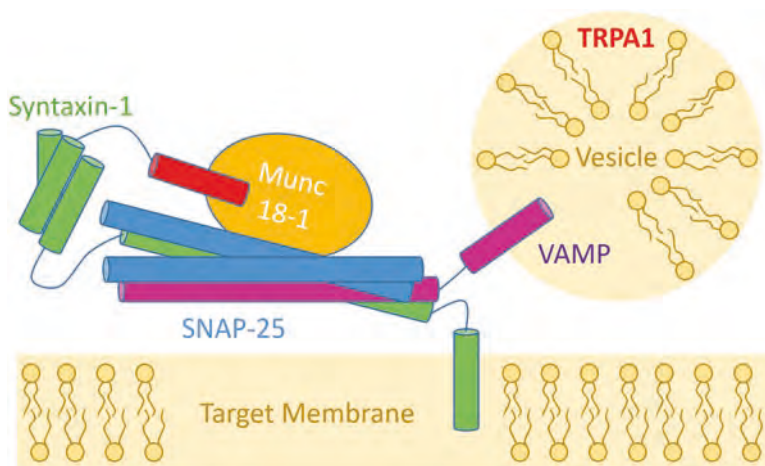


Fig. 10.8 TRPA1 packaged into VAMP-containing vesicles is delivered to the plasma membrane using the SNARE complex consisting of SNAP-25, Syntaxin 1, and Munc18-1

trafficking is somewhat dependent on its state, some form of direct interaction could likely occur between the large and mobile cytoplasmic portions of TRPA1 and the SNARE complex.

PKA/PLC, MO: Nociceptive Signals

TRPA1 membrane levels can be increased (and TRPA1-mediated nocifensive behavior can be sensitized) via protein kinase A/phospholipase C (PKA/PLC) signaling and by activating TRPA1 with the ligand mustard oil (MO) (Schmidt et al. 2009). Tetanus toxin attenuated the response to the second of two pulses of MO in neurons, suggesting that vesicle fusion increases functional surface TRPA1. Capacitance recordings suggested that MO induces exocytosis. Therefore, it was proposed that TRPA1 translocation to the membrane might represent one of the mechanisms controlling TRPA1 functionality upon acute activation or inflammatory signals (Schmidt et al. 2009).

Splice Variant TRPA1b

TRPA1b, an alternative splice variant of the mouse *Trpa1* gene, can physically interact with full-length TRPA1 (TRPA1a). Interestingly, it again increases the expression of TRPA1a in the plasma membrane as well as current density in response to different agonists, however, without affecting the single-channel conductance (Zhou et al. 2013).

IQGAP1 (CaM)

Further, TRPA1 undergoes increased trafficking to the membrane of dorsal root ganglia neurons catalyzed by the small GTPase Cdc42 associated with the IQ motif containing GTPase activating protein 1 (IQGAP1), leading to functional sensitization of the channel (Khan et al. 2022). Cdc42 anchored on IQGAP1 is crucial to TRPA1 trafficking. The distal C terminus of IQGAP1 (1361–1657) is the major binding region for TRPA1. Deletion of the distal 114 amino acids containing the coiled-coil domain from the C terminus of TRPA1 dramatically reduced the binding to C1-IQGAP1. Further truncation of 17AA containing the calmodulin (CaM)-binding region entirely abolished binding to C1-IQGAP1, showing that the distal C terminus of IQGAP1 binds primarily to the distal C terminus of TRPA1. Also, N-IQGAP1 binds to the proximal N terminus of TRPA1 between M450 and C652. IQGAP1 contains four tandem IQ motifs responsible for binding to Ca²⁺-loaded CaM and Ca²⁺-free CaM mediating the Ca²⁺-dependent regulation of IQGAP1. TRPA1 contains a noncanonical CaM-binding domain binding to Ca²⁺-loaded CaM but not to Ca²⁺-free CaM. CaM acts as a central negative regulator of IQGAP1 by preventing other partner proteins from binding to IQGAP1 (Khan et al. 2022). Indeed, overexpression of CaM or CaM1234 caused a robust reduction in the binding of IQGAP1 to TRPA1 even under the basal condition. Ca²⁺-loaded CaM after TRPA1 opening should exhibit reduced binding to IQGAP1, so that the inhibitory effect of CaM on IQGAP1 is relieved mediating increased TRPA1 binding and trafficking. Deletion of IQGAP1 markedly reduces mechanical and cold hypersensitivity in chronic inflammatory and neuropathic pain in mice. IQGAP1 thus promotes chronic pain by coupling the trafficking and signaling machinery to TRPA1 channels (Khan et al. 2022).

SgIII

Interaction trap analysis suggested that secretogranin III (SgIII) has multiple binding partners, including the mast cell ion channel TRPA1 (Prasad et al. 2008). Overexpression of SgIII suppresses calcium release via TRPA1. It likely reflects a disruption of TRPA1 localization to cholesterol-rich microdomains (Prasad et al. 2008).

AnxA2

Similarly, annexin A2 (AnxA2) limits the availability of TRPA1 channels to regulate nociceptive signaling in vertebrates (Avenali et al. 2014). An increase in TRPA1 membrane levels was observed in cultured sensory neurons from AnxA2-deficient mice. AnxA2 is enriched in a subpopulation of sensory neurons and co-expressed

with TRPA1. There is a physical association of AnxA2 with TRPA1 (Avenali et al. 2014). AnxA2 coimmunoprecipitates with TRPA1 in a heterologous expression system. Nevertheless, the truncated form of AnxA2 did not. It suggests that the first 15 amino acids of AnxA2 contribute to binding TRPA1 channels either directly (or, hypothetically, via yet-to-be-identified mediator proteins).

Protein Linkers

CaM

One of the most essential modulators of TRPA1 are calcium ions, which activate the channel at low concentrations and inactivate it at higher concentrations. These effects of calcium may occur by direct interaction with TRPA1 or as a result of association with a calcium-binding partner such as CaM (Moparathi et al. 2020), which pre-associates with the membrane-proximal C-terminal region of TRPA1 (L992-N1008) from where the Ca^{2+} -CaM enables the channel to distinctly respond to diverse Ca^{2+} signals (Hasan et al. 2017). As noted by Zimova et al. (2020), this region overlaps with the interaction site for membrane PIP_2 with which Ca^{2+} -CaM is likely to compete. Whereas the C-lobe of CaM acts as a tether anchoring CaM to the binding site at TRPA1, the N-lobe of conformationally promiscuous CaM is only partly involved in binding and does not affect the channel gating (Hasan et al. 2017). The N-lobe of conformationally extremely promiscuous CaM (Fig. 10.9) may either bridge different domains of TRPA1 or link the channel with some different target protein(s). This is possible because its two independently folded Ca^{2+} -binding lobes can interact differentially and, to some degree, separately (Zimova et al. 2020).

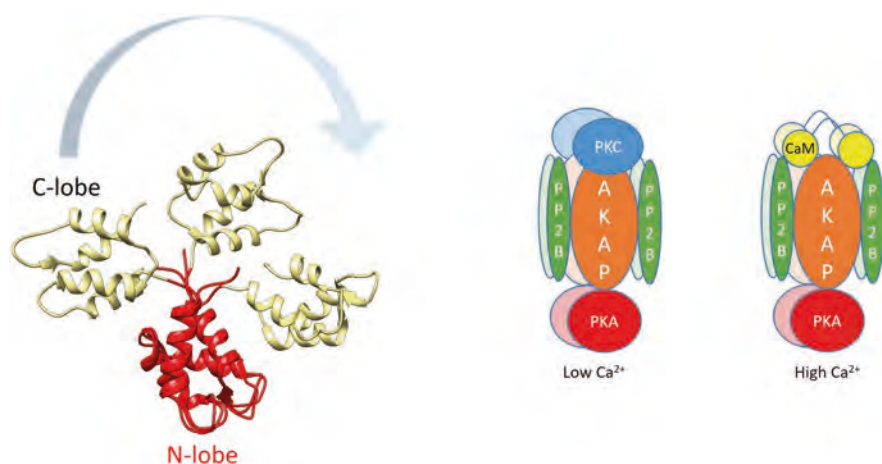


Fig. 10.9 Left, conformational variability of the bilobate calmodulin (CaM) linker involved in the stabilization of many multi-protein complexes. Right, AKAP79/150 regulating signaling multi-protein complexes. A-kinase anchoring protein (AKAP), protein phosphatase 2B (PP2B)

AKAP79/150

Another complexing partner of TRPA1 (and TRPV1) is the A-kinase anchoring protein (AKAP) 79/150 (human/rodent150) (Zimova et al. 2020; Gamper and Shah 2022). AKAP79/150 is a regulator of PKA and PKC. Nevertheless, it regulates a much broader list of targets and is involved in the formation of multi-protein signaling complexes. AKAP79/150 homo-dimerizes (upon addition of a lysine-reactive cross-linker, parallel homomeric dimers are stabilized through K328-K328 and K333-K333 cross-links and allow the recruitment of a single PKA or PKC per AKAP79/150 molecule (Gold et al. 2011). It is supposed that this dimerization has an impact on the arrangements within AKAP-mediated multi-protein complexes (Gamper and Shah 2022; Gold et al. 2011). An assembly of greater complexity comprising AKAP79, protein phosphatase 2B (PP2B), a type II regulatory subunit fragment (RII 1-45) of PKA, and CaM was reconstituted in vitro (Gold et al. 2011); see Fig. 10.9. Dimeric AKAP79 coordinates two RII 1-45 homodimers, four PP2B heterodimers, and two CaM molecules. The binding of Ca^{2+} /CaM to AKAP79 stabilizes the complex by generating a second interface for PP2B. This leads to the activation of the anchored phosphatases (Gold et al. 2011). Moreover, AKAP79 has been shown to interact directly with members of the MAGUK family of scaffolding molecules, which associate with many proteins in the postsynaptic densities, including glutamate receptors (Dodge and Scott 2000). Mechanisms of inflammatory mechanical hypersensitivity were found to involve the activation of mGluR5 which sensitized TRPA1 in an AKAP-dependent manner (Brackley et al. 2017). Zimova and colleagues presented evidence that TRPA1 can interact with AKAP to prime the TRPA1 channel for opening (Zimova et al. 2020). They asked whether the overexpression of AKAP in HEK293T cells may influence the functional response of TRPA1. Significant basal currents were observed at negative potentials, and they did not further increase upon repeated stimulation. They proposed several possibilities for how to explain this effect. AKAP may help to recruit TRPA1 to the plasma membrane, it may increase the activity of TRPA1 by increasing its basal phosphorylation, or it could induce a conformational change that primes the channel for activation.

Plasmatic Membrane

TRPV1 (TMEM100, CaM, AKAP)

In a large subset of sensory neurons, TRPA1 physically and functionally interacts/cooperates with the structurally related TRPV1 ion channel (Zimova et al. 2020; Gamper and Shah 2022; Akopian et al. 2007; Staruschenko et al. 2010). These two channels may sensitize or desensitize each other via the elevation of intracellular calcium ions (Akopian et al. 2007). TRPV1 activation by capsaicin sensitized TRPA1 and led to inflammatory cold hyperalgesia. TRPA1 activation enhanced TRPV1 activity via increased Ca^{2+} levels and in a PKA-dependent manner (Gamper

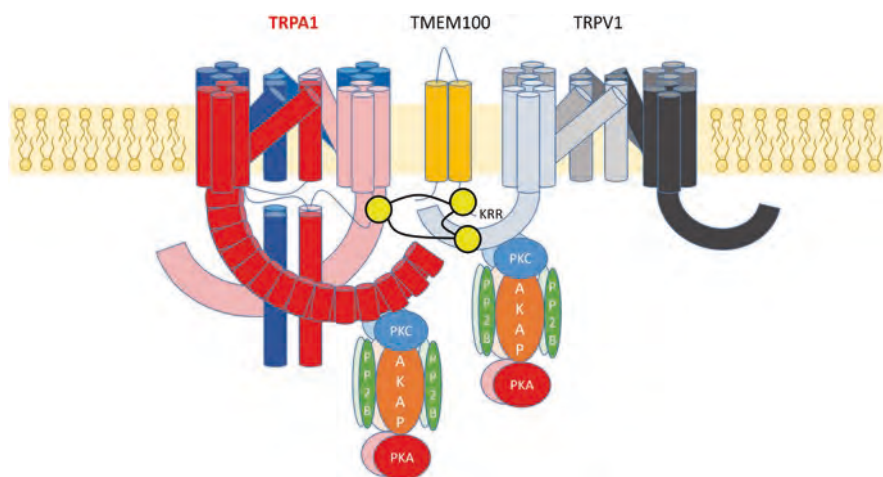


Fig. 10.10 TRPA1–TRPV1 complex stabilized by the transmembrane adaptor protein TMEM100 with possible involvement of CaM and/or AKAP79/150

and Shah 2022). A direct association of TRPA1 with TRPV1 strongly inhibits the responses of TRPA1 to electrophilic agonists, independently of Ca^{2+} (Staruschenko et al. 2010). TRPV1 counteracts the pharmacological desensitization of TRPA1 by preventing TRPA1 internalization. TRPV1:TRPA1 concatemers behave like TRPV1 channels; however, TRPA1 exerted an inhibitory effect on the activity of TRPV1. In peptidergic neurons, the interaction of TRPA1 and TRPV1 is tightly regulated by the transmembrane adaptor protein TMEM100. When TMEM100 is not present, TRPV1 exerts an inhibitory effect on TRPA1 activity, but the presence of TMEM100 enhances the activity of TRPA1 by spacing the channels (Fig. 10.10). Structurally, the regulation of this interaction depends on a KRR amino acid triplet on the C-terminus of TMEM100 (Weng et al. 2015). Mutant TMEM100 (KRR-QQQ) was unable to produce this spacing effect. Instead, the interaction within the TRPV1-TRPA1 complex was enhanced. In TRPV1-TRPA1 expressing cells, mutant TMEM100 was able to reduce TRPA1 conductance (Weng et al. 2015). Importantly, the highly positively charged C-terminus of TMEM100 also contains a putative CaM binding site (L99–L115). The N- (amino acids 220–260) and C- (684–720) terminal domains on TRPV1 are responsible for TRPA1-TRPV1 complex formation (Patil et al. 2020). Peptides encoding these regions were able to disrupt and reduce TRPV1-TRPA1 complex formation in sensory neurons. These domains partially overlap with the previously identified binding sites for CaM/ PIP_2 (189–222/682–725) (Rosenbaum et al. 2004; Ufret-Vincenty et al. 2011). Therefore, as proposed by Zimova et al. (2020), CaM could serve as a linker between TRPA1, TRPV1, and TMEM100. The C-lobe of CaM can bind to the C-terminus of TRPA1, and the N-lobe of CaM could bind to either TMEM100 or TRPV1 (Fig. 10.10).

NMDAR, MOR, DOR (HINT1, σ 1R) + TRPA1 (CaM)

The glutamate N-methyl-D-aspartate receptor (NMDAR), Ca^{2+} -permeable glutamate-gated ion channel whose physiological properties are regulated by intracellular Ca^{2+} , plays a relevant role in allodynia to mechanical stimuli. NMDAR is connected via the histidine triad nucleotide-binding protein 1 (HINT1) and type 1 sigma receptor (σ 1R) to μ -opioid receptors (MORs), which mediate the most potent pain relief (Cortés-Montero et al. 2021). TRPA1 channels are functionally associated with MORs, δ -opioid receptors (DORs), and NMDARs in the dorsal root ganglia, the spinal cord, and brain areas. The MOR-TRPA1 and NMDAR-TRPA1 associations do not require HINT1 or σ 1R but appear to be mediated by calcium-activated CaM (Cortés-Montero et al. 2021); see Fig. 10.11. Using a sequence-based motif search, Wang et al. (1999) have identified the third intracellular loop of the opioid receptor family as a possible direct contact point for interaction with CaM, in addition to its established role in G-protein activation. It means that CaM directly competes with G proteins for binding to opioid receptors and that CaM may itself serve as an independent second messenger molecule that is released upon receptor stimulation. Further, CaM interacts with the NR1 subunit of the NMDA receptor. CaM binding to NR1 causes a fourfold reduction in NMDA channel open probability (Ehlers et al. 1996).

MrgprA3 ($\text{G}_{\beta\gamma}$)

$\text{G}_{\beta\gamma}$ was proposed as a likely candidate for mediating the specific coupling of Mas-related G protein-coupled receptor A3 (MrgprA3) and TRPA1 (Fig. 10.12). $\text{G}_{\beta\gamma}$ modulates several ion channels via direct binding, including members of the G

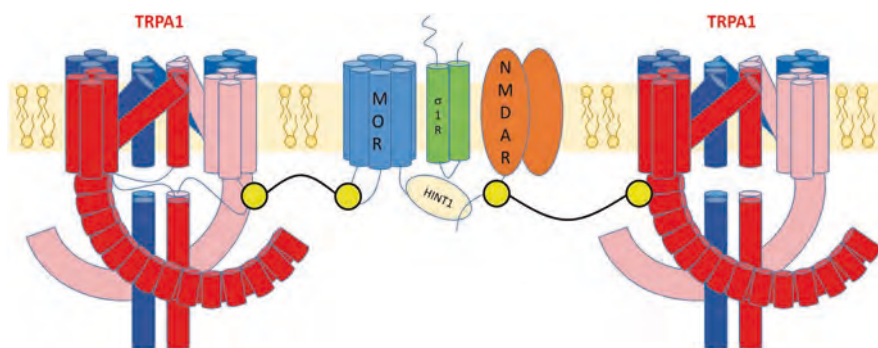


Fig. 10.11 NMDAR is connected via histidine triad nucleotide-binding protein 1 (HINT1) and type 1 sigma receptor (σ 1R) to μ -opioid receptors (MOR) and δ -opioid receptors (DOR). The MOR-TRPA1 and N-methyl-D-aspartate receptor (NMDAR)-TRPA1 association does not require HINT1 and σ 1R; it is mediated by calcium-activated CaM

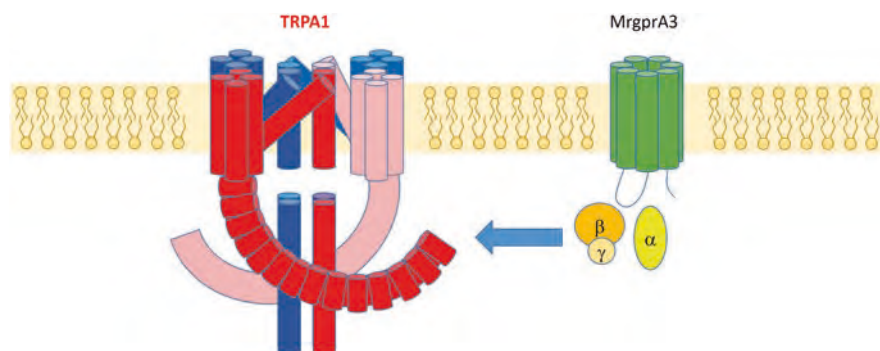


Fig. 10.12 G $\beta\gamma$ mediates the specific coupling of Mas-related G protein-coupled receptor A3 (MrgprA3) and TRPA1

protein-coupled inwardly rectifying potassium channel and voltage-gated calcium channel families (Wilson et al. 2011). Very recently, the cryo-EM structure of the complex between TRPM3 and G $\beta\gamma$ was resolved (Zhao and MacKinnon 2023). A major side effect of the MrgprA3 agonist and antimalarial drug chloroquine is intolerable itch. Therefore, inhibition of MrgprA3 or TRPA1 may enable chloroquine to remain a useful and relevant treatment (Wilson et al. 2011).

FGFR2

The membrane fibroblast growth factor receptor 2 (FGFR2) drives lung adenocarcinoma (LUAD) progression through aberrant protein-protein interactions mediated via its C-terminal proline-rich motif (Berrout et al. 2017). FGFR2 and TRPA1 form a complex, in which the ankyrin repeats of TRPA1 bind to the C-terminal proline-rich region of FGFR2 (Fig. 10.13). This interaction inhibits the TRPA1 channel yet acts as a possible conformational scaffold that induces FGFR2 receptor constitutive activation independent of extracellular stimulation, leading to aberrant proliferation and invasion. Therefore, TRPA1 could be considered an invasive and prognostic marker in LUAD cancer (Berrout et al. 2017).

Future Directions for the Study of TRPA1-Containing Protein Complexes

Gamper and colleague noted that the multitude of interactions reported in the literature raises the question of whether different proteins form individual clusters, where each complex is independent of other complex types or whether is there a large dynamic “super-complex” where there is interplay between multiple different proteins with TRPV1 at the heart of it (Gamper and Shah 2022). From what we

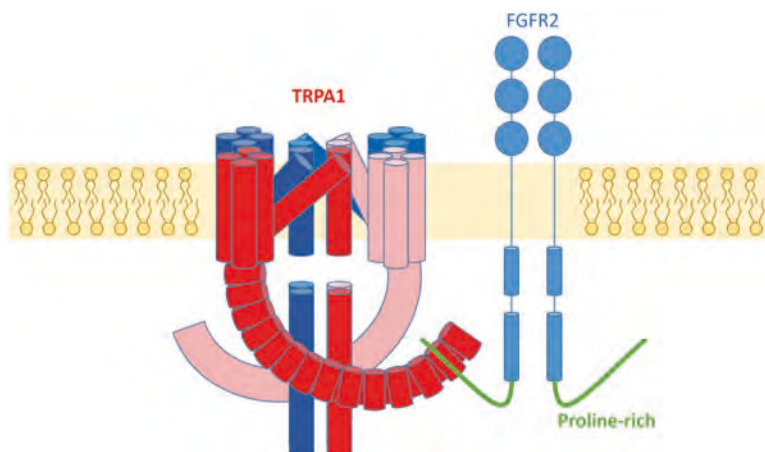


Fig. 10.13 The membrane fibroblast growth factor receptor 2, FGFR2, and TRPA1 form a complex, in which the ankyrin repeats of TRPA1 bind to the C-terminal proline-rich region of FGFR2

described above, it is clear that even TRPA1 or NMDAR have their corona of interaction partners stabilized and interconnected by more or less universal linker/scaffold proteins (CaM, AKAP, HINT1, or σ 1R). Therefore, they should be viewed and approached rather as a dynamic network of interacting super-complexes.

Conclusions and Perspectives

Recent incredible advances in single-particle cryo-electron microscopy have provided an essential framework for understanding the molecular basis of TRPA1 function. Without this key tool and the precise knowledge of the mechanisms of gating and regulation, it would not only be impossible to fully understand TRPA1-related disorders and the implications of genetic *TRPA1* polymorphisms in humans but also to study the channel's drug interaction and rationally screen for novel modulators. Importantly, structures in different conformational states are already available not only for TRPA1 but also for most major representatives of related TRP channels that share key structural features. Thus, much can be inferred on the basis of homology (Diver et al. 2022; Hilton et al. 2019). What is still lacking is a precise structural description of the transitions between individual conformational states and can only be predicted from ambiguous mutagenesis, truncation, and chimeric studies. Even from the large number of mutants and constructs that have been reported so far for TRPA1 (Meents et al. 2019; Benedikt et al. 2009; Samad et al. 2011; Sura et al. 2012; Cordero-Morales et al. 2011; Hynkova et al. 2016; Jabba et al. 2014), it is not possible to determine which regions of the channel protein complex are specific and essential and which are less important for processes such as activation/deactivation and sensitization/desensitization. Molecular dynamics simulations may provide

more clues, but these are still limited by the available computational capacity. Uniquely for the related TRPV1 channel, discrete substrates along the activation pathways have recently been characterized by the team of Prof. David Julius, the 2021 Nobel Prize Laureate (Zhang et al. 2021), so we can anticipate that similar progress may soon be achieved for the TRPA1 channel.

The number of articles published annually on TRPA1 has reached a plateau in recent years (about 250 entries in the PubMed database per year), and it is possible to some extent to predict the main directions of research in the coming years. In addition to further refining the 3D structure in various conformations, basic research interest will remain focused on understanding the physiological functions of the channel, mainly using experimental mice. The success of this research will depend on whether and possibly when the role of TRPA1 in temperature detection in humans is understood. This is one of the important but not the only parameter in which the human receptor differs from the mouse ortholog (Sinica et al. 2019; Koivisto et al. 2022). Pharmaceutical companies and investors will certainly want to avoid the complications in drug development that have occurred in the past with TRPV1 antagonists, which have shown adverse effects on temperature sensitivity and thermoregulation in human clinical trials (Bamps et al. 2021). This research is inconceivable without a general understanding of interspecies differences in the temperature sensitivity settings of the channel (Laursen et al. 2015). A recent seminal study on the heat-sensitive TRPA1 isoform in *Drosophila melanogaster* proposed a conserved role of the cytoplasmic C-terminal coiled-coil domain which shifts upward by ~12–13 Å upon heat stimulation (35 °C) and demonstrated that induced changes in the interaction of the distal C-terminus with the N-terminal ankyrin repeat domain are essential for thermal activation (Wang et al. 2023). Whether this is also the case for the human ortholog and what conformational changes are induced by temperatures felt by humans as painful remain to be explored.

In addition to these aforementioned directions of research on TRPA1, another has gradually emerged that goes beyond trying to understand the principles of channel function. It is represented by studies that aim to exploit the unique properties of TRPA1 as a molecular sensor. In 2014, thermal activation of artificially expressed *TRPA1* in different combinations of neurons allowed the study of locomotor behavior in *Drosophila*, thus providing an alternative to optogenetic approaches to explore and control the nervous system—thermogenetics (Bidaye et al. 2014). Since then, genetic techniques have advanced considerably and have recently even enabled the identification and use of human TRPA1 as a sonogenetic tool for noninvasive and selective targeting of deeper brain regions in mice using 7 MHz ultrasound (Duque et al. 2022). It turns out that the distal part of the N-terminus (amino acids 1–25) and the interaction with membrane cholesterol are essential for this activation modality. Thus, knowledge about the structure and function of TRPA1 gradually fits together. Once we know the channel to the last amino acid, an even more challenging phase of research will be undertaken to investigate exactly what structural interactions determine the channel's major functions in the natural environment of the human body and whether and how this information can be used to treat and improve health in humans.

Acknowledgments The authors would like to acknowledge support from the Czech Science Foundation (GACR 22-13750S) and Alexandra Ptakova for critical discussion of the manuscript.

References

- Achenbach J, Rhein M, Gombert S et al (2019) Childhood traumatization is associated with differences in TRPA1 promoter methylation in female patients with multisomatoform disorder with pain as the leading bodily symptom. *Clin Epigenetics* 11(1). <https://doi.org/10.1186/s13148-019-0731-0>
- Akopian AN, Ruparel NB, Jeske NA, Hargreaves KM (2007) Transient receptor potential TRPA1 channel desensitization in sensory neurons is agonist dependent and regulated by TRPV1-directed internalization. *J Physiol* 583(Pt 1):175–193
- Anand U, Otto WR, Facer P et al (2008) TRPA1 receptor localisation in the human peripheral nervous system and functional studies in cultured human and rat sensory neurons. *Neurosci Lett* 438(2):221–227. <https://doi.org/10.1016/j.neulet.2008.04.007>
- Andersson DA, Gentry C, Moss S, Bevan S (2008) Transient receptor potential A1 is a sensory receptor for multiple products of oxidative stress. *J Neurosci* 28(10):2485–2494. <https://doi.org/10.1523/JNEUROSCI.5369-07.2008>
- Andrade EL, Luiz AP, Ferreira J, Calixto JB (2008) Pronociceptive response elicited by TRPA1 receptor activation in mice. *Neuroscience* 152(2):511–520. <https://doi.org/10.1016/j.neuroscience.2007.12.039>
- Andrei SR, Sinharoy P, Bratz IN, Damron DS (2016) TRPA1 is functionally co-expressed with TRPV1 in cardiac muscle: co-localization at z-discs, costameres and intercalated discs. *Channels* 10(5):395–409. <https://doi.org/10.1080/19336950.2016.1185579>
- Anishkin A, Loukin SH, Teng J, Kung C (2014) Feeling the hidden mechanical forces in lipid bilayer is an original sense. *Proc Natl Acad Sci* 111(22):7898–7905. <https://doi.org/10.1073/pnas.1313364111>
- Arenas OM, Zaharieva EE, Para A et al (2017) Activation of planarian TRPA1 by reactive oxygen species reveals a conserved mechanism for animal nociception. *Nat Neurosci* 20(12):1686–1693. <https://doi.org/10.1038/s41593-017-0005-0>
- Atoyan R, Shander D, Botchkareva NV (2009) Non-neuronal expression of transient receptor potential type A1 (TRPA1) in human skin. *J Invest Dermatol* 129(9):2312–2315. <https://doi.org/10.1038/jid.2009.58>
- Avenali L, Narayanan P, Rouwette T et al (2014) Annexin A2 regulates TRPA1-dependent nociception. *J Neurosci* 34(44):14506–14516. <https://doi.org/10.1523/JNEUROSCI.1801-14.2014>
- Bahia PK, Parks TA, Stanford KR et al (2016) The exceptionally high reactivity of Cys 621 is critical for electrophilic activation of the sensory nerve ion channel TRPA1. *J Gen Physiol* 147(6):451–465. <https://doi.org/10.1085/jgp.201611581>
- Balestrini A, Joseph V, Dourado M et al (2021) A TRPA1 inhibitor suppresses neurogenic inflammation and airway contraction for asthma treatment. *J Exp Med* 218(4). <https://doi.org/10.1084/jem.20201637>
- Bali A, Schaefer SP, Trier I et al (2023) Molecular mechanism of hyperactivation conferred by a truncation of TRPA1. *Nat Commun* 14(1):2867. <https://doi.org/10.1038/s41467-023-38542-1>
- Bamps D, Vriens J, de Hoon J, Voets T (2021) TRP channel cooperation for nociception: therapeutic opportunities. *Annu Rev Pharmacol Toxicol* 61:655–677. <https://doi.org/10.1146/annurev-pharmtox-010919-023238>
- Bandell M, Story GM, Hwang SW et al (2004) Noxious cold ion channel TRPA1 is activated by pungent compounds and bradykinin. *Neuron* 41(6):849–857. [https://doi.org/10.1016/s0896-6273\(04\)00150-3](https://doi.org/10.1016/s0896-6273(04)00150-3)

- Barvikova K, Barvik I, Sinica V et al (2020) Phospho-mimetic mutation at Ser602 inactivates human TRPA1 channel. *Int J Mol Sci* 21(21). <https://doi.org/10.3390/ijms21217995>
- Bautista DM, Movahed P, Hinman A et al (2005) Pungent products from garlic activate the sensory ion channel TRPA1. *Proc Natl Acad Sci U S A* 102(34):12248–12252. <https://doi.org/10.1073/pnas.0505356102>
- Bautista DM, Jordt SE, Nikai T et al (2006) TRPA1 mediates the inflammatory actions of environmental irritants and proalgesic agents. *Cell* 124(6):1269–1282. <https://doi.org/10.1016/j.cell.2006.02.023>
- Bautista DM, Siemens J, Glazer JM et al (2007) The menthol receptor TRPM8 is the principal detector of environmental cold. *Nature* 448(7150):204–208. <https://doi.org/10.1038/nature05910>
- Belinskaia M, Wang J, Kaza SK et al (2023) Bipartite activation of sensory neurons by a TRPA1 agonist allyl isothiocyanate is reflected by complex Ca²⁺ influx and CGRP release patterns: enhancement by NGF and inhibition with VAMP and SNAP-25 cleaving botulinum neurotoxins. *Int J Mol Sci* 24(2):1338. <https://doi.org/10.3390/ijms24021338>
- Benedikt J, Samad A, Ettrich R et al (2009) Essential role for the putative S6 inner pore region in the activation gating of the human TRPA1 channel. *Biochim Biophys Acta* 1793(7):1279–1288. <https://doi.org/10.1016/j.bbamer.2009.04.014>
- Bernal L, Sotelo-Hitschfeld P, König C et al (2021) Odontoblast TRPC5 channels signal cold pain in teeth. *Sci Adv* 7(13). <https://doi.org/10.1126/sciadv.abf5567>
- Berrout J, Kyriakopoulou E, Moparathi L et al (2017) TRPA1-FGFR2 binding event is a regulatory oncogenic driver modulated by miRNA-142-3p. *Nat Commun* 8(1):947. <https://doi.org/10.1038/s41467-017-00983-w>
- Bidaye SS, Machacek C, Wu Y, Dickson BJ (2014) Neuronal control of Drosophila walking direction. *Science* 344(6179):97–101. <https://doi.org/10.1126/science.1249964>
- Binder A, May D, Baron R et al (2011) Transient receptor potential channel polymorphisms are associated with the somatosensory function in neuropathic pain patients. *PLoS One* 6(3):e17387. <https://doi.org/10.1371/journal.pone.0017387>
- Brackley AD, Gomez R, Guerrero KA et al (2017) A-kinase anchoring protein 79/150 scaffolds transient receptor potential A1 phosphorylation and sensitization by metabotropic glutamate receptor activation. *Sci Rep* 7(1):1842. <https://doi.org/10.1038/s41598-017-01999-4>
- Cao DS, Zhong L, Hsieh TH et al (2012) Expression of transient receptor potential ankyrin 1 (TRPA1) and its role in insulin release from rat pancreatic beta cells. *PLoS One* 7(5):e38005. <https://doi.org/10.1371/journal.pone.0038005>
- Cao E, Liao M, Cheng Y, Julius D (2013) TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* 504(7478):113–118. <https://doi.org/10.1038/nature12823>
- Chan P, Ding HT, Liederer BM et al (2021) Translational and pharmacokinetic-pharmacodynamic application for the clinical development of GDC-0334, a novel TRPA1 inhibitor. *Clin Transl Sci* 14(5):1945–1954. <https://doi.org/10.1111/cts.13049>
- Chen H, Terrett JA (2020) Transient receptor potential ankyrin 1 (TRPA1) antagonists: a patent review (2015–2019). *Expert Opin Ther Patents* 30(9):643–657. <https://doi.org/10.1080/013543776.2020.1797679>
- Chen J, Kang D, Xu J et al (2013) Species differences and molecular determinant of TRPA1 cold sensitivity. *Nat Commun* 4:2501. <https://doi.org/10.1038/ncomms3501>
- Chen Z, Mondal A, Abderemane-Ali F et al (2023) EMC chaperone–CaV structure reveals an ion channel assembly intermediate. *Nature*. <https://doi.org/10.1038/s41586-023-06175-5>
- Cordero-Morales JF, Gracheva EO, Julius D (2011) Cytoplasmic ankyrin repeats of transient receptor potential A1 (TRPA1) dictate sensitivity to thermal and chemical stimuli. *Proc Natl Acad Sci U S A* 108(46):E1184–E1191. <https://doi.org/10.1073/pnas.1114124108>
- Corey DP, Garcia-Anoveros J, Holt JR et al (2004) TRPA1 is a candidate for the mechanosensitive transduction channel of vertebrate hair cells. *Nature* 432(7018):723–730. <https://doi.org/10.1038/nature03066>

- Cortés-Montero E, Rodríguez-Muñoz M, Ruiz-Cantero MC et al (2021) Calmodulin supports TRPA1 channel association with opioid receptors and glutamate NMDA receptors in the nervous tissue. *Int J Mol Sci* 22(1):229. <https://doi.org/10.3390/ijms22010229>
- Dai Y, Wang S, Tominaga M et al (2007) Sensitization of TRPA1 by PAR2 contributes to the sensation of inflammatory pain. *J Clin Invest* 117(7):1979–1987. <https://doi.org/10.1172/JCI30951>
- de Araujo DSM, Nassini R, Geppetti P, De Logu F (2020) TRPA1 as a therapeutic target for nociceptive pain. *Expert Opin Ther Targets* 24(10):997–1008. <https://doi.org/10.1080/14728222.2020.1815191>
- de la Roche J, Eberhardt MJ, Klinger AB et al (2013) The molecular basis for species-specific activation of human TRPA1 protein by protons involves poorly conserved residues within transmembrane domains 5 and 6. *J Biol Chem* 288(28):20280–20292. <https://doi.org/10.1074/jbc.M113.479337>
- De Logu F, Nassini R, Materazzi S et al (2017) Schwann cell TRPA1 mediates neuroinflammation that sustains macrophage-dependent neuropathic pain in mice. *Nat Commun* 8(1):1887. <https://doi.org/10.1038/s41467-017-01739-2>
- Deering-Rice CE, Shapiro D, Romero EG et al (2015) Activation of transient receptor potential Ankyrin-1 by insoluble particulate material and association with asthma. *Am J Respir Cell Mol Biol* 53(6):893–901. <https://doi.org/10.1165/rcmb.2015-0086OC>
- del Camino D, Murphy S, Heiry M et al (2010) TRPA1 contributes to cold hypersensitivity. *J Neurosci* 30(45):15165–15174. <https://doi.org/10.1523/JNEUROSCI.2580-10.2010>
- Diver MM, King JVL, Julius D, Cheng Y (2022) Sensory TRP channels in three dimensions. *Annu Rev Biochem*. <https://doi.org/10.1146/annurev-biochem-032620-105738>
- Dodge K, Scott JD (2000) AKAP79 and the evolution of the AKAP model. *FEBS Lett* 476(1–2):58–61. [https://doi.org/10.1016/s0014-5793\(00\)01671-9](https://doi.org/10.1016/s0014-5793(00)01671-9)
- Doerner JF, Gisselmann G, Hatt H, Wetzel CH (2007) Transient receptor potential channel A1 is directly gated by calcium ions. *J Biol Chem* 282(18):13180–13189. <https://doi.org/10.1074/jbc.M607849200>
- Duque M, Lee-Kubli CA, Tufail Y et al (2022) Sonogenetic control of mammalian cells using exogenous transient receptor potential A1 channels. *Nat Commun* 13(1):600. <https://doi.org/10.1038/s41467-022-28205-y>
- Earley S, Gonzales AL, Crnich R (2009) Endothelium-dependent cerebral artery dilation mediated by TRPA1 and Ca²⁺-Activated K⁺ channels. *Circ Res* 104(8):987–994. <https://doi.org/10.1161/CIRCRESAHA.108.189530>
- Egan TJ, Acuna MA, Zenobi-Wong M et al (2016) Effects of N-glycosylation of the human cation channel TRPA1 on agonist-sensitivity. *Biosci Rep*. <https://doi.org/10.1042/BSR20160149>
- Ehlers MD, Zhang S, Bernhardt JP, Haganir RL (1996) Inactivation of NMDA receptors by direct interaction of calmodulin with the NR1 subunit. *Cell* 84(5):745–755. [https://doi.org/10.1016/S0092-8674\(00\)81052-1](https://doi.org/10.1016/S0092-8674(00)81052-1)
- El Karim IA, Linden GJ, Curtis TM et al (2011) Human odontoblasts express functional thermo-sensitive TRP channels: implications for dentin sensitivity. *Pain* 152(10):2211–2223. <https://doi.org/10.1016/j.pain.2010.10.016>
- Gallo V, Dijk FN, Holloway JW et al (2017) TRPA1 gene polymorphisms and childhood asthma. *Pediatr Allergy Immunol* 28(2):191–198. <https://doi.org/10.1111/pai.12673>
- Gamper N, Shah S (2022) Inferiority complex: why do sensory ion channels multimerize? *Biochem Soc Trans* 50(1):213–222. <https://doi.org/10.1042/BSR20211002>
- Gaudet R (2008) A primer on ankyrin repeat function in TRP channels and beyond. *Mol Biosyst* 4(5):372–379. <https://doi.org/10.1039/b801481g>
- Gemmer M, Chaillet ML, van Loenhout J et al (2023) Visualization of translation and protein biogenesis at the ER membrane. *Nature* 614(7946):160–167. <https://doi.org/10.1038/s41586-022-05638-5>
- Gold MG, Stengel F, Nygren PJ et al (2011) Architecture and dynamics of an A-kinase anchoring protein 79 (AKAP79) signaling complex. *Proc Natl Acad Sci* 108(16):6426–6431. <https://doi.org/10.1073/pnas.1014400108>

- Gouin O, L'Herondelle K, Lebonvallet N et al (2017) TRPV1 and TRPA1 in cutaneous neurogenic and chronic inflammation: pro-inflammatory response induced by their activation and their sensitization. *Protein Cell* 8(9):644–661. <https://doi.org/10.1007/s13238-017-0395-5>
- Guo W, Tang Q, Wei M et al (2022) Structural mechanism of human TRPC3 and TRPC6 channel regulation by their intracellular calcium-binding sites. *Neuron* 110(6):1023–1035.e1025. <https://doi.org/10.1016/j.neuron.2021.12.023>
- Gupta R, Saito S, Mori Y et al (2016) Structural basis of TRPA1 inhibition by HC-030031 utilizing species-specific differences. *Sci Rep* 6:37460. <https://doi.org/10.1038/srep37460>
- Habgood M, Seiferth D, Zaki AM et al (2022) Atomistic mechanisms of human TRPA1 activation by electrophile irritants through molecular dynamics simulation and mutual information analysis. *Sci Rep* 12(1):4929. <https://doi.org/10.1038/s41598-022-08824-7>
- Hall BE, Prochazkova M, Sapio MR et al (2018) Phosphorylation of the transient receptor potential Ankyrin 1 by Cyclin-dependent Kinase 5 affects Chemo-nociception. *Sci Rep* 8(1):1177. <https://doi.org/10.1038/s41598-018-19532-6>
- Hamilton NB, Kolodziejczyk K, Kougioumtzidou E, Attwell D (2016) Proton-gated Ca^{2+} -permeable TRP channels damage myelin in conditions mimicking ischaemia. *Nature* 529(7587):523–527. <https://doi.org/10.1038/nature16519>
- Hasan R, Leeson-Payne AT, Jaggar JH, Zhang X (2017) Calmodulin is responsible for Ca^{2+} -dependent regulation of TRPA1 Channels. *Sci Rep* 7:45098. <https://doi.org/10.1038/srep45098>
- Hilton JK, Kim M, Van Horn WD (2019) Structural and evolutionary insights point to allosteric regulation of TRP ion channels. *Acc Chem Res* 52(6):1643–1652. <https://doi.org/10.1021/acs.accounts.9b00075>
- Hinman A, Chuang HH, Bautista DM, Julius D (2006) TRP channel activation by reversible covalent modification. *Proc Natl Acad Sci U S A* 103(51):19564–19568. <https://doi.org/10.1073/pnas.0609598103>
- Hoffmann T, Kistner K, Miermeister F et al (2013) TRPA1 and TRPV1 are differentially involved in heat nociception of mice. *Eur J Pain* 17(10):1472–1482. <https://doi.org/10.1002/j.1532-2149.2013.00331.x>
- Hu Z, Zhang Y, Yu W et al (2023) Transient receptor potential ankyrin 1 (TRPA1) modulators: recent update and future perspective. *Eur J Med Chem* 257:115392. <https://doi.org/10.1016/j.ejmech.2023.115392>
- Hynkova A, Marsakova L, Vaskova J, Vlachova V (2016) N-terminal tetrapeptide T/SPLH motifs contribute to multimodal activation of human TRPA1 channel. *Sci Rep* 6:28700. <https://doi.org/10.1038/srep28700>
- Jabba S, Goyal R, Sosa-Pagan JO et al (2014) Directionality of temperature activation in mouse TRPA1 ion channel can be inverted by single-point mutations in ankyrin repeat six. *Neuron* 82(5):1017–1031. <https://doi.org/10.1016/j.neuron.2014.04.016>
- Jain SM, Balamurugan R, Tandon M et al (2022) Randomized, double-blind, placebo-controlled trial of ISC 17536, an oral inhibitor of transient receptor potential ankyrin 1, in patients with painful diabetic peripheral neuropathy: impact of preserved small nerve fiber function. *Pain* 163(6):e738–e747. <https://doi.org/10.1097/j.pain.0000000000002470>
- Jaquemar D, Schenker T, Trueb B (1999) An ankyrin-like protein with transmembrane domains is specifically lost after oncogenic transformation of human fibroblasts. *J Biol Chem* 274(11):7325–7333. <https://doi.org/10.1074/jbc.274.11.7325>
- Jhun EH, Hu X, Sadhu N et al (2018) Transient receptor potential polymorphism and haplotype associate with crisis pain in sickle cell disease. *Pharmacogenomics* 19(5):401–411. <https://doi.org/10.2217/pgs-2017-0198>
- Jordt SE, Bautista DM, Chuang HH et al (2004) Mustard oils and cannabinoids excite sensory nerve fibres through the TRP channel ANKTM1. *Nature* 427(6971):260–265. <https://doi.org/10.1038/nature02282>
- Kadkova A, Synytsya V, Krusek J et al (2017) Molecular basis of TRPA1 regulation in nociceptive neurons. A review. *Physiol Res* 66(3):425–439. <https://doi.org/10.33549/physiolres.933553>

- Kang K, Pulver SR, Panzano VC et al (2010) Analysis of *Drosophila* TRPA1 reveals an ancient origin for human chemical nociception. *Nature* 464(7288):597–600. <https://doi.org/10.1038/nature08848>
- Karashima Y, Talavera K, Everaerts W et al (2009) TRPA1 acts as a cold sensor in vitro and in vivo. *Proc Natl Acad Sci U S A* 106(4):1273–1278
- Khan S, Patra PH, Somerfield H et al (2022) IQGAP1 promotes chronic pain by regulating the trafficking and sensitization of TRPA1 channels. *Brain* 146(6):2595–2611. <https://doi.org/10.1093/brain/awac462>
- Kim H, Mittal DP, Iadarola MJ, Dionne RA (2006) Genetic predictors for acute experimental cold and heat pain sensitivity in humans. *J Med Genet* 43(8):e40. <https://doi.org/10.1136/jmg.2005.036079>
- Kittaka H, Tominaga M (2017) The molecular and cellular mechanisms of itch and the involvement of TRP channels in the peripheral sensory nervous system and skin. *Allergol Int* 66(1):22–30. <https://doi.org/10.1016/j.alit.2016.10.003>
- Knowlton WM, Bifolck-Fisher A, Bautista DM, McKemy DD (2010) TRPM8, but not TRPA1, is required for neural and behavioral responses to acute noxious cold temperatures and cold-mimetics in vivo. *Pain* 150(2):340–350. <https://doi.org/10.1016/j.pain.2010.05.021>
- Kochukov MY, McNearney TA, Fu Y, Westlund KN (2006) Thermosensitive TRP ion channels mediate cytosolic calcium response in human synoviocytes. *Am J Phys Cell Phys* 291(3):C424–C432. <https://doi.org/10.1152/ajpcell.00553.2005>
- Koivisto A, Jalava N, Bratty R, Pertovaara A (2018) TRPA1 antagonists for pain relief. *Pharmaceuticals* 11(4). <https://doi.org/10.3390/ph11040117>
- Koivisto AP, Belvisi MG, Gaudet R, Szallasi A (2022) Advances in TRP channel drug discovery: from target validation to clinical studies. *Nat Rev Drug Discov* 21(1):41–59. <https://doi.org/10.1038/s41573-021-00268-4>
- Kremeyer B, Lopera F, Cox JJ et al (2010) A gain-of-function mutation in TRPA1 causes familial episodic pain syndrome. *Neuron* 66(5):671–680. <https://doi.org/10.1016/j.neuron.2010.04.030>
- Kwan KY, Allchorne AJ, Vollrath MA et al (2006) TRPA1 contributes to cold, mechanical, and chemical nociception but is not essential for hair-cell transduction. *Neuron* 50(2):277–289. <https://doi.org/10.1016/j.neuron.2006.03.042>
- Kwan KY, Glazer JM, Corey DP et al (2009) TRPA1 modulates mechanotransduction in cutaneous sensory neurons. *J Neurosci* 29(15):4808–4819. <https://doi.org/10.1523/JNEUROSCI.5380-08.2009>
- Landini L, de Araujo DSM, Titiz M et al (2022) TRPA1 role in inflammatory disorders: what is known so far? *Int J Mol Sci* 23(9). <https://doi.org/10.3390/ijms23094529>
- Laursen WJ, Bagriantsev SN, Gracheva EO (2014) TRPA1 channels: chemical and temperature sensitivity. *Curr Top Membr* 74:89–112. <https://doi.org/10.1016/B978-0-12-800181-3.00004-X>
- Laursen WJ, Anderson EO, Hoffstaetter LJ et al (2015) Species-specific temperature sensitivity of TRPA1. *Temperature (Austin)* 2(2):214–226. <https://doi.org/10.1080/23328940.2014.1000702>
- Li J, Zhang H, Du Q et al (2023) Research progress on TRPA1 in diseases. *J Membr Biol*. <https://doi.org/10.1007/s00232-023-00277-x>
- Liao M, Cao E, Julius D, Cheng Y (2013) Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504(7478):107–112. <https://doi.org/10.1038/nature12822>
- Liu C, Reese R, Vu S et al (2021) A non-covalent ligand reveals biased agonism of the TRPA1 ion channel. *Neuron* 109(2):273–284.e274. <https://doi.org/10.1016/j.neuron.2020.10.014>
- Macikova L, Vyklicka L, Barvik I et al (2019) Cytoplasmic inter-subunit interface controls use-dependence of thermal activation of TRPV3 channel. *Int J Mol Sci* 20(16). <https://doi.org/10.3390/ijms20163990>
- Marchi M, Salvi E, Andelic M et al (2023) TRPA1 rare variants in chronic neuropathic and nociplastic pain patients. *Pain*. <https://doi.org/10.1097/j.pain.0000000000002905>

- Marcotti A, Fernandez-Trillo J, Gonzalez A et al (2022) TRPA1 modulation by Sigma-1 receptor prevents oxaliplatin-induced painful peripheral neuropathy. *Brain*. <https://doi.org/10.1093/brain/awac273>
- Marsakova L, Barvik I, Zima V et al (2017) The first extracellular linker is important for several aspects of the gating mechanism of human TRPA1 channel. *Front Mol Neurosci* 10:16. <https://doi.org/10.3389/fnmol.2017.00016>
- Meents JE, Fischer MJ, McNaughton PA (2017) Sensitization of TRPA1 by protein Kinase A. *PLoS One* 12(1):e0170097. <https://doi.org/10.1371/journal.pone.0170097>
- Meents JE, Ciotu CI, Fischer MJM (2019) TRPA1: a molecular view. *J Neurophysiol* 121(2):427–443. <https://doi.org/10.1152/jn.00524.2018>
- Meng J, Wang J, Steinhoff M, Dolly JO (2016) TNF α induces co-trafficking of TRPV1/TRPA1 in VAMP1-containing vesicles to the plasmalemma via Munc18–1/syntaxin1/SNAP-25 mediated fusion. *Sci Rep* 6(1):21226. <https://doi.org/10.1038/srep21226>
- Mesch S, Walter D, Laux-Biehlmann A et al (2023) Discovery of BAY-390, a selective CNS penetrant chemical probe as transient receptor potential Ankyrin 1 (TRPA1) antagonist. *J Med Chem* 66(2):1583–1600. <https://doi.org/10.1021/acs.jmedchem.2c01830>
- Miyano K, Shiraishi S, Minami K et al (2019) Carboplatin enhances the activity of human transient receptor potential Ankyrin 1 through the cyclic AMP-protein Kinase A-A-Kinase anchoring Protein (AKAP) pathways. *Int J Mol Sci* 20(13). <https://doi.org/10.3390/ijms20133271>
- Moparathi L, Zygmunt PM (2020) Human TRPA1 is an inherently mechanosensitive bilayer-gated ion channel. *Cell Calcium* 91. <https://doi.org/10.1016/j.ceca.2020.102255>
- Moparathi L, Survery S, Kreir M et al (2014) Human TRPA1 is intrinsically cold- and chemosensitive with and without its N-terminal ankyrin repeat domain. *Proc Natl Acad Sci U S A* 111(47):16901–16906. <https://doi.org/10.1073/pnas.1412689111>
- Moparathi L, Kichko TI, Eberhardt M et al (2016) Human TRPA1 is a heat sensor displaying intrinsic U-shaped thermosensitivity. *Sci Rep* 6:28763. <https://doi.org/10.1038/srep28763>
- Moparathi L, Moparathi SB, Wenger J, Zygmunt PM (2020) Calcium activates purified human TRPA1 with and without its N-terminal ankyrin repeat domain in the absence of calmodulin. *Cell Calcium* 90:102228. <https://doi.org/10.1016/j.ceca.2020.102228>
- Moparathi L, Sinica V, Moparathi VK et al (2022) The human TRPA1 intrinsic cold and heat sensitivity involves separate channel structures beyond the N-ARD domain. *Nat Commun* 13(1):6113. <https://doi.org/10.1038/s41467-022-33876-8>
- Morgan K, Sadofsky LR, Morice AH (2015) Genetic variants affecting human TRPA1 or TRPM8 structure can be classified in vitro as ‘well expressed’, ‘poorly expressed’ or ‘salvageable’. *Biosci Rep* 35(5). <https://doi.org/10.1042/BSR20150108>
- Mukhopadhyay I, Gomes P, Aranake S et al (2011) Expression of functional TRPA1 receptor on human lung fibroblast and epithelial cells. *J Recept Signal Transduct Res* 31(5):350–358. <https://doi.org/10.3109/10799893.2011.602413>
- Nadezhdin KD, Neuberger A, Trofimov YA et al (2021) Structural mechanism of heat-induced opening of a temperature-sensitive TRP channel. *Nat Struct Mol Biol* 28(7):564–572. <https://doi.org/10.1038/s41594-021-00615-4>
- Naert R, Talavera A, Startek JB, Talavera K (2020) TRPA1 gene variants hurting our feelings. *Pflugers Arch* 472(7):953–960. <https://doi.org/10.1007/s00424-020-02397-y>
- Naert R, Lopez-Requena A, Talavera K (2021) TRPA1 expression and pathophysiology in immune cells. *Int J Mol Sci* 22(21). <https://doi.org/10.3390/ijms222111460>
- Nagatomo K, Kubo Y (2008) Caffeine activates mouse TRPA1 channels but suppresses human TRPA1 channels. *Proc Natl Acad Sci U S A* 105(45):17373–17378. <https://doi.org/10.1073/pnas.0809769105>
- Namer B, Seifert F, Handwerker HO, Maihofner C (2005) TRPA1 and TRPM8 activation in humans: effects of cinnamaldehyde and menthol. *Neuroreport* 16(9):955–959. <https://doi.org/10.1097/00001756-200506210-00015>
- Nassini R, Pedretti P, Moretto N et al (2012) Transient receptor potential ankyrin 1 channel localized to non-neuronal airway cells promotes non-neurogenic inflammation. *PLoS One* 7(8):e42454. <https://doi.org/10.1371/journal.pone.0042454>

- Naumov DE, Kotova OO, Gassan DA et al (2021) Effect of TRPM8 and TRPA1 polymorphisms on COPD predisposition and lung function in COPD patients. *J Pers Med* 11(2). <https://doi.org/10.3390/jpm11020108>
- Nguyen THD, Itoh SG, Okumura H, Tominaga M (2021) Structural basis for promiscuous action of monoterpenes on TRP channels. *Commun Biol* 4(1):293. <https://doi.org/10.1038/s42003-021-01776-0>
- Nikolaev YA, Cox CD, Ridone P et al (2019) Mammalian TRP ion channels are insensitive to membrane stretch. *J Cell Sci* 132(23). <https://doi.org/10.1242/jcs.238360>
- Nilius B, Appendino G, Owsianik G (2012) The transient receptor potential channel TRPA1: from gene to pathophysiology. *Pflugers Arch* 464(5):425–458. <https://doi.org/10.1007/s00424-012-1158-z>
- Nirenberg MJ, Chaouni R, Biller TM et al (2018) A novel TRPA1 variant is associated with carbamazepine-responsive cramp-fasciculation syndrome. *Clin Genet* 93(1):164–168. <https://doi.org/10.1111/cge.13040>
- Nozawa K, Kawabata-Shoda E, Doihara H et al (2009) TRPA1 regulates gastrointestinal motility through serotonin release from enterochromaffin cells. *Proc Natl Acad Sci U S A* 106(9):3408–3413. <https://doi.org/10.1073/pnas.0805323106>
- Nummenmaa E, Hamalainen M, Moilanen LJ et al (2016) Transient receptor potential ankyrin 1 (TRPA1) is functionally expressed in primary human osteoarthritic chondrocytes. *Arthritis Res Ther* 18(1):185. <https://doi.org/10.1186/s13075-016-1080-4>
- Obata K, Katsura H, Mizushima T et al (2005) TRPA1 induced in sensory neurons contributes to cold hyperalgesia after inflammation and nerve injury. *J Clin Invest* 115(9):2393–2401. <https://doi.org/10.1172/JCI25437>
- Oh MH, Oh SY, Lu J et al (2013) TRPA1-dependent pruritus in IL-13-induced chronic atopic dermatitis. *J Immunol* 191(11):5371–5382. <https://doi.org/10.4049/jimmunol.1300300>
- Patil MJ, Salas M, Bialuhin S et al (2020) Sensitization of small-diameter sensory neurons is controlled by TRPV1 and TRPA1 association. *FASEB J* 34(1):287–302. <https://doi.org/10.1096/fj.201902026R>
- Paulsen CE, Armache JP, Gao Y et al (2015) Structure of the TRPA1 ion channel suggests regulatory mechanisms. *Nature* 520(7548):511–517. <https://doi.org/10.1038/nature14367>
- Pereira I, Mendes SJ, Pereira DM et al (2017) Transient receptor potential Ankyrin 1 channel expression on peripheral blood leukocytes from rheumatoid arthritic patients and correlation with pain and disability. *Front Pharmacol* 8:53. <https://doi.org/10.3389/fphar.2017.00053>
- Prasad P, Yanagihara AA, Small-Howard AL et al (2008) Secretogranin III directs secretory vesicle biogenesis in mast cells in a manner dependent upon interaction with chromogranin A1. *J Immunol* 181(7):5024–5034. <https://doi.org/10.4049/jimmunol.181.7.5024>
- Pumroy RA, Samanta A, Liu Y et al (2019) Molecular mechanism of TRPV2 channel modulation by cannabidiol. *eLife* 8:e48792. <https://doi.org/10.7554/eLife.48792>
- Rosenbaum T, Gordon-Shaag A, Munari M, Gordon SE (2004) Ca²⁺/calmodulin modulates TRPV1 activation by capsaicin. *J Gen Physiol* 123(1):53–62
- Samad A, Sura L, Benedikt J et al (2011) The C-terminal basic residues contribute to the chemical- and voltage-dependent activation of TRPA1. *Biochem J* 433(1):197–204. <https://doi.org/10.1042/BJ20101256>
- Satoh T, Ohba A, Liu Z et al (2015) dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in *Drosophila* photoreceptors. *eLife* 4:e06306. <https://doi.org/10.7554/eLife.06306>
- Sawada Y, Hosokawa H, Hori A et al (2007) Cold sensitivity of recombinant TRPA1 channels. *Brain Res* 1160:39–46. <https://doi.org/10.1016/j.brainres.2007.05.047>
- Schmidt M, Dubin AE, Petrus MJ et al (2009) Nociceptive signals induce trafficking of TRPA1 to the plasma membrane. *Neuron* 64(4):498–509. <https://doi.org/10.1016/j.neuron.2009.09.030>
- Schmidt HR, Zheng S, Gurbinar E et al (2016) Crystal structure of the human σ 1 receptor. *Nature* 532(7600):527–530. <https://doi.org/10.1038/nature17391>

- Schmidt HR, Betz RM, Dror RO, Kruse AC (2018) Structural basis for $\sigma 1$ receptor ligand recognition. *Nat Struct Mol Biol* 25(10):981–987. <https://doi.org/10.1038/s41594-018-0137-2>
- Shigetomi E, Tong X, Kwan KY et al (2011) TRPA1 channels regulate astrocyte resting calcium and inhibitory synapse efficacy through GAT-3. *Nat Neurosci* 15(1):70–80. <https://doi.org/10.1038/nn.3000>
- Sinica V, Vlachova V (2021) Transient receptor potential ankyrin 1 channel: an evolutionarily tuned thermosensor. *Physiol Res* 70(3):363–381. <https://doi.org/10.33549/physiolres.934697>
- Sinica V, Zimova L, Barvikova K et al (2019) Human and mouse TRPA1 are heat and cold sensors differentially tuned by voltage. *Cells* 9(1). <https://doi.org/10.3390/cells9010057>
- Skerratt S (2017) Recent progress in the discovery and development of TRPA1 modulators. *Prog Med Chem* 56:81–115. <https://doi.org/10.1016/bs.pmch.2016.11.003>
- Sleczkowska M, Almomani R, Marchi M et al (2022) Peripheral ion channel genes screening in painful small fiber neuropathy. *Int J Mol Sci* 23(22). <https://doi.org/10.3390/ijms232214095>
- Song K, Wei M, Guo W et al (2021) Structural basis for human TRPC5 channel inhibition by two distinct inhibitors. *eLife* 10. <https://doi.org/10.7554/eLife.63429>
- Soriani O, Kourrich S (2019) The Sigma-1 receptor: When adaptive regulation of cell electrical activity contributes to stimulant addiction and cancer. *Front Neurosci* 13. <https://doi.org/10.3389/fnins.2019.01186>
- Sotomayor M, Corey DP, Schulten K (2005) In search of the hair-cell gating spring elastic properties of ankyrin and cadherin repeats. *Structure* 13(4):669–682. <https://doi.org/10.1016/j.str.2005.03.001>
- Staruschenko A, Jeske NA, Akopian AN (2010) Contribution of TRPV1-TRPA1 interaction to the single channel properties of the TRPA1 channel. *J Biol Chem* 285(20):15167–15177. <https://doi.org/10.1074/jbc.M110.106153>
- Story GM, Peier AM, Reeve AJ et al (2003) ANKTM1, a TRP-like channel expressed in nociceptive neurons, is activated by cold temperatures. *Cell* 112(6):819–829. [https://doi.org/10.1016/s0092-8674\(03\)00158-2](https://doi.org/10.1016/s0092-8674(03)00158-2)
- Sulast MA, Ghosh M, Sinharoy P et al (2018) Modulation of TRPA1 channel activity by Cdk5 in sensory neurons. *Channels* 12(1):65–75. <https://doi.org/10.1080/19336950.2018.1424282>
- Suo Y, Wang Z, Zubecevic L et al (2020) Structural insights into electrophile irritant sensing by the human TRPA1 channel. *Neuron* 105. <https://doi.org/10.1016/j.neuron.2019.11.023>
- Sura L, Zima V, Marsakova L et al (2012) C-terminal Acidic cluster is involved in Ca^{2+} -induced regulation of human transient receptor potential Ankyrin 1 channel. *J Biol Chem* 287(22):18067–18077. <https://doi.org/10.1074/jbc.M112.341859>
- Takahashi N, Kuwaki T, Kiyonaka S et al (2011) TRPA1 underlies a sensing mechanism for O_2 . *Nat Chem Biol* 7(10):701–711. <https://doi.org/10.1038/nchembio.640>
- Takaishi M, Uchida K, Fujita F, Tominaga M (2014) Inhibitory effects of monoterpenes on human TRPA1 and the structural basis of their activity. *J Physiol Sci* 64(1):47–57. <https://doi.org/10.1007/s12576-013-0289-0>
- Talavera K, Gees M, Karashima Y et al (2009) Nicotine activates the chemosensory cation channel TRPA1. *Nat Neurosci* 12(10):1293–1299. <https://doi.org/10.1038/nn.2379>
- Talavera K, Startek JB, Alvarez-Collazo J et al (2020) Mammalian transient receptor potential TRPA1 channels: from structure to disease. *Physiol Rev* 100(2):725–803. <https://doi.org/10.1152/physrev.00005.2019>
- Talbot BE, Vondorpe DH, Stotter BR et al (2019) Transmembrane insertases and N-glycosylation critically determine synthesis, trafficking, and activity of the nonselective cation channel TRPC6. *J Biol Chem* 294(34):12655–12669. <https://doi.org/10.1074/jbc.RA119.008299>
- Terrett JA, Chen H, Shore DG et al (2021) Tetrahydrofuran-based transient receptor potential Ankyrin 1 (TRPA1) antagonists: ligand-based discovery, activity in a rodent asthma model, and mechanism-of-action via cryogenic electron microscopy. *J Med Chem* 64(7):3843–3869. <https://doi.org/10.1021/acs.jmedchem.0c02023>
- Thakore P, Alvarado MG, Ali S et al (2021) Brain endothelial cell TRPA1 channels initiate neurovascular coupling. *Elife* 10. <https://doi.org/10.7554/eLife.63040>

- Ufret-Vincenty CA, Klein RM, Hua L et al (2011) Localization of the PIP₂ sensor of TRPV1 ion channels. *J Biol Chem* 286(11):9688–9698. <https://doi.org/10.1074/jbc.M110.192526>
- Vandewauw I, De Clercq K, Mulier M et al (2018) A TRP channel trio mediates acute noxious heat sensing. *Nature* 555(7698):662–666. <https://doi.org/10.1038/nature26137>
- Viana F (2016) TRPA1 channels: molecular sentinels of cellular stress and tissue damage. *J Physiol* 594(15):4151–4169. <https://doi.org/10.1113/JP270935>
- Voolstra O, Huber A (2014) Post-translational modifications of TRP channels. *Cells* 3(2):258–287. <https://doi.org/10.3390/cells3020258>
- Wang D, Sadée W, Quillan JM (1999) Calmodulin binding to G protein-coupling domain of opioid receptors. *J Biol Chem* 274(31):22081–22088. <https://doi.org/10.1074/jbc.274.31.22081>
- Wang S, Dai Y, Fukuoka T et al (2008a) Phospholipase C and protein kinase A mediate bradykinin sensitization of TRPA1: a molecular mechanism of inflammatory pain. *Brain* 131(Pt 5):1241–1251. <https://doi.org/10.1093/brain/awn060>
- Wang YY, Chang RB, Waters HN et al (2008b) The nociceptor ion channel TRPA1 is potentiated and inactivated by permeating calcium ions. *J Biol Chem* 283(47):32691–32703. <https://doi.org/10.1074/jbc.M803568200>
- Wang H, Schupp M, Zurborg S, Heppenstall PA (2013) Residues in the pore region of Drosophila transient receptor potential A1 dictate sensitivity to thermal stimuli. *J Physiol* 591(Pt 1):185–201. <https://doi.org/10.1113/jphysiol.2012.242842>
- Wang S, Kobayashi K, Kogure Y et al (2018) Negative regulation of TRPA1 by AMPK in primary sensory neurons as a potential mechanism of painful diabetic neuropathy. *Diabetes* 67(1):98–109. <https://doi.org/10.2337/db17-0503>
- Wang Z, Ye D, Ye J et al (2019) The TRPA1 channel in the cardiovascular system: promising features and challenges. *Front Pharmacol* 10:1253. <https://doi.org/10.3389/fphar.2019.01253>
- Wang X, Li Y, Wei H et al (2023) Molecular architecture and gating mechanisms of the Drosophila TRPA1 channel. *Cell Discov* 9(1):36. <https://doi.org/10.1038/s41421-023-00527-1>
- Weng HJ, Patel KN, Jeske NA et al (2015) Tmem100 is a regulator of TRPA1-TRPV1 complex and contributes to persistent pain. *Neuron* 85(4):833–846. <https://doi.org/10.1016/j.neuron.2014.12.065>
- Wilson SR, Gerhold KA, Bifulco-Fisher A et al (2011) TRPA1 is required for histamine-independent, Mas-related G protein-coupled receptor-mediated itch. *Nat Neurosci* 14(5):595–602. <https://doi.org/10.1038/nn.2789>
- Wright DJ, Simmons KJ, Johnson RM et al (2020) Human TRPC5 structures reveal interaction of a xanthine-based TRPC1/4/5 inhibitor with a conserved lipid binding site. *Commun Biol* 3(1):704. <https://doi.org/10.1038/s42003-020-01437-8>
- Xiao B, Dubin AE, Bursulaya B et al (2008) Identification of transmembrane domain 5 as a critical molecular determinant of menthol sensitivity in mammalian TRPA1 channels. *J Neurosci* 28(39):9640–9651. <https://doi.org/10.1523/JNEUROSCI.2772-08.2008>
- Yariv B, Yariv E, Kessel A et al (2023) Using evolutionary data to make sense of macromolecules with a “face-lifted” ConSurf. *Protein Sci* 32(3):e4582. <https://doi.org/10.1002/pro.4582>
- Yin Y, Wu M, Zubcevic L et al (2018) Structure of the cold- and menthol-sensing ion channel TRPM8. *Science* 359(6372):237–241. <https://doi.org/10.1126/science.aan4325>
- Zhang X, Li L, McNaughton PA (2008) Proinflammatory mediators modulate the heat-activated ion channel TRPV1 via the scaffolding protein AKAP79/150. *Neuron* 59(3):450–461. <https://doi.org/10.1016/j.neuron.2008.05.015>
- Zhang K, Julius D, Cheng Y (2021) Structural snapshots of TRPV1 reveal mechanism of polymodal functionality. *Cell* 184(20):5138–5150. e5112. <https://doi.org/10.1016/j.cell.2021.08.012>
- Zhao C, MacKinnon R (2023) Structural and functional analyses of a GPCR-inhibited ion channel TRPM3. *Neuron* 111(1):81–91. e87. <https://doi.org/10.1016/j.neuron.2022.10.002>
- Zhao J, Lin King JV, Paulsen CE et al (2020) Irritant-evoked activation and calcium modulation of the TRPA1 receptor. *Nature* 585(7823):141–145. <https://doi.org/10.1038/s41586-020-2480-9>
- Zhao Y, McVeigh BM, Moiseenkova-Bell VY (2021) Structural pharmacology of TRP channels. *J Mol Biol* 433(17):166914. <https://doi.org/10.1016/j.jmb.2021.166914>

- Zhou Y, Suzuki Y, Uchida K, Tominaga M (2013) Identification of a splice variant of mouse TRPA1 that regulates TRPA1 activity. *Nat Commun* 4(1):2399. <https://doi.org/10.1038/ncomms3399>
- Zima V, Witschas K, Hynkova A et al (2015) Structural modeling and patch-clamp analysis of pain-related mutation TRPA1-N855S reveal inter-subunit salt bridges stabilizing the channel open state. *Neuropharmacology* 93:294–307. <https://doi.org/10.1016/j.neuropharm.2015.02.018>
- Zimova L, Sinica V, Kadkova A et al (2018) Intracellular cavity of sensor domain controls allosteric gating of TRPA1 channel. *Sci Signal* 11(514). <https://doi.org/10.1126/scisignal.aan8621>
- Zimova L, Barvikova K, Macikova L et al (2020) Proximal C-terminus serves as a signaling hub for TRPA1 channel regulation via its interacting molecules and supramolecular complexes. *Front Physiol* 11:189. <https://doi.org/10.3389/fphys.2020.00189>
- Zurborg S, Yurgionas B, Jira JA et al (2007) Direct activation of the ion channel TRPA1 by Ca^{2+} . *Nat Neurosci* 10(3):277–279. <https://doi.org/10.1038/nn1843>
- Zygmunt PM, Hogestatt ED (2014) TRPA1. *Handb Exp Pharmacol* 222:583–630. https://doi.org/10.1007/978-3-642-54215-2_23

Chapter 11

Structural Variations and Rearrangements in Bacterial Type II Toxin-Antitoxin Systems



Maria Raadkjær Nielsen and Ditlev E. Brodersen

Abstract Bacteria encode a wide range of survival and immunity systems, including CRISPR-Cas, restriction-modification systems, and toxin-antitoxin systems involved in defence against bacteriophages, as well as survival during challenging growth conditions or exposure to antibiotics. Toxin-antitoxin (TA) systems are small two- or three-gene cassettes consisting of a metabolic regulator (the “toxin”) and its associated antidote (the “antitoxin”), which also often functions as a transcriptional regulator. TA systems are widespread in the genomes of pathogens but are also present in commensal bacterial species and on plasmids. For mobile elements such as plasmids, TA systems play a role in maintenance, and increasing evidence now points to roles of chromosomal toxin-antitoxin systems in anti-phage defence. Moreover, the widespread occurrence of toxin-antitoxin systems in the genomes of pathogens has been suggested to relate to survival during host infection as well as in persistence during antibiotic treatment. Upon repeated exposure to antibiotics, TA systems have been shown to acquire point mutations as well as more dramatic rearrangements such as in-frame deletions with potential relevance for bacterial survival and pathogenesis. In this review, we present an overview of the known functional and structural consequences of mutations and rearrangements arising in bacterial toxin-antitoxin systems and discuss their relevance for survival and persistence of pathogenic species.

Keywords Plasmid maintenance · HipAB · CcdAB · VapXD · VapBC · RelBE · Conditional cooperativity

M. R. Nielsen · D. E. Brodersen (✉)

Department of Molecular Biology and Genetics, Aarhus University, Aarhus C, Denmark

e-mail: mrn@mbg.au.dk; deb@mbg.au.dk

© The Author(s), under exclusive license to Springer Nature

Switzerland AG 2024

J. R. Harris, J. Marles-Wright (eds.), *Macromolecular Protein Complexes V*, Subcellular Biochemistry 104, https://doi.org/10.1007/978-3-031-58843-3_11

245

Introduction

Bacteria harbour a wide range of defence and resistance systems that ensure survival of the population when encountering threats such as antibiotics, nutritional and oxidative stress, viral attack, or temperature fluctuations. For antibiotics, specific resistance enzymes, such as the β -lactamases that protect against penicillin-type antibiotics, have been known since the discovery of antibiotics themselves (Bradford 2001). Bacterial strains that encode proteins or enzymes counteracting the effect of specific antibiotics are said to be *antibiotic resistant*. Since beginning to use antibiotics to treat bacterial infections in the mid-twentieth century, the constant threat of antibiotic resistance has loomed as a dark cloud over infection medicine. Today, this challenge is greater than ever, and it is estimated that by 2050, ten million people will succumb due to resistant infections each year (O'Neill 2016). This is an immense challenge that calls for immediate action and should motivate our search for novel ways to fight bacterial infections. To succeed in this endeavour, we need a deep, molecular understanding of bacterial defence mechanisms, because only by knowing these we can devise ways of avoiding them. Recently, a wealth of hitherto uncharacterised bacterial defence systems mainly against bacteriophages (phages, viral particles infecting bacteria) have been discovered (Kelly et al. 2023). Phages are the most abundant biological entity on the planet (Brüssow and Hendrix 2002) and contain a large arsenal of advanced, molecular systems that have evolved to fight and kill bacteria. Therefore, phage genomes constitute a goldmine of potential antibiotics and ways to circumvent bacterial defence systems, and the idea of using phages for treating bacterial infections in humans has therefore recently been revived (Brives and Pourraz 2020). In a tit-for-tat fight for survival, bacteria have on their part evolved advanced defence mechanisms to counter phage attack, including restriction-modification (RM) systems, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas adaptive immune systems, and toxin-antitoxin (TA) systems. In addition, genome mining of so-called defence islands based on the “guilt-by association” principle combined with experimental validation has led to the identification of a wide range of novel phage recognition and response systems, the mechanisms of which are yet to be characterised (LeRoux et al. 2022; Bobonis et al. 2022; Fineran et al. 2009). In this review, we focus on the large and diverse class of toxin-antitoxin systems, many of which are well characterised, and discuss their role in bacterial survival (for an overview, see Table 11.1).

Toxin-antitoxin (TA) systems are small gene clusters originally discovered for their role in plasmid maintenance (Ogura and Hiraga 1983). Many TA systems are located on plasmids or other mobile genetic elements and can therefore migrate between organisms by horizontal gene transfer (Pandey and Gerdes 2005). TA loci are characterised by encoding a “toxic” protein, the function of which is inhibited during normal growth by an antitoxin, which can be either RNA or protein, by one of several mechanisms (for a review, see Harms et al. 2018). The various ways in which antitoxins inhibit their cognate toxins have been used to divide TA systems into at least six types, of which type II is the most studied and most abundant class (Page and Peti 2016). In this class, the antitoxin is expressed as a protein, and

Table 11.1 Overview of selected functional variations of type II toxin-antitoxin systems. The table lists the name of the TA system, the protein in which the variation occurs (toxin or antitoxin), details about the variation, as well as effects on toxin activity, toxin inhibition, DNA-complex interactions, and observed phenotype

TA system	Organism	Component	Observed variation	Toxin activity	Toxin inhibited?	DNA-complex interaction?	Phenotype
<i>hipBA7</i>	<i>E. coli</i>	Toxin	G22S, D291A	Less active	No	Impaired	High Persistence
<i>ccdAB_{ST}</i>	<i>S. Typhimurium</i>	Toxin	R99W	Inactive	Yes	Yes	Transcriptional Regulation
<i>vapXD_{NTH}</i>	<i>H. influenzae</i>	Toxin	46 amino acid deletion	Inactive	Unknown	Unknown	Dominant Negative
<i>vapBC12</i>	<i>M. tuberculosis</i>	Toxin	H32R	Decreased	Unknown	Unknown	Potential increased virulence or persistence
<i>vapBC37</i>	<i>M. tuberculosis</i>	Toxin	T16A	Increased	Unknown	Unknown	Potential increased virulence or persistence
<i>vapBC38</i>	<i>M. tuberculosis</i>	Toxin	V48A	Decreased	Unknown	Unknown	Potential increased virulence or persistence
<i>vapBC</i>	<i>E. coli</i>	Antitoxin	T3N/V5E/L7P/A13P/L16R/V20G/N-terminal inversion	Active	Expected	Unknown	Potential antibiotic tolerance
<i>relBE3_{ST}</i>	<i>S. Typhimurium</i>	Antitoxin	A14P	Active	No	Impaired	Unknown

inhibition of the toxin takes place via direct protein-protein interaction between toxin and antitoxin. However, the toxin can also be inhibited by binding of the antitoxin RNA (type III), inhibition of translation of the toxin mRNA (types I and V), or preventing interaction of the toxin with its cellular target (type IV) (Harms et al. 2018). The two genes encoding toxin and antitoxin are usually configured as an operon regulated by a single promoter, and protein-coding genes are often translationally coupled via overlapping open reading frames. For class II TA systems, a stable toxin and an unstable antitoxin are both translated into proteins. Some protein antitoxins contain intrinsically disordered protein (IDP) regions, which stabilise upon interaction with the toxin. This both stabilises the antitoxin and inhibits toxin activity in the stable TA complex (Fig. 11.1). In addition to the toxin-interacting domain, most type II antitoxins also contain a DNA binding domain, which folds through dimerisation and allows interaction with the TA promoter and, thus, transcriptional regulation of the operon (Harms et al. 2018). In many instances, the free antitoxin dimer is able to bind the promoter with lower affinity than the more stable TA complex, which allows for fine-grained transcriptional control in response to the ratio of toxin to antitoxin by a mechanism known as “conditional cooperativity” (Cataudella et al. 2012). Due to the intrinsic instability of the antitoxin, reduced antitoxin expression upon plasmid loss or targeted antitoxin degradation can lead to an excess of toxin and thus toxin activation in the cell. The active toxins are enzymes and belong to a range of different groups, including transferases active in post-translational modification of proteins (phosphorylation/ADP-ribosylation) and RNases involved in tRNA/mRNA/rRNA degradation. Although the effects of toxin activation are difficult to measure correctly in vivo and still debated, in most instances, it is believed that activation leads to growth arrest or even, in some cases, cell death (Song and Wood 2020).

TA systems have been implicated in both plasmid maintenance and survival during phage attack, but their biological roles remain a hot topic of debate. The *ccdAB* system is a classic example of a plasmid maintenance system and found on the *Escherichia coli* fertility factor F plasmid (Ogura and Hiraga 1983). For many plasmid-borne systems, the model states that plasmid loss is associated with a fitness cost, because continuous production of the unstable antitoxin is required to keep the stable toxin at bay. Thus, if a daughter cell does not receive a copy of the plasmid upon cell division, progressive toxin activation due to lack of antitoxin expression will result in growth reduction thus favouring cells containing the plasmid (Fig. 11.2). This model was first described as “post-segregational killing” based on experiments with the *hok/sok* genes of the *Salmonella paratyphi* R1 plasmid (Gerdes et al. 1986; Cox and Schildbach 2017), but maintenance systems can in principle function without killing plasmid-free cells simply by favouring those that harbour the plasmid. Plasmid-encoded TA systems are sometimes described as addiction modules that make cells addicted to a certain plasmid. Interestingly, this also means that chromosomal TAs can function as anti-addiction modules by encoding the corresponding antitoxin of a plasmid-encoded toxin and thus relieving the pressure to maintain the plasmid (Saavedra De Bast et al. 2008; Van Melderden and Saavedra De Bast 2009). For some chromosomally encoded TA systems, it has been

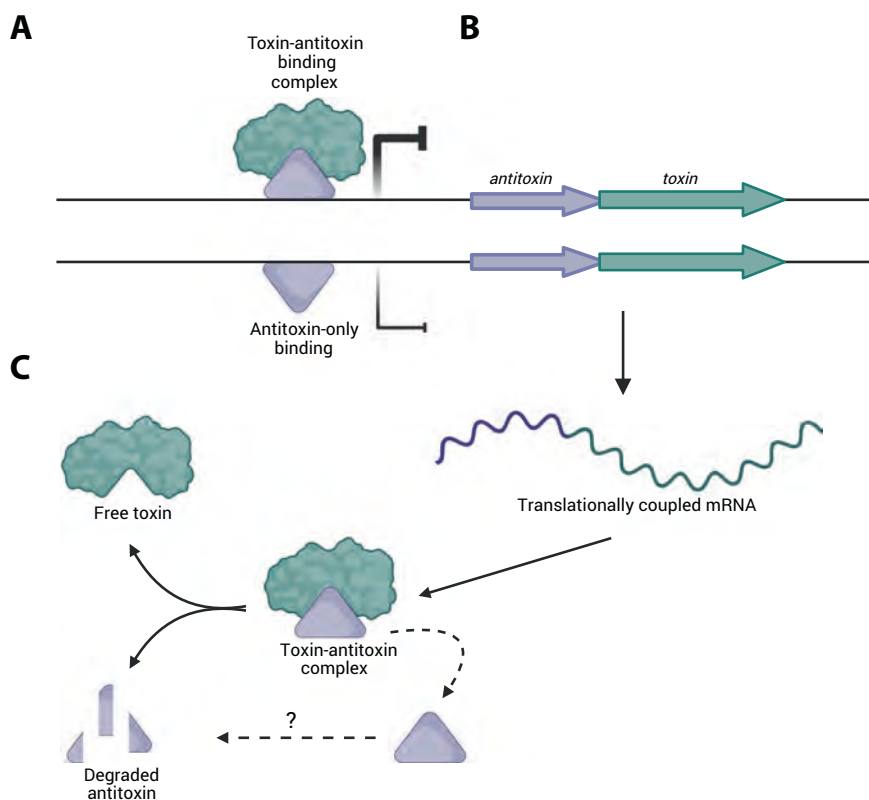


Fig. 11.1 Overview of type II toxin-antitoxin (TA) systems. **(a)** Transcription from TA loci is often regulated by a DNA-binding domain on the antitoxin, which binds the promoter either on its own (bottom) or in complex with the toxin (TA complex, top). Usually, transcriptional repression is stronger when the TA complex binds than antitoxin alone. **(b)** TA genes are transcribed from the same promoter, and protein-coding genes are often translationally coupled. **(c)** Upon expression, the antitoxin will tightly bind the toxin forming the TA complex, which inhibits the toxin. If at a later stage the antitoxin is degraded (in complex or after dissociation) or its expression is prevented, free and active toxin will be present in the cell. Created with BioRender

shown that degradation of antitoxins does not necessarily activate toxins but rather causes transcriptional de-repression and renewed production of both toxin and antitoxin (LeRoux et al. 2020).

Indications of a role for TA systems in phage defence came early on for the *hok/sok* system encoded on the *Salmonella paratyphi* R1 plasmid, which was shown to protect against T4 phage infection by activation of the Hok toxin, which causes cell lysis and altruistic suicide by a mechanism known as *abortive infection* (Pecota and Wood 1996). So far, only a few molecular mechanisms for the involvement of TA systems in phage defence have been described in detail, but more are currently being discovered at dramatic pace. One example is the ToxIN system that is based on mRNA degradation by the toxin causing inhibition of translation by the host,

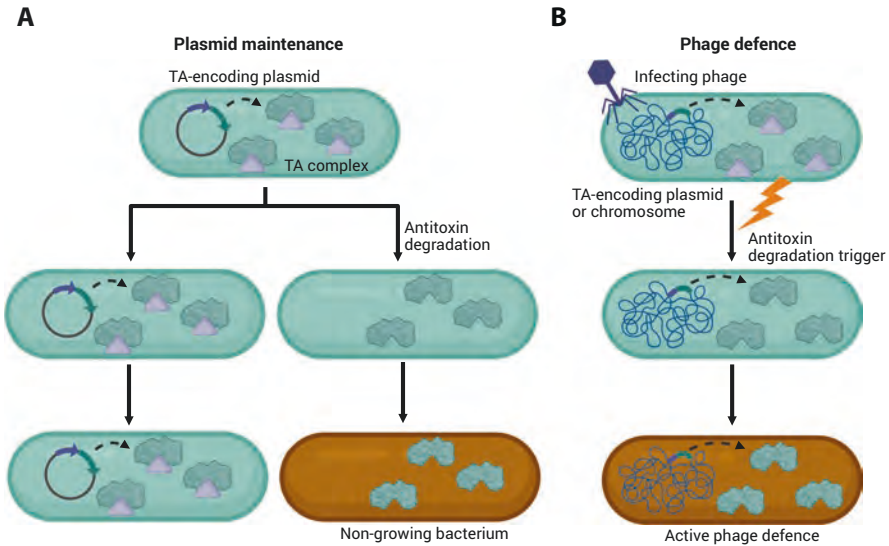


Fig. 11.2 Functions of type II TA systems. (a) Plasmid maintenance. A bacterial cell containing a plasmid encoding a TA system will express both toxin and antitoxin. If, during cell division, the plasmid is lost from one daughter cell, the instability of the antitoxin will result in the presence of free toxin and growth arrest or reduced fitness. (b) TA systems functioning in phage defence can be triggered in several ways by the presence of phage resulting in degradation of the antitoxin. Activation of the toxin will then cause cell death or growth arrest, effectively halting phage propagation. Created with BioRender

potentially causing abortive infection. Another example is DarTG in which phage DNA is modified by ADP-ribosylation, preventing its replication and thus phage propagation (LeRoux and Laub 2022; Fineran et al. 2009; LeRoux et al. 2022). TA systems have also been implicated in bacterial persister cell formation, whereby a subpopulation of cells exhibits an altered metabolism and slower kill rates in the presence of antibiotics but without any genomic mutations (Brauner et al. 2016). However, the significance of the observations supporting an involvement of TA systems in persister cell formation is debated (Song and Wood 2020). TAs are found in both commensal and pathogenic bacterial species as well as archaea and phages, but the number of loci in pathogens is often higher than in commensals (LeRoux and Laub 2022; Bendtsen and Brodersen 2017). For example, *Mycobacterium tuberculosis* contain a staggering 88 TA loci, whereas its non-pathogenic counterpart, *Mycobacterium smegmatis*, encodes a mere three (*vapBC*, *mazEF*, *phd/doc*) (Ramage et al. 2009; Jennifer et al. 2009). However, this pattern is not universal, and no clear correlation between the types and number of TA loci with bacterial lifestyle and host interaction has been found. TA systems are highly diverse and are generally clustered into families based on the protein fold and enzymatic activity of the toxin since sequence conservation can be quite low, even for closely related systems (Guglielmini and Van Melder 2011). In particular, antitoxins within the same TA family can vary extensively and are sometimes more similar to antitoxins

associated with other families, suggesting that interactions are modular and that gene shuffling takes place (Guglielmini and Van Melderden 2011). In some cases, point mutations and/or deletions have been observed between close relatives that often relate to functionality and sometimes antibiotic tolerance. In this review, we compare variations of related TA systems with a focus on the function of point mutations and deletions and discuss potential evolutionary benefits of these as well as the physiological benefit of encoding one or the other TA system.

Mutations and Deletions in Type II Toxins

Mutations in the HipA Kinase Increase Persister Cell Formation In 1983, several so-called High persistence (Hip) strains of *E. coli* were described, which displayed several orders of magnitude increase in the number of cells insensitive to inhibitors of murein synthesis (Moyed and Bertrand 1983). The phenotypes correlated with two mutations, *hipA7* and *hipA9*, which defined a new gene, *hipA* (Moyed and Bertrand 1983). *hipA* is encoded downstream of another gene *hipB*, and both are expressed as proteins (Scherrer and Moyed 1988). In *E. coli*, HipA was found to be a functional serine/threonine protein kinase specifically targeting glutamyl-tRNA synthetase (GltX) (Germain et al. 2013), and phosphorylation of the synthetase blocks aminoacylation of glutamyl-tRNA leading to translational inhibition and a toxic phenotype (Germain et al. 2013). HipB was shown to function as a type II antitoxin by directly binding and inhibiting HipA (Schumacher et al. 2009). The toxin-antitoxin interaction takes place via multiple hydrogen bonds between HipB and HipA (Schumacher et al. 2009). This interaction is important for inactivation of HipA toxicity and at the same time help form the DNA-binding higher-order TA complex (Schumacher et al. 2009). The higher-order structure of the HipBA complex consists of a HipB homodimer forming the DNA-interacting domain, with each HipB chain also interacting with one inhibited HipA monomer (Fig. 11.3) (Schumacher et al. 2009). HipA can likewise form homodimers and allows the complex to polymerise along DNA through interactions between HipA in one TA complex and HipA in the next complex (Schumacher et al. 2015). In an additional layer of regulatory complexity, HipA can auto-phosphorylate itself in *trans* at S150, which is important for regulation of kinase activity as auto-phosphorylation inactivates the kinase (Schumacher et al. 2012). Structural studies have furthermore shown that auto-phosphorylation results in ejection of the so-called Gly-rich loop (the structural equivalent of the P loop) from the active site of the kinase, which prevents ATP binding (Schumacher et al. 2012). The HipA7 variant found in high persister phenotypes contains two mutations, G22S and D291A (Korch et al. 2003), and is found less in the auto-phosphorylated state (10%) than wild type (wt) HipA (90%) (Semanjski et al. 2018). Since phosphorylation inactivates the kinase, this means that only 10% of HipA is in an active state while 90% of HipA7 is active. However, this is countered by wt HipA being enzymatically more active than HipA7. In vivo, wt HipA was found to have phosphorylated 76% of its target (GltX)

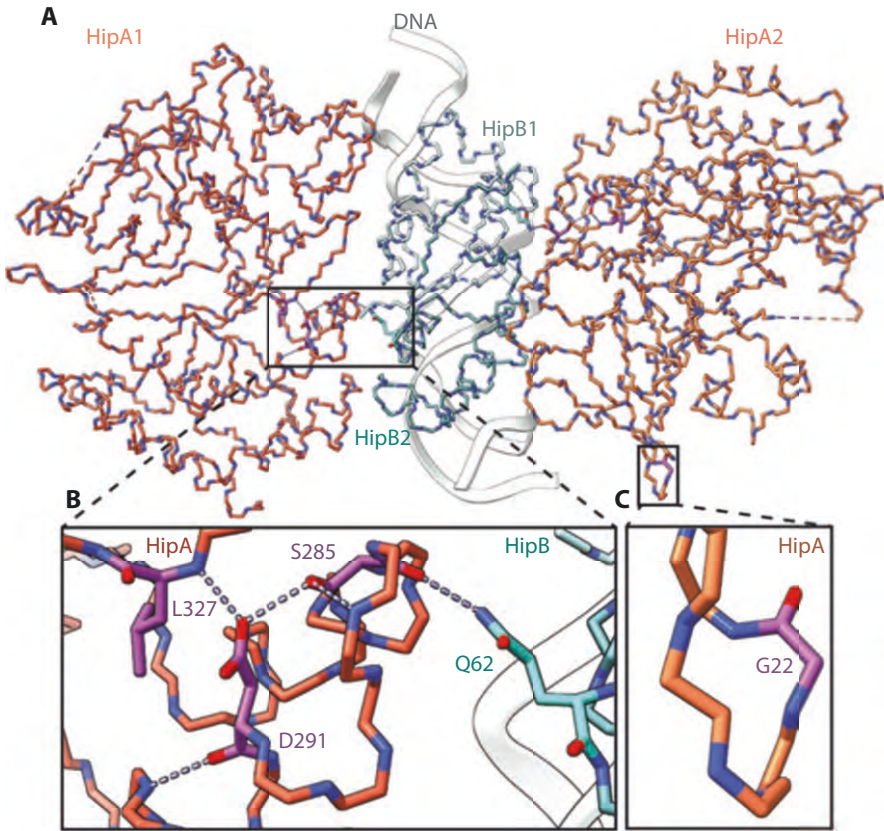


Fig. 11.3 Structure of the DNA-bound form of *E. coli* HipBA. (a) Overall structure of the DNA-bound HipBA complex shown in backbone representation, where HipB (light blue) forms a homodimer bound to DNA (white) (PDB: 5 K98) (Schumacher et al. 2015). (b) Aspartic acid D291 (magenta) in HipA (orange) positions the main chain carbonyl group of S285 to interact with Q62 from HipB. (c) Glycine G22 (magenta) in HipA (orange) is proposed to be important for dimerisation of HipA and higher-order structure of the HipBA complex on DNA (not shown)

after 95 minutes of induction, compared to 48% for HipA7, suggesting that HipA7 is a less effective kinase (Semanjski et al. 2018). On the other hand, HipA7 appears to catalyse fewer off-target phosphorylation reactions than wt HipA, demonstrating the complexity of these regulatory networks (Semanjski et al. 2018). Of the two mutations in HipA7, G22S renders HipA nontoxic but not inactive, while the D291A mutation alone is responsible for the high persistence phenotype (Korch et al. 2003). Evidence suggests that D291 is important for positioning the carbonyl group of S285 to interact with Q62 on HipB and consequently that the D291A mutation weakens the interaction between toxin and antitoxin (Fig. 11.3) (Schumacher et al. 2009). The relevance of the G22S mutation is not known, but it has been hypothesised that this residue is important for HipA dimerisation (Schumacher et al. 2015). Another possible effect of G22S is the introduction of a new serine and potential

auto-phosphorylation site (Semanjski et al. 2018), but the implications of this have not been studied. It is also not fully understood how *hipA7* might lead to antibiotic persistence. However, it has been speculated that the more active wt kinase is likely more toxic once activated and hence will be more difficult to counteract in order to resume growth (Semanjski et al. 2018). The less active HipA7, on the other hand, could lead to a dormant state that is easier to revert from. Even though the *hipBA* locus is named after its high persistence phenotype, it seems that only the mutant alleles (*hipA7* and *hipA9*) show higher persistence, which might be due to inherent heterogeneous growth of cells (Balaban et al. 2004). This leaves open the possibility that wt *hipBA* has another function entirely which is supported by the observation that deletion of *hipBA* has no effect on persistence (Hansen et al. 2008). Moreover, the *hipA7* and *hipA9* alleles were found in an artificial laboratory experiment and have not been found in free-living bacteria, raising questions as to their biological relevance. Finally, a wide range of HipA-like kinases have been found in bacterial genomes, some of which are coupled to various DNA binding and recombination domains, suggesting that a role in recognition of foreign (i.e. phage) DNA might be worth investigating (Gerdes et al. 2021).

Inactivation of the CcdB Toxin by Mutation of an Active Site Residue *Escherichia coli* and *Salmonella enterica* serovar Typhimurium are believed to derive from a common ancestor that existed around 140 million years ago (Ochman and Wilson 1987). Even today, the two species carry similar plasmids related to the F plasmid of *E. coli*, and high homology between certain genes suggests that genetic exchange between these two species have occurred (Boyd and Hartl 1997). The F (fertility) plasmid is a circular piece of extrachromosomal DNA of 100 kb present in a low copy number in the cell (Boyd and Hartl 1997), maintenance of the plasmid is known to partly depend on the type II TA system *ccdAB*, which was discovered due to its coupling to cell division (*ccd* = coupled cell division) (Ogura and Hiraga 1983). Similarly, *S. enterica* Typhimurium contains the F-like virulence plasmid (pSLT), roughly 90 kb in size and also present in low copy numbers (Lobato-Marquez et al. 2016; McClelland et al. 2001). Here, plasmid maintenance is conferred by the *parAB* partitioning system, which helps plasmids segregate equally into daughter cells, and partly by the Type II TA systems *ccdAB_{ST}* and *vapBC2_{ST}* (Lobato-Marquez et al. 2016). Of these, the *ccdAB* system encodes the CcdB toxin, which inhibits DNA gyrase upon activation (Bernard and Couturier 1992) and the antitoxin CcdA that directly binds and inhibits CcdB. CcdA is a known target of the bacterial Lon protease, which has been proposed as a means for toxin activation (Van Melderden et al. 1994). Degradation by Lon usually leads to CcdB activation, but in the case of CcdAB_{ST}, surprisingly, the toxin is inactive (Lobato-Márquez et al. 2015, 2016). In this variant, the important active site residue W99 has been replaced by a hydrophilic arginine, which only affects the toxin activity and not toxin-antitoxin interactions (Lobato-Marquez et al. 2016). Based on structural analysis of the CcdB structure, it is not expected that the W99R mutation affects overall fold or stability of the toxin (Fig. 11.4). Lobato-Márques et al. hypothesised that this mutation arose to maintain the transcriptional regulation of the TA locus while not

requiring toxicity (Lobato-Marquez et al. 2016). This would make it possible for the TA system to function as a tight transcriptional regulator but without the ability of activating the toxin. Transcriptional regulation by TA systems can be highly complex and include so-called conditional cooperativity, a mechanism in which the affinity of the antitoxin for DNA increases with toxin binding at lower toxin/antitoxin (<1.0) ratios but decreases at higher (>1.0) ratios (Overgaard et al. 2008; Loris and Garcia-Pino 2014). This phenomenon could explain the benefit of maintaining an inactive toxin rather than losing the gene completely because the toxin would still be required to maintain a high level of transcriptional repression by the TA complex. On the virulence plasmid of *Salmonella* Typhimurium, the *ccdAB* locus is part of an operon containing four downstream genes, *pSLT029*, *pSLT030*, *rsdB*, and *pSLT032*, which could be regulated by the TA system (Fig. 11.4), supporting the hypothesis (McClelland et al. 2001; Lobato-Marquez et al. 2016). Interestingly, it has been shown that reverting CcdB_{ST} to an active toxin neither increases nor decreases plasmid maintenance by this TA system, indicating that transcriptional regulation (of other genes) likely is the most important function (Lobato-Marquez et al. 2016). It is currently not known how CcdAB_{ST} can function in plasmid main-

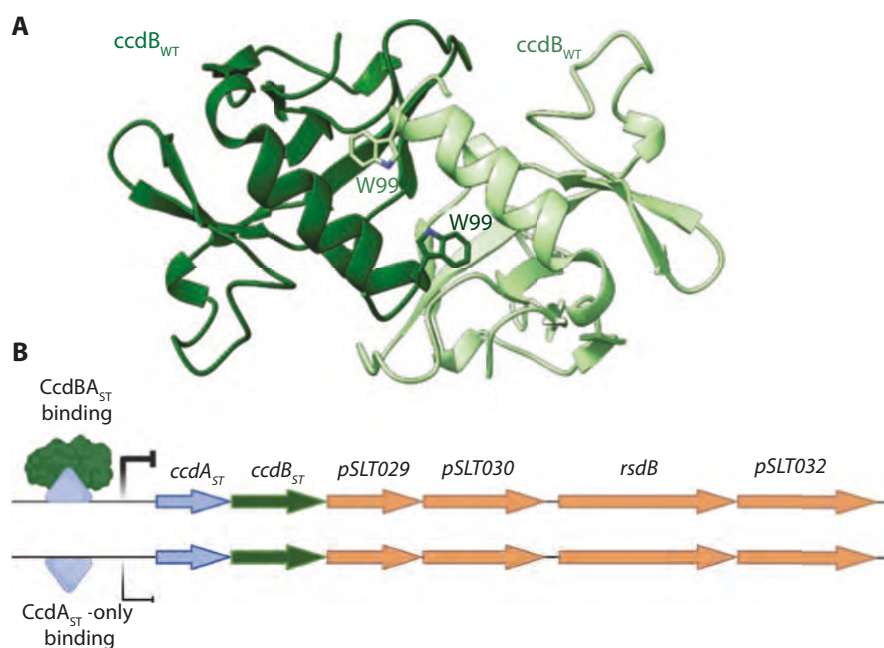


Fig. 11.4 Structure of the CcdB toxin. (a) An experimentally determined crystal structure of CcdB (PDB: 1VUB) (Loris et al. 1999). (b) Overview of the *Salmonella* Typhimurium *ccdAB_{ST}* operon showing the four downstream genes in orange. The cartoon on the bottom illustrates that the CcdAB_{ST} complex will inhibit transcription stronger at low toxin concentration than the antitoxin alone. Partly created with BioRender

tenance with an inactive toxin, but the answer might lie in the four downstream genes and their functions, which are also unexplored (McClelland et al. 2001).

An In-Frame Deletion in the *vapD* Toxin Results in a Dominant Negative Phenotype Non-typeable (non-encapsulated) *Haemophilus influenzae* (NTHi) is a human pathogen known to infect the ear and respiratory tract and causing a range of disorders, including otitis media, sinusitis, and pulmonary disease, particularly in infants (Murphy 2003). *H. influenzae* encodes four TA systems, *toxAvapA*, *vapBC-1*, *vapBC-2*, and *vapXD*, of which *vapBC* (see below) and *vapXD* loci encode type II TA systems. The toxin encoded by *vapXD*, VapD, is unusual because in some organisms it is found as an orphan toxin gene, i.e. in the absence of an antitoxin, suggesting it is a non-classical type II toxin (Daines et al. 2004; Yee et al. 2023; Kwon et al. 2012). The VapD protein has a Cas2-like RNase fold and has been reported to display non-specific endoribonuclease activity in vitro (Fig. 11.5) (Makarova et al. 2012; Kwon et al. 2012; Ren et al. 2012). The first structural insights into VapD came from studies of the protein expressed from an orphan locus in *Helicobacter pylori*, which revealed a homodimeric structure consistent with the dimer seen for Cas2 inside the Cas1-Cas2 complex (Kwon et al. 2012; Nuñez et al. 2014). Intriguingly, VapD from *H. pylori* is encoded in an operon along with another protein, but this does not appear to be an antitoxin (Kwon et al. 2012). The subsequent structural analysis of *H. influenzae* VapD in complex with its cognate VapX antitoxin revealed a similar fold to *H. pylori* VapD despite the bound antitoxin (Bertelsen et al. 2021). However, the structure also revealed a unique TA complex higher-order structure in which VapD and VapX interact in an unusual 2:1 complex, with VapX

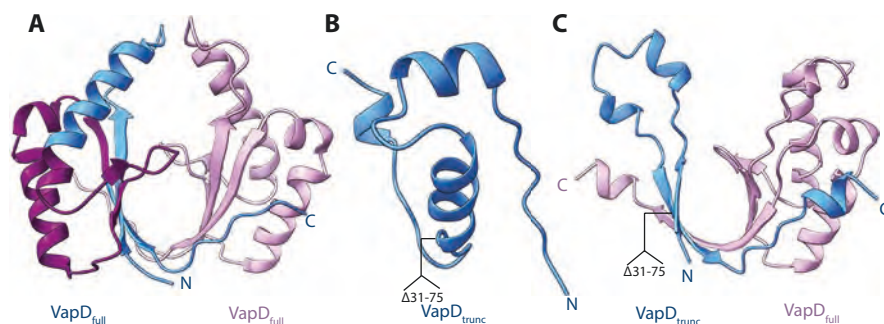


Fig. 11.5 The structure of VapD and variants. (a) Experimentally determined crystal structure of the VapD dimer from *Haemophilus influenzae* (PDB: 6ZI0) (Bertelsen et al. 2021) with chain A (blue and purple) and chain B (pink). The residues deleted in the truncated VapD are represented in chain A in purple. (b) AlphaFold2 prediction of the truncated VapD_{trunc} monomer (pLDDTs = 86.26) (Jumper et al. 2021) with Δ31–75 indicating the point of deletion. (c) AlphaFold2 prediction of the heterodimer between full-length VapD_{full} (pink) and truncated VapD_{trunc} (blue) (ipTM+pTM = 0.70) (Jumper et al. 2021). VapD_{full} is predicted to fold in a similar way to VapD in the experimental structure, while the truncated monomer is predicted to adapt its structure compared to the isolated VapD_{trunc} upon interacting to mimic the fold of full-length VapD. Overall, this means that the region including the deletion changes from α-helical to β-sheet structure

intercalating between two VapD toxin molecules. This heterotrimeric complex again forms a ring-shaped and nonameric structure consisting of three VapXD₂ complexes, which was found in solution as well (Bertelsen et al. 2021). The structure also showed that VapX displays an Oligonucleotide/Oligosaccharide-Binding (OB) fold and unlike other known antitoxins does not contain an IDP region, nor a classical DNA-binding domain (Bertelsen et al. 2021). Interestingly, some clinically relevant isolates of *H. influenzae* encode a truncated version of VapD, in which 46 residues have been removed due to an in-frame deletion (residues 31–75) (Daines et al. 2004). This truncated version of VapD (VapD_{trunc}) cannot form a homodimer but appears to be able to form a heterodimer with full-length VapD (VapD_{full}). Moreover, this interaction between a truncated and full-length toxin results in a dominant negative phenotype in vivo, whereby the toxicity of wt VapD is countered in a similar way to when antitoxin binds (Fig. 11.5). In silico folding shows that the predicted fold of VapD_{trunc} likely changes in the presence of a full-length VapD in a way that makes the heterodimer similar to the wt homodimer (Fig. 11.5). This supports the observed dominant negative phenotype and suggests that the truncated protein is able to sequester the wt toxin in an inactive heterodimer. Importantly, the truncation appears to have arisen multiple times in independent clinical isolates clearly pointing to a potential functional relevance of this variant (Daines et al. 2004). Since the truncated variant inhibits VapD_{full}, it can be speculated that the deletion arose as a way to counteract the conjugation of plasmids encoding VapD, i.e. by functioning as an anti-addiction module. It is also possible that this mutation was selected to completely inhibit the toxicity of VapD.

Variations in the VapC Toxin Affect *Mycobacterium tuberculosis* Virulence The *vapBC* family is one of the most widespread families of type II TA systems (Pandey and Gerdes 2005). This type of TA system is found in a range of bacterial organisms ranging from *E. coli* to *M. tuberculosis* and even in some archaea (Bendtsen and Brodersen 2017). Most impressively, *M. tuberculosis* encodes at least 47 *vapBC* systems out of a total of 88 identified TA systems in its genome (Ramage et al. 2009). It is currently not known what the advantage of encoding such a vast number of TA systems might be nor why having such a high redundancy of similar TA systems is beneficial. The *vapBC* TA system consists of the antitoxin *vapB* encoding a DNA-binding protein, which forms a homodimer and its associated toxin, *vapC*, which encodes a homodimeric toxin, and together, this dual dimerisation results in a large, heterooctameric VapB₄C₄ protein complex (Dienemann et al. 2011). The VapC toxin belongs to the family of PilT N-terminal (PIN) domain proteins, which are found across all domains of life and often associated with RNase activity (Levin et al. 2004). This is also the case for VapC, which displays endoribonucleolytic activity and inhibits protein translation through degradation of specific transfer or ribosomal RNA species, including tRNA and rRNA (Winther et al. 2016). Detailed studies of VapC targets have revealed that each toxin is highly specific for a single RNA species providing a possible rationale for the large number of loci in, e.g. *M. tuberculosis* (Winther et al. 2016). Several structures of VapBC complexes from a range of organisms have been determined by X-ray crystallography, generally

showing similar heterooctameric structures with a 1:1 molecular interaction between VapB and VapC, with VapB inactivating its cognate VapC active site (Dienemann et al. 2011; Martyn et al. 2022; Park et al. 2020). Interestingly, however, in the VapBC complexes from *Caulobacter crescentus* and *Rickettsia felis*, which show heterooctameric structures similar to other VapBC complexes, two VapB antitoxins were found to use their extended C-terminal IDP regions to inhibit all four VapC active sites, leaving two of the antitoxins in the complex unused for VapC inactivation (Bendtsen et al. 2017; Maté et al. 2012). Moreover, in *Caulobacter crescentus*, upon interaction with DNA, the C-terminal tails of all four VapB antitoxins were found to swap positions, so that the previously unused VapB termini inactivate the four VapC toxins in the DNA-bound state, while the other VapB antitoxins have free C-terminal extensions (Bendtsen et al. 2017). This inactivation mechanism was hypothesised to depend on near-perfect palindromic protein sequence in the VapB C-terminus that allows a single antitoxin tail to interact with two adjacent VapC active sites that are related by a twofold symmetry. The palindromic motif was also found to be present in other VapB proteins, suggesting this principle might be more widespread than what has previously been anticipated.

As mentioned above, *Mycobacterium tuberculosis* encodes a large number of TA systems, especially *vapBC* systems. The Beijing-lineage of *M. tuberculosis* is known to display high virulence and exhibit higher levels of drug resistance (Van Soolingen et al. 1995; Glynn et al. 2002). Here, sequence analysis has revealed three single nucleotide polymorphisms (SNPs) in *vapC* toxins of the sublineage Beijing modern (*vapC37*, *vapC38*) and one in *vapC12* of Beijing-B0/W148 (Zaychikova et al. 2018). Interestingly, for VapC12 and VapC38, these mutations were shown to display decreased ribonuclease activity in vitro. How a less active toxin can benefit a bacterial cell is currently unknown, but like for *ccdAB*, it could involve a mechanism of transcriptional regulation. For VapC37, it seems that the toxin activity increases slightly in vitro. In any case, it must be borne in mind that any physiological effect is the result of the total set of VapBC in the organism and therefore also that potential cross-inhibition might occur and any hypothesis of the functions of these VapC variations is difficult to make.

Mutations and Deletions in Type II Antitoxins

Variations in the *vapB* Antitoxin Might Confer Antibiotic Tolerance The F plasmid of *Escherichia coli* consists of four regions, a leading region, stability region, a region with cargo genes, and a DNA transfer region encoding a *vapBC* locus (Koraimann 2018). The importance of the *vapBC* TA systems for maintenance of the F plasmid is not well understood; however, the locus is encoded on the opposite strand of the transfer genes and overlaps with a potential *trbH* gene, which is only conserved in 10% of F-like plasmids (Cox and Schildbach 2017; Fridman et al. 2014). This unusual genetic organisation with *vapBC* on the antisense and *trbH* on

the sense strand is conserved among the virulence plasmids of the pathogens *Shigella flexneri* and *Shigella sonnei*, with a few point mutations separating the three sets of VapBC proteins. It is known that *vapBC* is important for maintenance of the virulence plasmid of *S. flexneri* (Sayeed et al. 2000; McVicker et al. 2019), less in *S. sonnei* (Martyn et al. 2022), while it does not seem to be important for plasmid maintenance of the F plasmid in *E. coli* (Radnedge et al. 1997). It is not currently known why this is the case and why only a few point mutations can cause this difference. In 2017, a study using artificial evolution based on repeated antibiotic exposure resulted in an *E. coli* strain (*E. coli* KLY) with several independent point mutations in the *vapB* gene found on the F plasmid (Levin-Reisman et al. 2017). Importantly, the authors could show that the mutations caused an antibiotic-tolerant phenotype that correlated with mutations in the VapB protein (T3N, V5E, L7P, A13P, L16R, V20G, and an N-terminal inversion) (Levin-Reisman et al. 2017; Fridman et al. 2014). Whether these mutations are also a causal factor of the observed tolerant phenotype is not known, and in some instances, there were also mutations observed elsewhere; however, mutations in VapB seemed to consistently show up. Experimental crystal structures of the VapBC complexes encoded on the virulence plasmids of *S. flexneri* and *S. sonnei* have been determined, which allows us to make solid predictions about potential functional importance of the mutations on the F plasmid (Dienemann et al. 2011; Martyn et al. 2022). All of the observed tolerance mutations of VapB map to the DNA-binding domain in the N-terminus of the antitoxin suggesting initially that transcriptional regulation might be affected (Fig. 11.6). It is also possible that the mutations prevent VapB dimerisation and therefore disrupt the entire VapBC complex, potentially leading to increased activation of VapC. Activation of VapC is normally believed to occur by degradation of VapB by the Lon protease, which is known to be important for the role of VapBC in plasmid maintenance in *S. flexneri* (Winther and Gerdes 2012; McVicker et al. 2019). In the experimental structures of VapBC, the VapB N-terminal tails extend away from the complex and are stabilised by crystal contacts (Fig. 11.6a). Given that these extensions overlap with the DNA-binding region of the VapB dimers, it can be speculated that they need to fold up upon or before interaction with DNA and that this ability might be affected by mutations (Levin-Reisman et al. 2017). It is important to keep in mind that these mutations were observed in an artificial evolution experiment and therefore might not represent mutations that would arise naturally or in a clinical setting. Therefore, any biological relevance of these potential tolerance mutations would need to be proven experimentally or similar mutations would need to be observed in clinically isolated tolerant cultures.

***Salmonella* Typhimurium Encodes a Non-neutralising Antitoxin RelB_{ST}** The *relBE* loci were among the first characterised type II TA systems in 1998 (Gottfredsen and Gerdes 1998). The name derives from the fact that mutations in the *relB* gene show a “delayed relaxed” phenotype, characterised by a delay in accumulation of tRNA and rRNA during amino acid starvation compared to wild-type strains (Mosteller 1978; Diderichsen et al. 1977; Diderichsen and Desmarez 1980). Genetic analysis revealed two nearby genes, *relE* and *relF*, located in the same operon, for

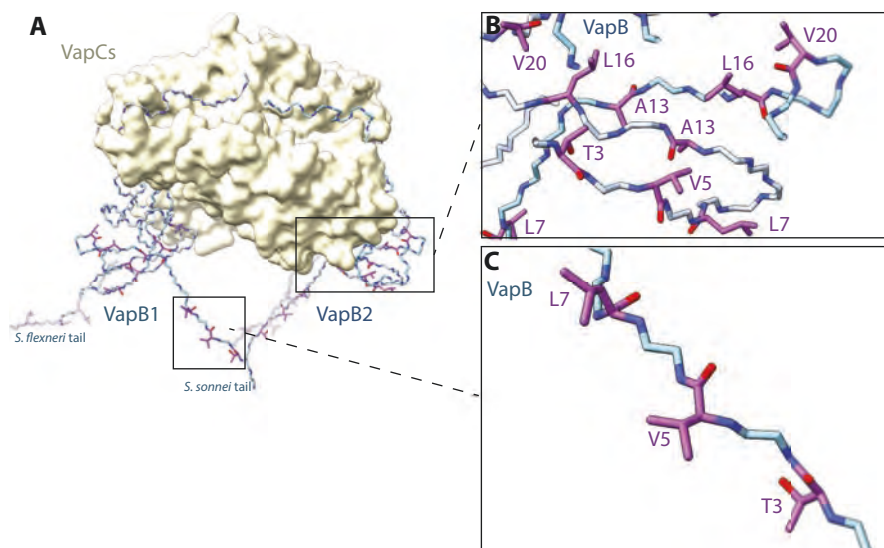


Fig. 11.6 Structure of the *Shigella sonnei* VapBC complex. (a) The heterooctameric VapBC complex from *S. sonnei* (PDB: 6SD6) (Martyn et al. 2022) shown in combined surface/stick representation with VapB (blue) and VapC (white). The potential tolerance mutations (purple) are all located in the N-terminal extension and DNA-binding domain of VapB, of which the extension is stabilised by lattice contacts in the crystal. The corresponding extension from the *S. flexneri* structure (PDB: 3TND) (Dienemann et al. 2011) is shown as a ghost chain. (b) Closeup of the residues affected by tolerance mutations in the VapB DNA-binding domain. (c) Closeup of the residues affected by tolerance mutations in the VapB extended region stabilised by crystal contacts

which *relE* and its associated protein product, RelE, was shown to be toxic upon expression (Bech et al. 1985; Gotfredsen and Gerdes 1998). RelE forms a complex with RelB, and together they defined *relBE* as a type II TA system due to activation of RelE, induced by Lon protease-specific RelB degradation (Pedersen et al. 2003; Christensen et al. 2001). The *Escherichia coli* RelBE TA higher-order complex has an overall 2:2 stoichiometry consisting of a homodimer of RelB, each with a single distally bound RelE monomer (Bøggild et al. 2012). Like other type II antitoxins, RelB contains an N-terminal DNA-binding domain formed by tightly interacting β -sheets, while the C-terminus is extended and interacts with and inhibits RelE by sterically preventing it from entering the ribosomal A site where it otherwise cleaves mRNA in a semi-sequence dependent fashion, causing translational arrest (Cherny et al. 2007; Bøggild et al. 2012; Neubauer et al. 2009; Pedersen et al. 2003). *Salmonella enterica* serovar Typhimurium encodes 27 TA systems (type I and II) including 4 *relBE* systems (Lobato-Márquez et al. 2015). Interestingly, one of these, *relBE*_{3ST}, encodes a functional toxin; however, the associated RelB_{3ST} antitoxin is unable to inhibit it (Lobato-Márquez et al. 2015). The reason for encoding an inactive antitoxin and why that does not immediately cause translational arrest is not currently understood. A single residue (proline 14) is hypothesised to be responsible for the inactive antitoxin, and mutation of this residue to alanine allows RelB_{3ST} to

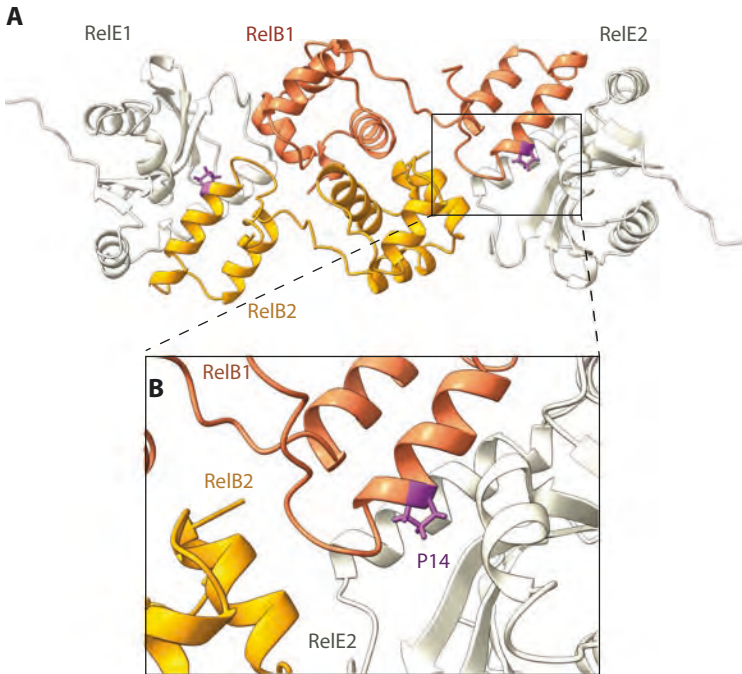


Fig. 11.7 Predicted structure of the RelBE_{3ST} complex. (a) Structure of the predicted *S. enterica* RelBE_{3ST} complex (iptm+ptm = 0.497) shown as cartoon with RelE1 and RelE2 in white and the antitoxins RelB1 and RelB2 in yellow and orange. (b) Closeup of proline P14 located at the interface with RelE2

bind and inhibit RelE_{3ST} (Yusof et al. 2022). Currently, all experimentally determined structures of RelBE adopt a similar higher-order structure. Surprisingly, however, in silico prediction of the structure of the RelBE_{3ST} complex results in a different organisation than the experimental crystal structure of *E. coli* RelBE (Bøggild et al. 2012). While the predicted structure of the RelE_{3ST} toxin is compact and similar to other known RelE structures, it is the C terminus rather than the N terminus of RelB_{3ST}, which is predicted to form a DNA-binding domain and organise the higher-order 2:2 complex (Fig. 11.7). Consequently, the N terminus of RelB_{3ST} is predicted to interact with the toxin. RelBE_{3ST} is more closely related to HigBA than *E. coli* RelBE, and in the HigBA2 complex from *Vibrio cholerae*, the HigA2 antitoxin also displays a C-terminal DNA binding domain (Hadži et al. 2017). The RelBE_{3ST} prediction is further supported by the observation that isolated RelB_{3ST} can still bind DNA, suggesting that only the interaction with the toxin is affected by the presence of P14, which is close to the N-terminus of the antitoxin (Yusof et al. 2022). The toxin-antitoxin interface contains an α -helix, which has been predicted to be shortened by the presence of P14 (Yusof et al. 2022); however, this is not borne out in the AlphaFold2 prediction. Regardless of structure, it is known that RelB_{3ST} cannot inhibit the toxicity of RelE_{3ST} in vivo (Yusof et al.

2022). Given that *S. Typhimurium* encodes four different *relBE* systems, an attractive hypothesis would be that RelE3_{ST} could be inhibited by a RelB antitoxin from another system, thus making its cognate antitoxin redundant. There is currently no evidence for these speculations, and the phenotype caused by the inactive antitoxin, if any, is unknown.

Summary

Toxin-antitoxin systems are widespread in bacterial genomes and on plasmids, and especially type II TA systems are functionally and structurally quite well understood. Here we have presented some of the spontaneous variations observed in Type II TA systems that can affect either the function of the toxin or antitoxin or the phenotype they confer. Some of these variations occur naturally, and others have been artificially evolved in the lab. *hipA* was discovered by analysis of an increased persistence phenotype in *E. coli*, and it was only later discovered that the gene belongs to a type II TA system. The mutated allele *hipA7* was found to contain two mutations causing the toxin kinase HipA7 to be less active and preventing inhibition by the antitoxin HipB. How this causes persistence is unknown and also debated. In any case, it seems that the wt *hipA* gene is not important for persister formation, and its natural function therefore remains elusive. The *hipA7* variant, on the other hand, does infer a persistence phenotype, but it might not be a result of exposure to antibiotics but rather allowing for heterogeneous growth rates within a bacterial population. The virulence plasmid of *Salmonella Typhimurium* contains a *ccdAB* TA system, which expresses an inactive toxin. Despite this, the TA system appears to be important for plasmid stability and is potentially also important for regulation of downstream genes encoded within the same operon. For the endoribonuclease toxin, VapD, from *Haemophilus influenzae*, some clinical isolates express a toxin containing a large, internal deletion that removes roughly half of the protein. This was shown to result in a shorter VapD protein that is unable to dimerise and therefore most likely will be an inactive toxin on its own. However, when co-expressed with full-length VapD, this causes a dominant negative phenotype, likely through inactivation of the wild-type toxin. Whether this variant arose to counteract the effect of other VapD toxins or simply as a way to inactivate VapD is currently unknown. The largest family of type II TA systems is VapBC, which is also present in large numbers in the *M. tuberculosis* genome. The highly virulent Beijing modern strain of *M. tuberculosis* encodes two unique variant *vapBC* loci, *vapC37* and *vapC38*, while the strain Beijing-B0/W148 contains one unique *vapC12* variant. The products of these genes show altered RNase activity rates compared to the wild-type genes of the same toxins. Whether these changes have an effect on *M. tuberculosis* phenotype or virulence is unknown. The F plasmid of *E. coli* contains a *vapBC* locus, which might cause antibiotic tolerance if mutations occur in the DNA binding domain of VapB. These mutations could prevent both DNA binding by the VapBC complex but could also cause instability of the complex. How this might

confer tolerance is unknown. Finally, the RelBE_{3ST} of *Salmonella enterica* Typhimurium encodes an antitoxin, which cannot inhibit its cognate toxin, RelE_{3ST}. However, *Salmonella* encodes three other RelBE TA systems, which could encode antitoxins that could inhibit RelE_{3ST}, making the RelB_{3ST} protein redundant. Of the TA variations described here, some have only been found within a lab setting, such as the persister mutations of *hipA* and the tolerance mutations of *vapB*. The phenotypes described for these mutations might be true, such as the increased persistence observed for *hipA7*, but we have no idea if these variations would occur in nature. Other gene variations such as the *ccdB* and *relB3_{ST}* variations of *Salmonella* are known to be conserved within this serovar (Typhimurium) and closely related strains, which indicates that evolution has selected these variants because they are somehow beneficial. One might speculate, for example, that since *ccdB* is important for autoregulation of the *ccdAB* operon, it might also regulate other downstream genes. For *relB3_{ST}*, it can be speculated that the inability of the toxin to inhibit RelB3 results in more active toxin and thus a quicker dormancy response when necessary. Likewise, the fact that the same large deletion of the VapD toxin from *Haemophilus influenzae* was observed in isolates from several patients suggests that this toxin has a unique function. We speculate that the dominant negative phenotype that arises when truncated VapD dimerises with full-length VapD results in an anti-addiction effect, whereby the truncated VapD can function as an “antitoxin” that inhibits dimerisation of full-length VapD. Finally, for the variations found in VapCs from the Beijing sublineage of *M. tuberculosis*, it can be speculated whether these changes in the toxicity can have an effect on persistence of the pathogen. The Beijing sublineage is known to be more virulent and more persistent than other sublineages of *M. tuberculosis*, but there could potentially be other variations in the bacterial genome that could explain this difference in pathogenicity, and it remains possible that the toxicity of VapC could play a role in this.

References

- Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S (2004) Bacterial persistence as a phenotypic switch. *Science* 305(5690):1622–1625. <https://doi.org/10.1126/science.1099390>
- Bech FW, Jørgensen ST, Diderichsen B, Karlström OH (1985) Sequence of the *relB* transcription unit from *Escherichia coli* and identification of the *relB* gene. *EMBO J* 4(4):1059–1066. <https://doi.org/10.1002/j.1460-2075.1985.tb03739.x>
- Bendtsen KL, Brodersen DE (2017) Higher-order structure in bacterial VapBC toxin-antitoxin complexes. In: Harris JR, Marles-Wright J (eds) *Macromolecular protein complexes: structure and function*. Springer International Publishing, Cham, pp 381–412. https://doi.org/10.1007/978-3-319-46503-6_14
- Bendtsen KL, Xu K, Luckmann M, Winther KS, Shah SA, Pedersen CNS, Brodersen DE (2017) Toxin inhibition in *C. crescentus* VapBC1 is mediated by a flexible pseudo-palindromic protein motif and modulated by DNA binding. *Nucleic Acids Res* 45(5):2875–2886. <https://doi.org/10.1093/nar/gkw1266>

- Bernard P, Couturier M (1992) Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol* 226(3):735–745. [https://doi.org/10.1016/0022-2836\(92\)90629-x](https://doi.org/10.1016/0022-2836(92)90629-x)
- Bertelsen MB, Senissar M, Nielsen MH, Bisiak F, Cunha MV, Molinaro AL, Daines DA, Brodersen DE (2021) Structural basis for toxin inhibition in the VapXD toxin-antitoxin system. *Structure* 29(2):139–150.e133. <https://doi.org/10.1016/j.str.2020.10.002>
- Bobonis J, Mitosch K, Mateus A, Karcher N, Kritikos G, Selkrig J, Zietek M, Monzon V, Pfalz B, Garcia-Santamarina S, Galardini M, Sueki A, Kobayashi C, Stein F, Bateman A, Zeller G, Savitski MM, Elfenbein JR, Andrews-Polymenis HL, Typas A (2022) Bacterial retrons encode phage-defending tripartite toxin-antitoxin systems. *Nature* 609(7925):144–150. <https://doi.org/10.1038/s41586-022-05091-4>
- Bøggild A, Sofos N, Andersen KR, Feddersen A, Easter AD, Passmore LA, Brodersen DE (2012) The crystal structure of the intact *E. coli* RelBE toxin-antitoxin complex provides the structural basis for conditional cooperativity. *Structure* 20(10):1641–1648. <https://doi.org/10.1016/j.str.2012.08.017>
- Boyd EF, Hartl DL (1997) Recent horizontal transmission of plasmids between natural populations of *Escherichia coli* and *Salmonella enterica*. *J Bacteriol* 179(5):1622–1627. <https://doi.org/10.1128/jb.179.5.1622-1627.1997>
- Bradford PA (2001) Extended-spectrum β -lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. *Clin Microbiol Rev* 14(4):933–951. <https://doi.org/10.1128/cmr.14.4.933-951.2001>
- Brauner A, Fridman O, Gefen O, Balaban NQ (2016) Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat Rev Microbiol* 14(5):320–330. <https://doi.org/10.1038/nrmicro.2016.34>
- Brives C, Pourraz J (2020) Phage therapy as a potential solution in the fight against AMR: obstacles and possible futures. *Palgrave Communications* 6(1):100. <https://doi.org/10.1057/s41599-020-0478-4>
- Brüssow H, Hendrix RW (2002) Phage genomics: small is beautiful. *Cell* 108(1):13–16. [https://doi.org/10.1016/s0092-8674\(01\)00637-7](https://doi.org/10.1016/s0092-8674(01)00637-7)
- Cataudella I, Trusina A, Snuppen K, Gerdes K, Mitarai N (2012) Conditional cooperativity in toxin-antitoxin regulation prevents random toxin activation and promotes fast translational recovery. *Nucleic Acids Res* 40(14):6424–6434. <https://doi.org/10.1093/nar/gks297>
- Cherny I, Overgaard M, Borch J, Bram Y, Gerdes K, Gazit E (2007) Structural and thermodynamic characterization of the *Escherichia coli* RelBE toxin-antitoxin system: indication for a functional role of differential stability. *Biochemistry* 46(43):12152–12163. <https://doi.org/10.1021/bi701037e>
- Christensen SK, Mikkelsen M, Pedersen K, Gerdes K (2001) RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci* 98(25):14328–14333. <https://doi.org/10.1073/pnas.251327898>
- Cox KEL, Schilbach JF (2017) Sequence of the R1 plasmid and comparison to F and R100. *Plasmid* 91:53–60. <https://doi.org/10.1016/j.plasmid.2017.03.007>
- Daines DA, Jarisch J, Smith AL (2004) Identification and characterization of a nontypeable *Haemophilus influenzae* putative toxin-antitoxin locus. *BMC Microbiol* 4:30. <https://doi.org/10.1186/1471-2180-4-30>
- Diderichsen B, Desmarez L (1980) Variations in phenotype of relB mutants of *Escherichia coli* and the effect of *pus* and *sup* mutations. *Mol Gen Genet* 180(2):429–437. <https://doi.org/10.1007/bf00425859>
- Diderichsen B, Fiil NP, Lavallé R (1977) Genetics of the relB locus in *Escherichia coli*. *J Bacteriol* 131(1):30–33. <https://doi.org/10.1128/jb.131.1.30-33.1977>
- Dienemann C, Bøggild A, Winther KS, Gerdes K, Brodersen DE (2011) Crystal structure of the VapBC toxin-antitoxin complex from *Shigella flexneri* reveals a hetero-octameric DNA-binding assembly. *J Mol Biol* 414(5):713–722. <https://doi.org/10.1016/j.jmb.2011.10.024>

- Fineran PC, Blower TR, Foulds IJ, Humphreys DP, Lilley KS, Salmond GPC (2009) The phage abortive infection system, ToxIN, functions as a protein–RNA toxin–antitoxin pair. *Proc Natl Acad Sci* 106(3):894–899. <https://doi.org/10.1073/pnas.0808832106>
- Fridman O, Goldberg A, Ronin I, Shores N, Balaban NQ (2014) Optimization of lag time underlies antibiotic tolerance in evolved bacterial populations. *Nature* 513(7518):418–421. <https://doi.org/10.1038/nature13469>
- Gerdes K, Rasmussen PB, Molin S (1986) Unique type of plasmid maintenance function: post-segregational killing of plasmid-free cells. *Proc Natl Acad Sci* 83(10):3116–3120. <https://doi.org/10.1073/pnas.83.10.3116>
- Gerdes K, Baerentsen R, Brodersen DE (2021) Phylogeny reveals novel HipA-homologous kinase families and toxin-antitoxin gene organizations. *mBio* 12(3):e0105821. <https://doi.org/10.1128/mBio.01058-21>
- Germain E, Castro-Roa D, Zenkin N, Gerdes K (2013) Molecular mechanism of bacterial persistence by HipA. *Mol Cell* 52(2):248–254. <https://doi.org/10.1016/j.molcel.2013.08.045>
- Glynn JR, Whiteley J, Bifani PJ, Kremer K, Van Soolingen D (2002) Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg Infect Dis* 8(8):843–849. <https://doi.org/10.3201/eid0805.020002>
- Gotfredsen M, Gerdes K (1998) The *Escherichia coli* relBE genes belong to a new toxin-antitoxin gene family. *Mol Microbiol* 29(4):1065–1076. <https://doi.org/10.1046/j.1365-2958.1998.00993.x>
- Guglielmini J, Van Melder L (2011) Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. *Mob Genet Elements* 1(4):283–290. <https://doi.org/10.4161/mge.18477>
- Hadži S, Garcia-Pino A, Haesaerts S, Jurėnas D, Gerdes K, Lah J, Loris R (2017) Ribosome-dependent *Vibrio cholerae* mRNAse HigB2 is regulated by a β -strand sliding mechanism. *Nucleic Acids Res* 45(8):4972–4983. <https://doi.org/10.1093/nar/gkx138>
- Hansen S, Lewis K, Vulić M (2008) Role of global regulators and nucleotide metabolism in antibiotic tolerance in *Escherichia coli*. *Antimicrob Agents Chemother* 52(8):2718–2726. <https://doi.org/10.1128/aac.00144-08>
- Harms A, Brodersen DE, Mitarai N, Gerdes K (2018) Toxins, targets, and triggers: an overview of toxin-antitoxin biology. *Mol Cell* 70(5):768–784. <https://doi.org/10.1016/j.molcel.2018.01.003>
- Jennifer R, Joanna LM, Ray C, Gregory MC, Vickery LA (2009) The vapBC operon from *Mycobacterium smegmatis* is an autoregulated toxin–antitoxin module that controls growth via inhibition of translation. *J Mol Biol* 390(3):353–367. <https://doi.org/10.1016/j.jmb.2009.05.006>
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873):583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kelly A, Arrowsmith TJ, Went SC, Blower TR (2023) Toxin-antitoxin systems as mediators of phage defence and the implications for abortive infection. *Curr Opin Microbiol* 73:102293. <https://doi.org/10.1016/j.mib.2023.102293>
- Koraimann G (2018) Spread and persistence of virulence and antibiotic resistance genes: a ride on the F plasmid conjugation module. *EcoSal Plus* 8(1). <https://doi.org/10.1128/ecosalplus.ESP-0003-2018>
- Korch SB, Henderson TA, Hill TM (2003) Characterization of the hipA7 allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Mol Microbiol* 50(4):1199–1213. <https://doi.org/10.1046/j.1365-2958.2003.03779.x>
- Kwon A-R, Kim J-H, Park SJ, Lee K-Y, Min Y-H, Im H, Lee I, Lee K-Y, Lee B-J (2012) Structural and biochemical characterization of HP0315 from *Helicobacter pylori* as a VapD protein with an endoribonuclease activity. *Nucleic Acids Res* 40(9):4216–4228. <https://doi.org/10.1093/nar/gkr1305>

- LeRoux M, Laub MT (2022) Toxin-antitoxin systems as phage defense elements. *Ann Rev Microbiol* 76:21–43. <https://doi.org/10.1146/annurev-micro-020722-013730>
- LeRoux M, Culviner PH, Liu YJ, Littlehale ML, Laub MT (2020) Stress can induce transcription of toxin-antitoxin systems without activating toxin. *Mol Cell* 79(2):280–292.e288. <https://doi.org/10.1016/j.molcel.2020.05.028>
- LeRoux M, Srikant S, Teodoro GIC, Zhang T, Littlehale ML, Doron S, Badiie M, Leung AKL, Sorek R, Laub MT (2022) The DarTG toxin-antitoxin system provides phage defence by ADP-ribosylating viral DNA. *Nat Microbiol* 7(7):1028–1040. <https://doi.org/10.1038/s41564-022-01153-5>
- Levin I, Schwarzenbacher R, Page R, Abdubek P, Ambing E, Biorac T, Brinen LS, Campbell J, Canaves JM, Chiu HJ, Dai X, Deacon AM, DiDonato M, Elsliger MA, Floyd R, Godzik A, Grittini C, Grzechnik SK, Hampton E, Jaroszewski L, Karlak C, Klock HE, Koesema E, Kovarik JS, Kreusch A, Kuhn P, Lesley SA, McMullan D, McPhillips TM, Miller MD, Morse A, Moy K, Ouyang J, Quijano K, Reyes R, Rezezadeh F, Robb A, Sims E, Spraggon G, Stevens RC, van den Bedem H, Velasquez J, Vincent J, von Delft F, Wang X, West B, Wolf G, Xu Q, Hodgson KO, Wooley J, Wilson IA (2004) Crystal structure of a PIN (PilT N-terminus) domain (AF0591) from *Archaeoglobus fulgidus* at 1.90 Å resolution. *Proteins* 56(2):404–408. <https://doi.org/10.1002/prot.20090>
- Levin-Reisman I, Ronin I, Gefen O, Braniss I, Shores N, Balaban NQ (2017) Antibiotic tolerance facilitates the evolution of resistance. *Science* 355(6327):826–830. <https://doi.org/10.1126/science.aaj2191>
- Lobato-Márquez D, Moreno-Córdoba I, Figueroa V, Díaz-Orejas R, García-Del Portillo F (2015) Distinct type I and type II toxin-antitoxin modules control *Salmonella* lifestyle inside eukaryotic cells. *Sci Rep* 5. <https://doi.org/10.1038/srep09374>
- Lobato-Marquez D, Molina-Garcia L, Moreno-Cordoba I, Garcia-Del Portillo F, Diaz-Orejas R (2016) Stabilization of the virulence plasmid pSLT of *salmonella typhimurium* by three maintenance systems and its evaluation by using a new stability test. *Front Mol Biosci* 3:66. <https://doi.org/10.3389/fmolb.2016.00066>
- Loris R, Garcia-Pino A (2014) Disorder- and dynamics-based regulatory mechanisms in toxin-antitoxin modules. *Chem Rev* 114(13):6933–6947. <https://doi.org/10.1021/cr400656f>
- Loris R, Dao-Thi M-H, Bahassi EM, Van Melderden L, Poortmans F, Liddington R, Couturier M, Wyls L (1999) Crystal structure of CcdB, a topoisomerase poison from *E. coli*. *J Mol Biol* 285(4):1667–1677. <https://doi.org/10.1006/jmbi.1998.2395>
- Makarova KS, Anantharaman V, Aravind L, Koonin EV (2012) Live virus-free or die: coupling of antiviral immunity and programmed suicide or dormancy in prokaryotes. *Biol Direct* 7:40. <https://doi.org/10.1186/1745-6150-7-40>
- Martyn JE, Pilla G, Hollingshead S, Winther KS, Lea S, McVicker G, Tang CM (2022) Maintenance of the *Shigella sonnei* virulence plasmid is dependent on its repertoire and amino acid sequence of toxin-antitoxin systems. *J Bacteriol* 204(3):e0051921. <https://doi.org/10.1128/jb.00519-21>
- Maté MJ, Vincentelli R, Foos N, Raoult D, Cambillau C, Ortiz-Lombardía M (2012) Crystal structure of the DNA-bound VapBC2 antitoxin/toxin pair from *Rickettsia felis*. *Nucleic Acids Res* 40(7):3245–3258. <https://doi.org/10.1093/nar/gkr1167>
- McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413(6858):852–856. <https://doi.org/10.1038/35101614>
- McVicker G, Hollingshead S, Pilla G, Tang CM (2019) Maintenance of the virulence plasmid in *Shigella flexneri* is influenced by Lon and two functional partitioning systems. *Mol Microbiol* 111(5):1355–1366. <https://doi.org/10.1111/mmi.14225>
- Mosteller RD (1978) Evidence that glucose starvation-sensitive mutants are altered in the *relB* locus. *J Bacteriol* 133(2):1034–1037. <https://doi.org/10.1128/jb.133.2.1034-1037.1978>

- Moyed HS, Bertrand KP (1983) *hipA*, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *J Bacteriol* 155(2):768–775. <https://doi.org/10.1128/jb.155.2.768-775.1983>
- Murphy TF (2003) Respiratory infections caused by non-typeable *Haemophilus influenzae*. *Curr Opin Infect Dis* 16(2):129–134. <https://doi.org/10.1097/00001432-200304000-00009>
- Neubauer C, Gao Y-G, Andersen KR, Dunham CM, Kelley AC, Hentschel J, Gerdes K, Ramakrishnan V, Brodersen DE (2009) The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell* 139(6):1084–1095. <https://doi.org/10.1016/j.cell.2009.11.015>
- Núñez JK, Kranzusch PJ, Noeske J, Wright AV, Davies CW, Doudna JA (2014) Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity. *Nat Struct Mol Biol* 21(6):528–534. <https://doi.org/10.1038/nsmb.2820>
- O'Neill J (2016) Tackling drug-resistant infections globally: final report and recommendations. Review on antimicrobial resistance. Wellcome Trust, London. https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf
- Ochman H, Wilson AC (1987) Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J Mol Evol* 26(1–2):74–86. <https://doi.org/10.1007/bf02111283>
- Ogura T, Hiraga S (1983) Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc Natl Acad Sci USA* 80(15):4784–4788. <https://doi.org/10.1073/pnas.80.15.4784>
- Overgaard M, Borch J, Jørgensen MG, Gerdes K (2008) Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Mol Microbiol* 69(4):841–857. <https://doi.org/10.1111/j.1365-2958.2008.06313.x>
- Page R, Peti W (2016) Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat Chem Biol* 12(4):208–214. <https://doi.org/10.1038/nchembio.2044>
- Pandey DP, Gerdes K (2005) Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* 33(3):966–976. <https://doi.org/10.1093/nar/gki201>
- Park D, Yoon HJ, Lee KY, Park SJ, Cheon SH, Lee HH, Lee SJ, Lee BJ (2020) Crystal structure of proteolyzed VapBC and DNA-bound VapBC from *Salmonella enterica* Typhimurium LT2 and VapC as a putative Ca(2+)-dependent ribonuclease. *FASEB J* 34(2):3051–3068. <https://doi.org/10.1096/fj.201901989R>
- Pecota DC, Wood TK (1996) Exclusion of T4 phage by the *hok/sok* killer locus from plasmid R1. *J Bacteriol* 178(7):2044–2050. <https://doi.org/10.1128/jb.178.7.2044-2050.1996>
- Pedersen K, Zavialov AV, Pavlov MY, Elf J, Gerdes K, Ehrenberg M (2003) The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell* 112(1):131–140. [https://doi.org/10.1016/s0092-8674\(02\)01248-5](https://doi.org/10.1016/s0092-8674(02)01248-5)
- Radnedge L, Davis MA, Youngren B, Austin SJ (1997) Plasmid maintenance functions of the large virulence plasmid of *Shigella flexneri*. *J Bacteriol* 179(11):3670–3675. <https://doi.org/10.1128/jb.179.11.3670-3675.1997>
- Ramage HR, Connolly LE, Cox JS (2009) Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genet* 5(12):e1000767. <https://doi.org/10.1371/journal.pgen.1000767>
- Ren D, Walker AN, Daines DA (2012) Toxin-antitoxin loci *vapBC-1* and *vapXD* contribute to survival and virulence in nontypeable *Haemophilus influenzae*. *BMC Microbiol* 12:263. <https://doi.org/10.1186/1471-2180-12-263>
- Saavedra De Bast M, Mine N, Van Melderen L (2008) Chromosomal toxin-antitoxin systems may act as antiaddiction modules. *J Bacteriol* 190(13):4603–4609. <https://doi.org/10.1128/jb.00357-08>
- Sayeed S, Reeves L, Radnedge L, Austin S (2000) The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. *J Bacteriol* 182(9):2416–2421. <https://doi.org/10.1128/jb.182.9.2416-2421.2000>

- Scherrer R, Moyed HS (1988) Conditional impairment of cell division and altered lethality in hipA mutants of *Escherichia coli* K-12. *J Bacteriol* 170(8):3321–3326. <https://doi.org/10.1128/jb.170.8.3321-3326.1988>
- Schumacher MA, Piro KM, Xu W, Hansen S, Lewis K, Brennan RG (2009) Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. *Science* 323(5912):396–401. <https://doi.org/10.1126/science.1163806>
- Schumacher MA, Min J, Link TM, Guan Z, Xu W, Ahn YH, Soderblom EJ, Kurie JM, Evdokimov A, Moseley MA, Lewis K, Brennan RG (2012) Role of unusual P loop ejection and autophosphorylation in HipA-mediated persistence and multidrug tolerance. *Cell Rep* 2(3):518–525. <https://doi.org/10.1016/j.celrep.2012.08.013>
- Schumacher MA, Balani P, Min J, Chinnam NB, Hansen S, Vulic M, Lewis K, Brennan RG (2015) HipBA-promoter structures reveal the basis of heritable multidrug tolerance. *Nature* 524(7563):59–64. <https://doi.org/10.1038/nature14662>
- Semanjski M, Germain E, Bratl K, Kiessling A, Gerdes K, Macek B (2018) The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance. *Sci Signal* 11(547). <https://doi.org/10.1126/scisignal.aat5750>
- Song S, Wood TK (2020) Toxin/Antitoxin system paradigms: toxins bound to antitoxins are not likely activated by preferential antitoxin degradation. *Adv Biosyst* 4(3):1900290. <https://doi.org/10.1002/adbi.201900290>
- Van Melderden L, Saavedra De Bast M (2009) Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genet* 5(3):e1000437. <https://doi.org/10.1371/journal.pgen.1000437>
- Van Melderden L, Bernard P, Couturier M (1994) Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Mol Microbiol* 11(6):1151–1157. <https://doi.org/10.1111/j.1365-2958.1994.tb00391.x>
- Van Soolingen D, Qian L, De Haas PEW, Douglas JT, Traore H, Portaels F, Enkhsaikan D, Nymadawa P, Van Embden JDA (1995) Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol* 33(12):3234–3238. <https://doi.org/10.1128/jcm.33.12.3234-3238.1995>
- Winther KS, Gerdes K (2012) Regulation of enteric vapBC transcription: induction by VapC toxin dimer-breaking. *Nucleic Acids Res* 40(10):4347–4357. <https://doi.org/10.1093/nar/gks029>
- Winther K, Tree JJ, Tollervey D, Gerdes K (2016) VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation. *Nucleic Acids Res* 44(20):9860–9871. <https://doi.org/10.1093/nar/gkw781>
- Yee WX, Yasir M, Turner AK, Baker DJ, Cehovin A, Tang CM (2023) Evolution, persistence, and host adaption of a gonococcal AMR plasmid that emerged in the pre-antibiotic era. *PLoS Genet* 19(5):e1010743. <https://doi.org/10.1371/journal.pgen.1010743>
- Yusuf TY, Ong EBB, Teh AH (2022) RelEB3 toxin-antitoxin system of *Salmonella Typhimurium* with a ribosome-independent toxin and a mutated non-neutralising antitoxin. *Int J Biol Macromol* 219:1080–1086. <https://doi.org/10.1016/j.ijbiomac.2022.08.144>
- Zaychikova MV, Mikhecheva NE, Belay YO, Alekseeva MG, Melerzanov AV, Danilenko VN (2018) Single nucleotide polymorphisms of Beijing lineage *Mycobacterium tuberculosis* toxin-antitoxin system genes: Their role in the changes of protein activity and evolution. *Tuberculosis (Edinb)* 112:11–19. <https://doi.org/10.1016/j.tube.2018.06.011>

Chapter 12

Structures and Functions of the Human GATOR1 Complex



Ilina Ivanova and Kuang Shen

Abstract Eukaryotic cells coordinate available nutrients with their growth through the mechanistic target of rapamycin complex 1 (mTORC1) pathway, in which numerous evolutionarily conserved protein complexes survey and transmit nutrient inputs toward mTORC1. mTORC1 integrates these inputs and activates downstream anabolic or catabolic programs that are in tune with cellular needs, effectively maintaining metabolic homeostasis. The *GAP* activity toward *Rags-1* (GATOR1) protein complex is a critical negative regulator of the mTORC1 pathway and, in the absence of amino acid inputs, is activated to turn off mTORC1 signaling. GATOR1-mediated inhibition of mTORC1 signaling is tightly regulated by an ensemble of protein complexes that antagonize or promote its activity in response to the cellular nutrient environment. Structural, biochemical, and biophysical studies of the GATOR1 complex and its interactors have advanced our understanding of how it regulates cellular metabolism when amino acids are limited. Here, we review the current research with a focus on GATOR1 structure, its enzymatic mechanism, and the growing group of proteins that regulate its activity. Finally, we discuss the implication of GATOR1 dysregulation in physiology and human diseases.

Keywords GATOR1 · mTORC1 · GTPase activating protein (GAP) · Rag GTPases · Enzyme mechanism · Protein structure

I. Ivanova

Program in Molecular Medicine, University of Massachusetts Chan Medical School,
Worcester, MA, USA

e-mail: ilina.ivanova@umassmed.edu

K. Shen (✉)

Program in Molecular Medicine, University of Massachusetts Chan Medical School,
Worcester, MA, USA

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Chan
Medical School, Worcester, MA, USA

e-mail: Kuang.Shen@umassmed.edu

Introduction

The mechanistic target of rapamycin complex 1 (mTORC1) pathway balances cell growth with resource availability by surveying the environment for nutrients, energy, and growth factors. At the core of mTORC1 is a highly conserved serine/threonine protein kinase, mTOR, which acts as a central switch between anabolic and catabolic pathways (Brown et al. 1994; Sabatini et al. 1994; Sabers et al. 1995). When environmental resources are abundant, mTORC1 phosphorylates key effectors that drive downstream anabolic processes essential for growth, development, and proliferation (Ma and Blenis 2009). Conversely, when resources are limited, mTORC1 inhibition shifts the cell toward a catabolic state, promoting the autophagic breakdown and recycling of cellular components (Rabinowitz and White 2010). Cells must tightly regulate mTORC1 activity in response to resource fluctuations stimulated by feeding or fasting to prevent futile growth under malnutrition. To this end, mTORC1 integrates upstream inputs from an ensemble of dedicated nutrient-sensing proteins that respond to cellular levels of amino acids, metabolites, growth factors, ATP, and glucose, and activate or inhibit mTORC1 signaling depending on the nutrient state of the cell (Condon and Sabatini 2019; Dibble and Manning 2013) (Fig. 12.1).

In mammalian cells, the mTORC1 pathway signaling hub is situated at the lysosomal surface where the switch to mTORC1 activation is coordinated by two small guanosine triphosphatase (GTPase) units: the amino acid-sensitive Rag GTPases and the growth factor-sensitive Rheb GTPase (Inoki et al. 2003; Kim et al. 2008; Sancak et al. 2008; Zhang et al. 2003). When amino acids become abundant, mTORC1 is translocated from the cytosol and recruited to the lysosomal surface by the Rag GTPases where, under the conditional presence of growth factors, mTORC1

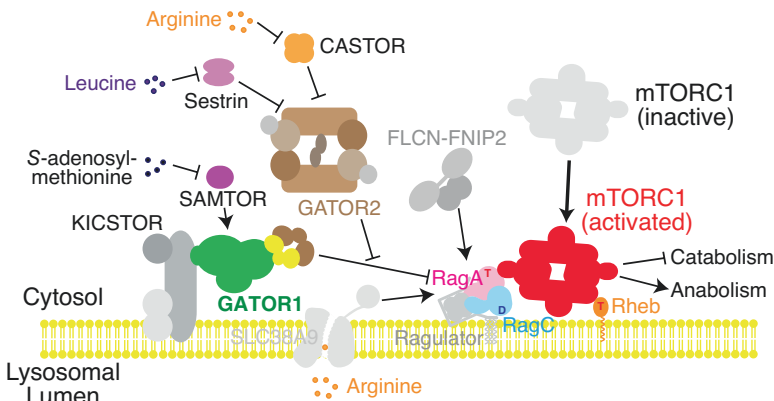


Fig. 12.1 Overview of the mTORC1 signaling pathway. Selected protein components of the mTORC1 pathway are shown, focusing on the GATOR1 complex, its downstream effectors, and its upstream regulators. Amino acid sensors either directly modulate the activity of GATOR1, such as SAMTOR, or through GATOR2, such as Sestrin2 and Castor1

kinase activity is stimulated by the Rheb GTPase (Efeyan et al. 2013; Long et al. 2005; Sancak et al. 2010). In this arrangement, the physical recruitment of mTORC1 to the lysosome allows it to simultaneously assess the levels of dietary (cytosolic) and recycled (lysosomal) nutrients, while the coordination between the Rag and Rheb GTPases ensures that mTORC1-mediated anabolism is initiated only when the cell has adequate and complementary resources.

Nutrient-sensing proteins converge on the Rag GTPases to transduce amino acid signals to mTORC1 and activate the pathway when nutrient stores are full. The Rag GTPases function as heterodimers composed of a RagA or RagC subunit bound to a RagB or RagD subunit (Nakashima et al. 1999; Schürmann et al. 1995; Sekiguchi et al. 2001). Each subunit contains a diametrically loaded N-terminal nucleotide binding domain (NBD) and a C-terminal roadblock domain (CRD), the latter of which is essential for their stable heterodimerization (Gong et al. 2011; Jeong et al. 2012; Levine et al. 2013; Shen et al. 2017). In contrast to canonical signaling GTPases that function as monomers, the Rag GTPases are obligatory heterodimers (Sekiguchi et al. 2001). Intersubunit communication within the Rag GTPase heterodimer locks it into one of two distinct conformational states, an “off” state and an “on” state, only the latter of which can bind mTORC1 and effectively tether it to the lysosome. In the on-state conformation, the NBD of RagA is loaded with GTP, and the NBD of RagC is loaded with GDP ($\text{GTP RagA-RagC}^{\text{GDP}}$). Conversely, in the off-state conformation, the respective nucleotide loading states are reversed ($\text{GDP RagA-RagC}^{\text{GTP}}$), resulting in release of mTORC1 into the cytosol where its kinase domain cannot be activated by the Rheb GTPase (Kim et al. 2008; Sancak et al. 2008; Shen et al. 2017).

The combination of intersubunit communication and an extremely low rate of intrinsic GTP hydrolysis by the Rag GTPase heterodimer results in a stably locked nucleotide loading state that must be reversed by external catalysis (Shen et al. 2017). This built-in lock imparts stability to mTORC1 signaling, reducing susceptibility to mis-activation or mis-inactivation that is out of tune with the nutrient state of the cell. Moreover, it adds a new layer of mTORC1 regulation by leading the Rag GTPases to outsource the catalysis of their nucleotide binding pockets to specialized GTPase activating proteins (GAPs) or guanine nucleotide exchange factors (GEFs). Several putative GEFs for RagA/B have been proposed, including Regulator/LAMTOR (late endosomal/lysosomal adaptor and MAPK and mTOR activator) and SLC38A9 (solute Carrier Family 38 Member 9) (Bar-Peled et al. 2012; Jung et al. 2015; Shen and Sabatini 2018; Wang et al. 2015). Folliculin (FLCN) in complex with folliculin-interacting protein 2 (FNIP2) is a GAP for RagC/D and promotes the on-state conformation of the Rag GTPases (Petit et al. 2013; Shen et al. 2019a; Tsun et al. 2013). On the other hand, GAP Activity Toward Rags-1 (GATOR1) is the only known GAP for RagA/B, promoting the Rag GTPases to adopt the off-conformation that inhibits mTORC1 signaling (Bar-Peled et al. 2013; Egri et al. 2022; Panchaud et al. 2013a). Here, we review the history and current state of knowledge regarding the mechanism of mTORC1 inhibition by the human GATOR1 protein complex.

Discovery of GATOR1

The function of GATOR1 in the mTORC1 pathway was identified between 2009 and 2013 by several independent research groups searching for negative regulators of mTORC1 signaling in response to nutrient starvation. At that time, it was understood that the asymmetrically loaded Rag GTPase heterodimer, with GTP-loaded RagA and GDP-loaded RagC, is critical for mTORC1 activation in the presence of certain amino acids, such as leucine or arginine (Kim et al. 2008; Sancak et al. 2008), and it was hypothesized that GEFs and GAPs are necessary for switching the GTPase nucleotide loading state. However, identifying a GAP for RagA/B that functioned to relay amino acid deficiency to mTORC1 eluded researchers. This challenge stemmed from the inherently transient nature of enzyme-substrate interactions *in vivo*, and its solution would require robust approaches to investigating protein-protein interactions (Back et al. 2003; Chichili et al. 2013; Mattson et al. 1993).

The first clue arrived when a systemic screen of the yeast genome revealed two proteins, *npr2* and *npr3*, which are required for mTORC1 inhibition in the absence of a rich nitrogen source such as ammonia (Neklesa and Davis 2009). Unlike mammals, yeast are prototrophic and can synthesize all amino acids *de novo* so long as nitrogen is present. Thus, either nitrogen limitation or the use of amino acid-specific auxotrophs is necessary to study mTORC1 inhibition in yeast. Subsequent research identified that *npr2* and *npr3* function in complex with a novel protein, *iml1*, to positively regulate autophagy (Wu and Tu 2011), and individual knockdown of *npr2*, *npr3* or *iml1* led to mTORC1 hyperactivation that is resistant to nutrient starvation (Panchaud et al. 2013a). The *npr2-npr3-iml1* complex was found to inhibit mTORC1 signaling through GAP activity toward Gtr1, which is the yeast ortholog of RagA (Panchaud et al. 2013a). At the same time, a similar mechanism was discovered in human cells when crosslinking of the Rag GTPase in human cancer cells resulted in the co-immunoprecipitation (co-IP) of NPRL2, NPRL3, and DEPDC5 (the human orthologs of *npr2*, *npr3*, and *iml1*, respectively). Subsequent *in vitro* assays further validated that NPRL2, NPRL3, and DEPDC5 work in complex to strongly stimulate GTP hydrolysis by RagA/B, but not by RagC/D (Bar-Peled et al. 2013). Thus, the GATOR1 complex, named aptly after its mechanism of action, was revealed as the long sought-after negative regulator of the amino acid-sensing branch of the mTORC1 pathway.

Architecture of GATOR1

The human GATOR1 protein complex is a 289 kDa heterotrimer composed of three evolutionarily conserved subunits: DEP domain-containing protein-5 (DEPDC5), nitrogen permease regulator 2-like (NPRL2) protein, and nitrogen permease regulator 3-like (NPRL3) protein. Recent structural investigations using cryogenic

electron microscopy (cryo-EM) have begun to resolve the domain architecture and quaternary assembly of the GATOR1 subunits (Egri et al. 2022; Shen et al. 2018).

DEP Domain-Containing Protein-5

DEPDC5 is the largest subunit (181 kDa) of the GATOR1 complex and is encoded by the ubiquitously expressed 48-exon *DEPDC5* gene (*KIAA0645*) located on chromosome 22 (22q12.2-q12.3). At least eight isoforms have been reported by alternative splicing, the longest of which is a 1603-amino-acid sequence organized into five presently defined domains: an N-terminal domain (NTD); a structural axis for binding arrangement (SABA) domain; a steric hindrance for enhancement of nucleotidase activity (SHEN) domain; a dishevelled, Egl-10, and Plekstrin (DEP) domain; and a C-terminal domain (CTD) (Fig. 12.2). In addition to its five domains, DEPDC5 contains two long flexible linkers, Lasso A and Lasso B, which connect the SABA-SHEN and SHEN-DEP domains, respectively.

The NTD (aa. 1–165) of DEPDC5 contains two asymmetric lobes, lobe A and lobe B, which are connected by a short flexible loop. Lobe B, characterized by six antiparallel β -strands and a single α -helix, packs into the structure to make contacts with the SHEN and SABA domains. The smaller lobe A, characterized by three antiparallel β -strands and a single α -helix, extends out of the structure and may function as a protein adaptor (Shen et al. 2018; Shiozawa et al. 2004).

Immediately following the NTD is the globular SABA domain (a.a. 166–427), which is arranged in six mostly parallel β -strands that are encircled by four α -helices. As its name suggests, this domain mediates interactions with NPRL2 and NPRL3

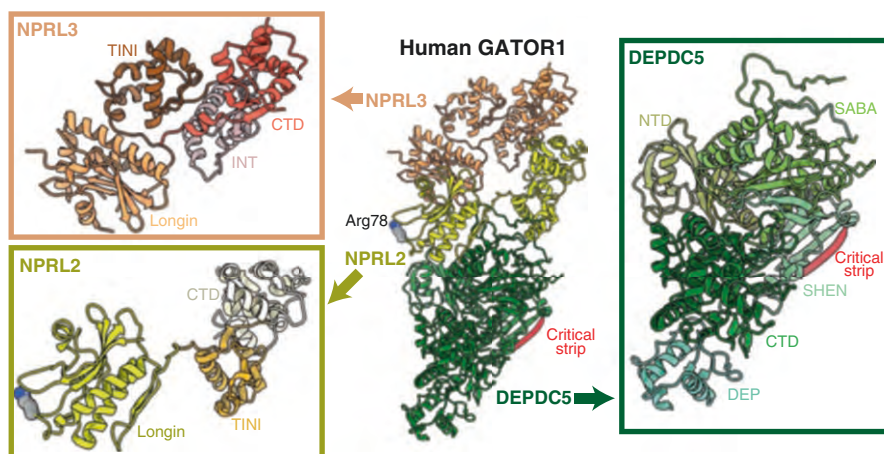


Fig. 12.2 Architecture of the human GATOR1 complex. GATOR1 is a heterotrimeric protein complex consisting of DEPDC5 (green), NPRL2 (yellow), and NPRL3 (brown). The domain arrangements for each subunit are shown accordingly

that are critical for GATOR1 assembly. Interestingly, this domain also shares structural similarity with the ligand-binding domains of flavodoxin reductase and integrin α -L (Ingelman et al. 1997; Qu and Leahy 1995; Shen et al. 2018), a property that may hint at the evolutionary origin of the DEPDC5 protein before its interactions with NPRL2 and NPRL3 were established.

The SHEN domain (a.a. 720–1010) is connected to the SABA domain via Lasso A. It consists of four antiparallel β -strands arranged in a β -meander that is flanked by two α -helices. It contains two disordered regions, loop S and linker S, the latter of which mediates interactions between the NTD and SABA domains. An eight residue β -strand situated farthest from the SHEN/SABA/NTD interface has been named the “critical strip” for its interaction with the nucleotide-binding domain of RagA at the inhibitory mode (see section “The GATOR1 Inhibitory Mode”).

The DEP domain (a.a. 1175–1270) is connected to the SHEN domain via Lasso B. This small domain is characterized by a triple-helix bundle and an antiparallel β -hairpin arm at its C-terminus. The DEP domain is a common feature of GAPs involved in regulating signal transduction (Burchett 2000). But despite previous observations, the DEP domain within DEPDC5 neither mediates interaction with the Rag GTPase heterodimer, nor catalyzes the GTP hydrolysis reaction. Alternatively, it has been suggested that DEP domains may function to localize proteins to intracellular membrane signaling hubs (Wong et al. 2000), although this remains to be tested in the case of DEPDC5.

The CTD (a.a. 1291–1603), which immediately follows the DEP domain, contains two analogous lobes arranged in a pseudo twofold rotational symmetry. Each lobe forms a five-stranded β -meander, which cradles a single α -helix. The CTD is centrally wedged within the DEPDC5 core and makes extensive contacts with the DEP, SHEN, and SABA domains.

NPRL2 and NPRL3

NPRL2 (44 kDa) and NPRL3 (64 kDa) form a stable heterodimer within the GATOR1 complex. The human NPRL2 protein is encoded by an 11-exon gene (*NPRL2*) located on chromosome 3 (3p21.3), while the human NPRL3 protein is encoded by a 15-exon gene (*NPRL3*) on chromosome 16 (16p13.3), located adjacent to the alpha-globin gene cluster. At least two isoforms of NPRL2 and five isoforms of NPRL3 have been reported by alternative splicing.

NPRL2 and NPRL3 exhibit structural homology and share many of the same domains (Fig. 12.2). Starting at the N-terminus, both proteins contain an N-terminal Longin domain (LD), followed by a tiny intermediary of NPRL [2] that interacts [with DEPDC5] (TINI) domain, and end with a CTD. In addition, NPRL3 contains an intermediary (INT) domain that links its CTD and TINI domains and accounts for its larger mass.

The Longin Domain (LD) is a conserved feature of mammalian GAP family protein complexes, and LD-containing proteins have been associated with nutrient

sensing, membrane organization, vesicular formation, and vesicular trafficking (De Franceschi et al. 2014; Levine et al. 2013). Structurally, the LD is characterized by a five stranded β -meander sandwiched between a short α -helix on its top face and two longer α -helices on its bottom face (Tochio et al. 2001). LDs tend to participate in side-by-side dimerization to produce a larger functional platform consisting of a ten-stranded β -meander with four α -helices forming a hydrophobic-leaning surface and two α -helices sequestering to the opposite surface. The hydrophobic surface may associate with membranes or form a hydrophobic core of a protein complex, while the more hydrophilic surface may mediate protein-protein interactions.

LDs share structural and functional similarity with roadblock domains (RBD), as RBDs are cited as the evolutionary ancestors of LDs (Levine et al. 2013; Zhang et al. 2017). Both LDs and RBDs can provide binding platforms for small GTPases and/or associate with membrane bound organelles; however, there are no known LDs in prokaryotes, suggesting that LD-containing proteins may have adapted the topology of bacterial and archaeal RBDs to accommodate expanding roles in eukaryotic membrane trafficking and vesicular formation (Levine et al. 2013; Miertzschke et al. 2011; Schlenker et al. 2006). Structural models of LD-containing GAP subunits have revealed a conserved arginine residue, for example, Arg78 in NPRL2 (Shen et al. 2019b), Arg164 in FLCN (Lawrence et al. 2019; Shen et al. 2019a), and Arg147 in SMCR8 (Nörpel et al. 2021), which functions as a catalytic arginine finger to complete the active site of its target GTPase and stimulate GTP hydrolysis. Functional assays mutating this critical arginine residue have consistently ablated GAP activity. This critical residue will be discussed below in the context of GATOR1.

Following the LD, both NPRL2 and NPRL3 contain a structurally similar, but functionally different, TINI domain. Both TINI domains contain four short α -helices that form a bowl shape upon which two antiparallel β -strands form a lid. The TINI domain of NPRL2 mediates interactions with DEPDC5, while the TINI domain of NPRL3 points outward and may mediate interaction with other mTORC1 pathway proteins. Likewise, the exposed INT domain of NPRL3, which lacks a homolog in NPRL2 and is characterized by a similar fold as the TINI domain, may also be involved in facilitating protein-protein interactions.

At their C-terminus, both NPRL2 and NPRL3 contain a homologous CTD characterized by five and four α -helices, respectively, and through which the two subunits form extensive contacts.

Binding Interfaces Between the GATOR1 Subunits

Cryo-EM structural models of the human GATOR1 complex (PDB: 7T3A, 7T3B, 7T3C) have resolved the binding interfaces between DEPDC5, NPRL2, and NPRL3. The overall structure of the complex loosely resembles a flower with a stem in which the NPRL2-NPRL3 heterodimer forms the flowerhead and DEPDC5 forms the stem. NPRL2 and NPRL3 heterodimerize by making extensive complementary

contacts at their respective N- and C-terminals (Fig. 12.3a). This interface is stabilized by three intersubunit contact sites: (1) NPRL2-NLD and NPRL3-NLD (NLD interface), (2) NPRL2-CTD and NPRL3-CTD (CTD interface), and (3) NPRL2-CTD and NPRL3-INT. The NLD interface follows a standard Longin domain dimerization pattern and is mediated by two β -strands, one from each subunit (Leu53-Thr58 of NPRL2 and Lys83-Ile88 of NPRL3), which form a stable antiparallel arrangement. The NLD interface is further reinforced by hydrogen bonding between Asp38 on NPRL2 and Lys76 on NPRL3. The CTD interface forms through extensive hydrogen bonding involving Arg279, Gln252, and Asp256 of NPRL2, and Arg521, Arg523, Ser549, and Val550 of NPRL3, while proline-rich disordered regions within the CTD of NPRL2 and the INT domain of NPRL3 further reinforce the heterodimer.

The NPRL2-NPRL3 heterodimer is anchored to DEPDC5 through extensive contacts between the NLD and TINI domains of NPRL2 and the globular SABA domain of DEPDC5 (Fig. 12.3b). Stretches within three disordered loops of the

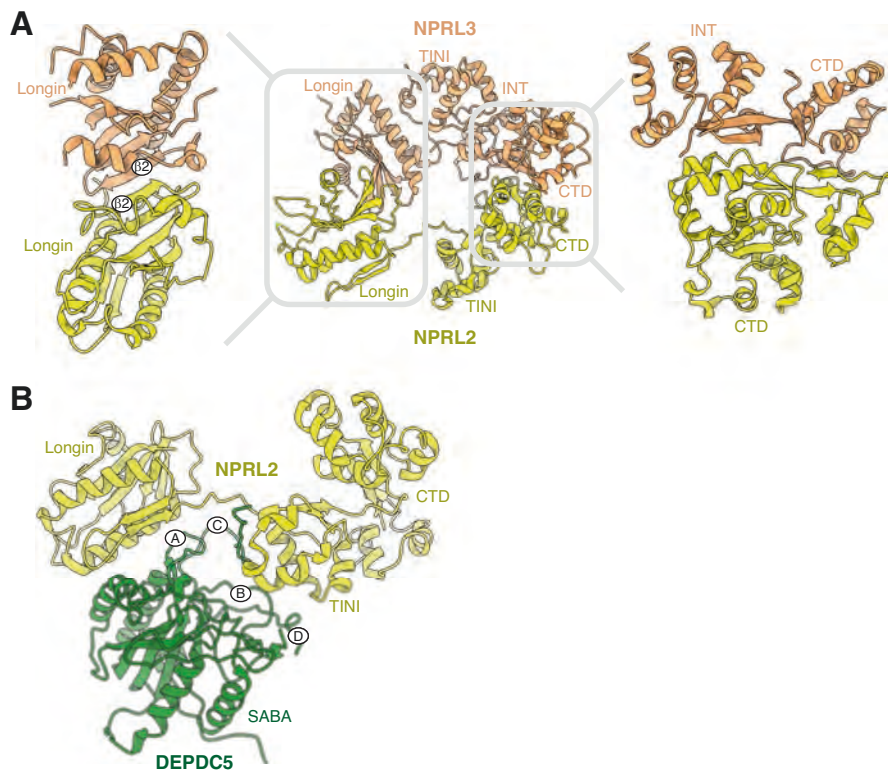


Fig. 12.3 Interaction between the GATOR1 subunits. (a) NPRL2-NPRL3 interactions. NPRL2 (yellow) and NPRL3 (brown) form extensive interactions through three pairs of domains, Longin-Longin, CTD-CTD, and CTD-INT. (b) DEPDC5-NPRL2 interactions. DEPDC5 (green) utilizes the four loops (loop A-loop D) to form redundant interactions with NPRL2 (yellow)

SABA domain, comprising residues 373–375 (loop B), 418–427 (loop C), and 730–735 (loop D), tuck into the α -helices of the TIN1 domain, while portions of loop C (415–417) form additional contacts with a disordered region situated between the NLD and the TIN1 domain of NPRL2. Another sequence (loop A, 187–193) on the SABA domain and a nearby α -helix (273–279) stably interface with a β -sheet hairpin motif immediately following the NLD of NPRL2.

Mutational studies probing segments of the SABA domain have identified redundancy within the contacts between DEPDC5 and NPRL2. Specifically, while mutations within individual loops fail to disrupt the NPRL2-DEPDC5 interaction, simultaneous mutations at both loop A and loop B are required to abolish binding, suggesting the enhanced binding interactions to maintain the complex integrity.

GATOR1 Binds the Rag GTPases in Two Modes

GATOR1 is a dedicated GAP for the RagA subunit of the Rag GTPase heterodimer and is required for the negative regulation of mTORC1 signaling when nutrients are limited. Although GAPs have been shown to employ a variety of catalytic mechanisms, canonical GAPs typically utilize a conserved arginine finger (Ahmadian et al. 1997; Nagy et al. 2016; Scheffzek et al. 1997) or an asparagine thumb (Daumke et al. 2004) to stimulate GTP hydrolysis of their target GTPase. Genetic approaches identified that GATOR1 uses a canonical GAP mechanism governed by a conserved arginine finger (Shen et al. 2019b) to accelerate the conversion of active $\text{GTP RagA-RagC}^{\text{GDP}}$ to inactive $\text{GDP RagA-RagC}^{\text{GTP}}$, resulting in the detachment of mTORC1 from the lysosomal surface.

Cryo-EM structural models of GATOR1 bound to the Rag GTPase heterodimer have resolved two distinct modes in which GATOR1 interacts with the Rag GTPase heterodimer (Fig. 12.4a): (1) an inhibitory mode that lacks GAP activity and (2) a GAP mode that specifically accelerates GTP hydrolysis by RagA. The structural model of the inhibitory mode was the first to be discovered, and its lack of GAP activity temporarily confounded the search for the catalytic GAP site on GATOR1. It was the later discovery of the GAP structure that solidified Arg78, located on a dynamic loop within the NLD of NPRL2, as the catalytic arginine finger of the GATOR1 complex.

The GATOR1 Inhibitory Mode

The inhibitory binding mode was the first structural model to be resolved (Shen et al. 2018) and is configured such that DEPDC5 is the only GATOR1 subunit that shares an interface with the Rag GTPases. In this mode, the so-called critical strip of the SHEN domain of DEPDC5, comprising residues 771–778, directly contacts the GTP-loaded nucleotide binding pocket of RagA (Fig. 12.4b). The interface is

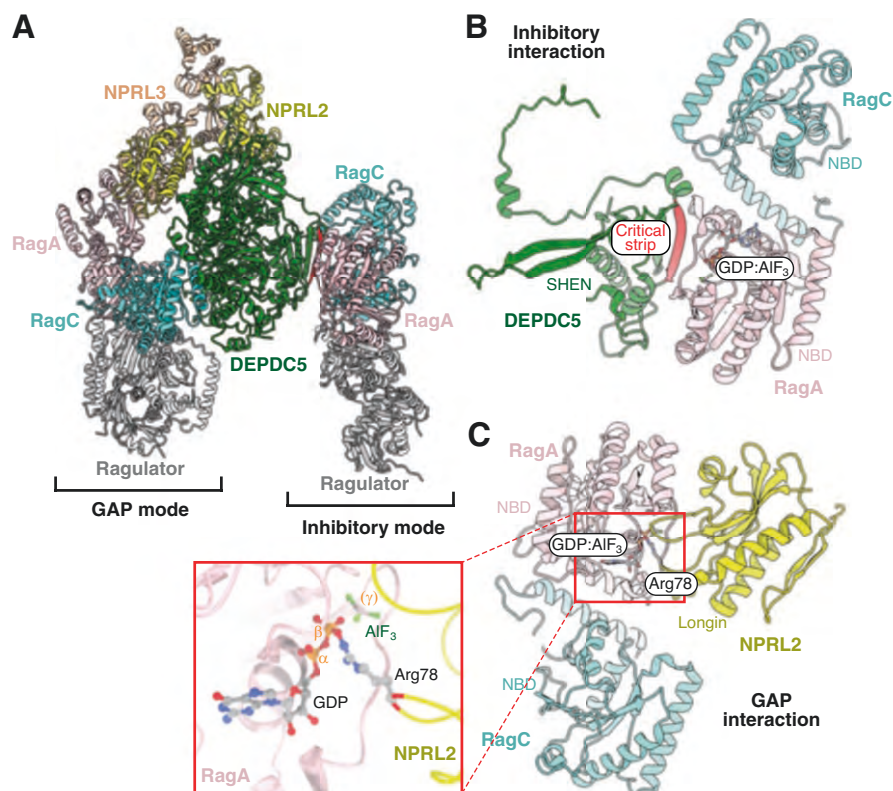


Fig. 12.4 GATOR1 contacts the Rag GTPase heterodimer via two modes. (a) Two binding sites on GATOR1 can be occupied by the Rag GTPase heterodimer: an inhibitory mode mediated by the critical strip (red) on DEPDC5 and a GAP mode mediated by the Arg78 residue on NPRL2. The catalytic finger of NPRL2 stimulates GTP hydrolysis by RagA only in the GAP mode. (b) GATOR1-Rag interaction in the inhibitory mode. The critical strip (red, a.a. 770–777 of DEPDC5) contacts the nucleotide binding pocket of RagA (pink). GDP:AlF₃ is a non-hydrolyzable GTP analog used to stabilize the complex. (c) GATOR1-Rag interaction in the GAP mode. Arg78 of NPRL2 inserts into the nucleotide binding pocket of RagA and points toward the β- and γ-phosphate of the bound nucleotide to stimulate GTP hydrolysis

stabilized by four hydrogen bonds: (1) RagA-Arg38 and DEPDC5-Glu772, (2) RagA-Arg38 and DEPDC5-Thr771, (3) RagA-Leu39 and DEPDC5-Gly773, and (4) RagA-Ser48 and DEPDC5-Asp776. Co-immunoprecipitation assays in which various residues within the critical strip were individually mutated revealed that Tyr775 was essential for the DEPDC5-RagA interaction. However, Tyr775 does not directly participate in hydrogen bonding with any RagA residues. Instead, it likely plays an essential role in maintaining the β-strand conformation of the DEPDC5 critical strip, bending it in such a way as to bring the other residues within hydrogen-bond proximity of each other (Shen et al. 2018).

Still, there was a problem: the initial cryo-EM structural model of GATOR1 bound to the Rag GTPases did not reveal a GATOR1 GAP mechanism; indeed, the

critical strip of GATOR1 that directly binds RagA lacks a catalytic residue that can stimulate GTP hydrolysis by RagA. Thus, how does GATOR1 function as a GAP for RagA?

Searching for the GAP in GATOR1

To interrogate this question, researchers employed kinetics-based assays to probe the binding and enzymatic activity of the GATOR1[DEPDC5(Y775A)] mutant, which had disrupted the DEPDC5-RagA interaction (Shen et al. 2018). As expected, the GATOR1[DEPDC5(Y775A)] mutant and RagA exhibited weaker binding kinetics, as measured by a 15-fold increase in the dissociation constant (K_d). Yet, surprisingly, the mutant exhibited enhanced GAP activity toward RagA, as measured by a 20-fold increase in its chemical step (k_{cat}). This finding suggested the existence of an alternative binding site, one which enables GATOR1 GAP activity toward RagA and in which the DEPDC5 critical strip is not involved.

To search for a different binding mode, the GAP activity and binding propensity of individual GATOR1 components (DEPDC5 or the NPRL2-NPRL3 heterodimer) was measured against the Rag GTPases using a series of kinetics-based assays (Shen et al. 2018). In vitro, the NPRL2-NPRL3 heterodimer was found to exhibit strong GAP activity toward RagA, while DEPDC5 alone had no GAP activity. In addition, much higher concentrations of NPRL2-NPRL3 than full GATOR1 were needed to stimulate hydrolysis, suggesting that DEPDC5 may play a stabilizing role within the complex. Moreover, the addition of the NPRL2-NPRL3 heterodimer was found to increase GTP hydrolysis by RagA in the presence of full GATOR1, indicating that full GATOR1 exists in a heterogeneous population with some complexes engaging in GAP activity while others inhibit it. Overall, this supported the structural finding that DEPDC5 mediates an inhibitory role.

These in vitro results corroborated cell-based findings that had localized the GATOR1 GAP site to the NPRL2-NPRL3 heterodimer (Shen et al. 2019b). DEPDC5 knockout cells expressing the NPRL2-NPRL3 heterodimer were found to inhibit mTORC1 signaling, and while this inhibition was further increased by the presence of DEPDC5, it was not dependent on it, indicating that the NPRL2-NPRL3 heterodimer carries out GAP activity while DEPDC5 functions to stabilize and promote the interaction. The GAP site was further localized to NPRL2 by introducing NPRL2 arginine mutants into an NPRL2-knockout cell line and testing their ability to inhibit mTORC1 signaling in response to nutrient depletion. The result of a single loss-of-function mutation pointed toward Arg78 as the likely catalytic residue of GATOR1, but this finding remained to be structurally confirmed.

Altogether, these data strongly suggested that the catalytic GAP site of GATOR1 is located on NPRL2 and that DEPDC5 functions to either stabilize the GATOR1-Rag GTPase bond or modulate GAP activity by blocking the RagA nucleotide binding pocket when amino acids are abundant.

The GATOR1 GAP Mode

The mystery was solved when a new cryo-EM structural model revealed a second GATOR1 and Rag GTPase binding mode, which was aptly named “the GAP mode” for its depiction of the GATOR1 GAP mechanism (Fig. 12.4c) (Egri et al. 2022). The GAP mode is drastically different from the inhibitory mode. In the inhibitory mode, DEPDC5 is the only subunit of GATOR1 to contact the Rag GTPases. In contrast, in the GAP mode, both DEPDC5 and NPRL2-NPRL3 contact the Rag GTPases: NPRL2-NPRL3 binds RagA, while DEPDC5 binds RagC. Flexible β -strand hairpin loops extending from the NPRL2 and NPRL3 NLDs form a cradle for the NBD of RagA. This configuration situates the NLD of NPRL2 proximal to the RagA nucleotide binding pocket, while the NLD of NPRL3 is positioned distal to it.

The DEPDC5-RagC interface is defined by two binding sites. At the first binding site, the NTD of DEPDC5 and the external face of the RBD of RagC make extensive contacts mediated by a series of hairpin loops. At the second binding site, an α -helix within the CTD of DEPDC5 and an α -helix within the NBD of RagC form hydrogen bonds. As GATOR1 is a specific GAP for RagA, this additional, auxiliary interface was not expected (Doxsey et al. 2023).

In the GAP mode, the catalytic arginine finger was captured in action (Fig. 12.4c). Here, Arg78 of NPRL2 inserts into the nucleotide binding pocket of RagA and points directly toward the β - and γ -phosphate of the bound guanine nucleotide. This configuration is similar to other GAP-GTPase pairs (Schröter et al. 2015), as the positively charged arginine side chain can effectively disperse the negative charges that build up on the oxygen atoms during GTP hydrolysis and thus lower the activation barrier for the S_N2 reaction to accelerate the reaction rate (Coleman et al. 1994; Privé et al. 1992, p. 21).

The two binding modes are compatible in vitro, as a cryo-EM structural model shows that two copies of the Rag GTPases bind to the same GATOR1 complex simultaneously (Egri et al. 2022). However, whether such a configuration is permitted in vivo is currently unknown. Due to the opposite orientation of the Rag GTPases and their membrane anchor, Ragulator (Bar-Peled et al. 2012; Sancak et al. 2010), it is proposed that the spatial constraints imposed by the lysosomal membrane may allow only one copy of the Rag GTPases to bind GATOR1 at any given time.

Regulators of GATOR1

Thus far, we have learned that GATOR1 is an essential negative regulator of mTORC1 signaling through its GAP activity toward RagA of the Rag GTPase heterodimer. But GATOR1 does not work alone. It is, in turn, regulated by an assortment of upstream effectors. Work by several groups has identified several essential regulators of GATOR1 function in the mTORC1 pathway, including GATOR2 (Bar-Peled et al. 2013; Panchaud et al. 2013b; Peng et al. 2017; Valenstein et al. 2022; Wei et al. 2014; Wolfson et al. 2017), Sestrin2 (Chantranupong et al. 2014; Kim

et al. 2015; Parmigiani et al. 2014; Saxton et al. 2016b; Wolfson et al. 2016), Castor1 (Chantranupong et al. 2016; Saxton et al. 2016a), KICSTOR (Wolfson et al. 2017), and SAMTOR (Gu et al. 2017; Tang et al. 2022).

GATOR1 Is Antagonized by GATOR2

The human GATOR2 complex antagonizes GATOR1 to positively regulate mTORC1 signaling in the presence of amino acids. It inhibits GATOR1 GAP activity by transducing signals from dedicated amino acid sensors toward GATOR1. The two complexes were discovered through co-immunoprecipitation (Dokudovskaya et al. 2011; Neklesa and Davis 2009); yet despite sharing a similar name, they are vastly different. While GATOR1 is a 289 kDa heterotrimer, the pentameric GATOR2 complex is nearly four times in mass (1.1 MDa).

The architecture of the GATOR2 complex reveals a modular, octagonally framed cage-like structure composed of multiple copies of five subunits: Mios, WDR24, WDR59, Seh1L, and Sec13 (Tafur et al. 2022; Valenstein et al. 2022). Its five subunits associate in three different combinations to form eight heterodimers consisting of four Mios-Seh1L, two WDR24-Seh1L, and two WDR59-Sec13 pairs. Despite sharing little sequence homology, Mios, WDR24, and WDR59 display surprisingly complementary domain architecture: an N-terminal WD40 domain is connected to a 3-stranded β -blade, followed by an α -solenoid trunk, and culminating in a C-terminal zinc finger-RING domain. The prominent motif of the remaining two subunits, Seh1L and Sec13, is an open 6-bladed β -propeller. The domains of Mios, WDR24, and WDR59 fold into an *ancestral coatomer element* (ACE)-1-like motif in which the elongated superhelical trunk of the α -solenoid domain folds back on itself to form an α -solenoid “crown.” These ACE1-like motifs form the stable core of GATOR2 through complementary crown-crown and RING-RING interactions which circularize in a spartan-like handshake. Around this stable circular core, GATOR2 coordinates two distinct faces: reaching toward each other from opposing sides of the ring, the β -propellers of two Mios subunits connect to form a central dimer, while on the reverse face of the ring, another pair of extended Mios β -propellers fails to make contact. To complete the scaffold, the 6-bladed β -propellers of Seh1L and Sec13 provide joint reinforcements by accepting a β -blade from Mios and WDR24 or WDR59, respectively. The remaining β -propellers protrude from the scaffold in various orientations and are likely involved in the recruitment of accessory proteins at the GATOR2 periphery. Specifically, the β -propellers of WDR59, WDR24, and Mios are likely docking sites for GATOR1, Sestrin2, and Castor1, respectively (see next section) (Tafur et al. 2022; Valenstein et al. 2022).

It has been speculated that the GATOR2 RING domains, which are a common feature of E3 ubiquitin ligases, enable GATOR2 to regulate GATOR1 via ubiquitin-mediated degradation. However, as of yet, functional studies have failed to identify such a catalytic function, and structural models support an architectural role for the RING domains (Valenstein et al. 2022).

Amino Acid Sensors Upstream of GATOR2

To date, “sensors” that signal amino acid availability to the Rag GTPases through a GATOR2-GATOR1 cascade have been identified for two amino acids: leucine and arginine. Leucine is sensed by the Sestrin family of proteins (Sestrin2), while arginine is sensed by Castor1 (Chantranupong et al. 2016; Wolfson et al. 2016). When leucine or arginine are limiting, Sestrin2 and Castor1 in their apo form inhibit the GATOR2 complex by binding to dedicated β -propeller domains: Castor1 binds the Mios β -propeller, and Sestrin2 binds the WDR24-Seh1L β -propeller (Jiang et al. 2023). Of note, the WDR59-Sec13 single chain β -propeller was found to mediate the GATOR2-GATOR1 interaction (Tafur et al. 2022; Valenstein et al. 2022).

The Sestrins and Castors harbor self-regulatory mechanisms that mediate their binding to GATOR2. Crystal structures of the Sestrin2 (55 kDa) protein reveal two globular domains (an NTD and CTD) linked by an inter-domain loop (Saxton et al. 2016b). Portions of the NTD and CTD contain the GATOR2 binding site, while a nearby region of the CTD harbors a hydrophobic binding pocket with a “lid-latch” mechanism regulated by the inter-domain loop (Saxton et al. 2016b). When leucine binds, the lid-latch conceals the bound leucine with a patch of hydrophilic residues (Saxton et al. 2016b). It has been proposed that activation of the lid-latch puts strain on the inter-domain loop and leads to allosteric inhibition of the Sestrin2-GATOR2 interaction (Yao et al. 2017). The loss of the Sestrin2-GATOR2 interaction in the presence of leucine allows GATOR2 to antagonize GATOR1. In turn, when Sestrin2 binds GATOR2, the GATOR2-GATOR1 interaction is disrupted (Kim et al. 2015). Thus, Sestrin2 promotes GATOR1 GAP activity by inhibiting GATOR2.

The Castor family proteins exist as homo- and heterodimers, and each subunit contains two tandem ACT domains that interface to form an arginine binding pocket (Chantranupong et al. 2016; Saxton et al. 2016a). Castor1, but not Castor2, is able to potently bind arginine (Chantranupong et al. 2016). The GATOR2 binding site on Castor1 is at the opposite-facing side of the arginine binding pocket (Saxton et al. 2016a). Because Castor1 ACT domains share structural homology with regulatory aspartate kinase domains involved in allosteric feedback inhibition (Grant 2006; Lang et al. 2014), it has been hypothesized that the binding of arginine reorganizes the dimer to inhibit the GATOR2 interaction. Indeed, a key asparagine residue essential for the GATOR2 interaction appears deeply buried in the arginine-bound structure of Castor1, providing an explanation for why the binding of arginine disrupts the Castor1-GATOR2 interaction (Saxton et al. 2016a). In turn, the disruption of the Castor1-GATOR2 interaction by arginine leads to mTORC1 activation, presumably through GATOR1 inhibition.

Many questions remain regarding the sensor-GATOR2-GATOR1 axis. What happens when the sensors detach from GATOR2, and how does this lead to GATOR1 inhibition? Conversely, how does the sensor-GATOR2 interaction promote GATOR1 activity? Are these interactions direct, or do they involve a third-party interactor essential for function or localization? How many sensor proteins can bind the large GATOR2 scaffold at once, and do their binding sites compete with GATOR1? Unraveling the mechanisms behind GATOR1 and GATOR2 interactions remains a compelling challenge.

GATOR1 Is Localized to the Lysosome by KICSTOR

Much of the nutrient sensing machinery required for mTORC1 signaling localizes to the lysosomal surface, including the Rag GTPases, which are tethered to it by Ragulator (Sancak et al. 2010; Zhang et al. 2017). Thus, in order for GATOR1 to negatively regulate the Rag GTPases, it must also localize to the lysosome. In mammalian cells, GATOR1 localization is dictated by the lysosome-associated KICSTOR complex (Peng et al. 2017; Wolfson et al. 2017), named so after its four subunits (KPTN, ITFG2, c12orf66, and SZT2). KICSTOR interacts with GATOR1 through its SZT2 subunit independent of amino acid levels (Wolfson et al. 2017) and is essential for mediating the GATOR1 interaction with the Rag GTPases or its antagonist, GATOR2 (Wolfson et al. 2017). In cells lacking KICSTOR, GATOR1 does not interact with the Rag GTPases (Wolfson et al. 2017). Furthermore, co-IP and crosslinking experiments have revealed that KICSTOR can interact with GATOR1 alone or in complex with GATOR1 and GATOR2 in an SZT2-dependent manner (Peng et al. 2017; Wolfson et al. 2017). The cooperative binding of SZT2-GATOR1-GATOR2 was thus named the STZ2-Orchestrated GATOR (SOG) complex (Peng et al. 2017). Disruption of the SOG complex compromised mTORC1 activation in the presence of amino acids, indicating that SOG formation inhibits GATOR1 GAP activity (Peng et al. 2017). Thus, KICSTOR functions as a modular scaffold that positions GATOR1 at the lysosome enabling it to rapidly respond to amino acid insufficiency and, when amino acids are abundant, enables GATOR2 to suppress GATOR1 (Valenstein et al. 2022).

GATOR1 and SAMTOR

SAMTOR is a dedicated *S*-adenosylmethionine (SAM) sensor that binds to and promotes GATOR1 GAP activity in the absence of SAM (Gu et al. 2017). Unlike Sestrin2 and Castor1, which bind and inhibit GATOR2 in their apo states, when SAM is limiting, SAMTOR binds GATOR1 and promotes its GAP activity. Conversely, when SAM levels rise, SAM binds to SAMTOR and disrupts its interaction with GATOR1, leading to mTORC1 activation. The bigger consequence of this interaction comes into focus when we consider that GATOR1 can exist within two mutually exclusive supercomplexes: (1) the GATOR1-KICSTOR supercomplex, which promotes GATOR1 GAP activity toward RagA, or (2) the SOG supercomplex, which inhibits GATOR1 GAP activity. Interestingly, SAMTOR only binds the GAP-promoting complex, as evidenced by co-IP studies, and it was found to disrupt the SOG complex in a dose-dependent manner (Gu et al. 2017). This indicates that SAMTOR may function to promote GATOR1 GAP activity by interfering with GATOR1 suppression by GATOR2.

A crystal structure of the *Drosophila melanogaster* SAMTOR ortholog depicts SAMTOR in its ligand-bound and ligand-free conformations (Tang et al. 2022). Similar to Sestrin2 and Castor1, SAMTOR appears to utilize allosteric feedback inhibition to regulate its binding interactions. Mutational studies have proposed that

the binding of SAMTOR to GATOR1 is coordinated, in part, by a critical phenylalanine residue (Phe175) at its MTase domain, which is concealed following a conformational change induced by the binding of SAM (Tang et al. 2022).

SAM is the methylated product of methionine and, thus, serves as a proxy for cellular methionine levels. SAM-mediated methylation reactions are essential for many physiological processes involved in cell proliferation, including histone modification (Law and Jacobsen 2010), one-carbon metabolism (Mentch et al. 2015), and phospholipid modification (Ye et al. 2017). Therefore, having a dedicated mechanism by which to inhibit mTORC1 signaling when SAM is limiting is critical. Additional structural and functional information is needed to interrogate the nature of the SAMTOR-GATOR1-KICSTOR interaction and the mechanism by which it promotes GATOR1 GAP activity.

GATOR1 and Disease

Remarkably, at the organismal level, the mTORC1 pathway must balance the complex anabolic and catabolic needs of a variety of tissues and cell types throughout all stages of development (Efeyan et al. 2013; Sengupta et al. 2010). Given that mTORC1 regulation is essential for normal health and physiology, it is of little surprise that its dysregulation is associated with a range of human diseases, including disorders of metabolism (Lee et al. 2016), cancer (Menon and Manning 2008), neurodegeneration, (Lipton and Sahin 2014; Takei et al. 2004), and premature aging (Carroll et al. 2017; Mannick et al. 2018). Hyperactivation of mTORC1 signaling is particularly deleterious to the developing brain and cell-cycle control (Baulac 2016; Crino 2015), and this is evidenced by a startling number of *DEPDC5*, *NPRL2*, and *NPRL3* mutations associated with epilepsy (Dibbens et al. 2013; Ricos et al. 2016; Sim et al. 2016; Weckhuysen et al. 2016) and cancer (Bar-Peled et al. 2013).

Epilepsy

Mutations in GATOR1 genes are linked to a spectrum of brain-related pathologies characterized to some degree by hyperactive neurons, structural brain abnormalities, and seizures (Iffland et al. 2019). Nearly two hundred GATOR1 pathogenic variants have been linked to 37% of all familial focal epilepsy cases (Baldassari et al. 2019; Baulac 2016), representing the most frequently mutated genes seen in this disorder. Most of these mutations (85%) can be localized to regions of *DEPDC5* (Baldassari et al. 2019; Dibbens et al. 2013), but variants in *NPRL2* or *NPRL3* have been found to produce a similar phenotype (Baldassari et al. 2019; Ricos et al. 2016; Sim et al. 2016). Most cases of GATOR1-related epilepsies are inherited from an asymptomatic parent and result from loss-of-function or missense mutations within coding regions, with just 4% arising de novo (Moloney et al. 2021). Whether the mutation is somatic or originates in the germline appears to affect the pathogenic

outcome, with germline or somatic mutations contributing to non-lesional or lesional focal epilepsy, respectively. Severe disease phenotypes characterized by drug-resistant focal cortical dysplasia are the likely result of a “second-hit” somatic mutation in the brain co-occurring with an inherited germline mutation (Ribierre et al. 2018). It is reasoned that because *DEPDC5* is ubiquitously expressed starting at early embryogenesis and exhibits particularly high expression in the developing brain, the intense synaptic proliferation during fetal development produces a favorable environment for somatic *DEPDC5* neuronal mutations (Insel 2014; Ribierre et al. 2018). Startlingly, over half of GATOR1-related epilepsy cases are drug resistant and account for 10% of all sudden unexpected death in epilepsy (SUDEP) cases (Bagnall et al. 2016; Baldassari et al. 2019). Some GATOR1-associated SUDEP cases have exhibited the co-occurrence of an *NPRL3* variant linked to cardiovascular defects (Bacq et al. 2022), but subsequent investigations failed to link GATOR1-related SUDEP with a primary cardiac etiology. Instead, the present data suggests that GATOR1-related SUDEP is the result of abnormalities in the brainstem, a hub which controls cardiac and respiratory pacemakers (Devinsky et al. 2016; Faingold 2012). Of note, mutations in the GATOR1-interacting KICSTOR components, SZT2 and KPTN, are both linked to epileptic encephalopathy (Baple et al. 2014; Basel-Vanagaite et al. 2013; Pajusalu et al. 2015; Venkatesan et al. 2016).

Postmortem brain tissue specimens collected from GATOR1-epilepsy cases demonstrate hyperactivation of the mTORC1 pathway (Sim et al. 2016; Weckhuysen et al. 2016), but the exact mechanism by which GATOR1 mutations alter neuronal brain development is poorly understood. One explanation is that mTORC1 hyperactivation disrupts the normal formation of neural circuits in the brain (Baldassari et al. 2016). Since GATOR1 genes demonstrate elevated expression in brain tissue early on in development, dysregulated mTORC1 inhibition may be particularly catastrophic to the developing neuronal program.

Interestingly, a ketogenic diet, which mimics a fasting state and is characterized by low sugar, low protein, and high fat intake, has been a long prescribed therapy for idiopathic epilepsy, especially in children and adolescents (Hassan et al. 1999; Kinsman et al. 1992). More recent evidence has found that a ketogenic diet is helpful in cases of intractable focal epilepsy (Maydell et al. 2001; McDaniel et al. 2011; Rho 2017). Since glucose and amino acids are potent activators of mTORC1 signaling, and since the brain can make efficient use of ketones in place of glucose, dietary limitation may function to reduce mTORC1 activation in certain GATOR1 mutant backgrounds.

Cancer

A hyperactive mTORC1 pathway promotes tumor growth and metastasis and is a characteristic of over 80% of human cancers (Menon and Manning 2008). Perhaps as a result of reduced nutrient flow within the isolated tumor microenvironment, selective pressure may favor its uncoupling from the limitations imposed on it by the mTORC1 pathway. This is in opposition with GATOR1 function, which is to

potently inhibit mTORC1 when amino acids are low. Thus, it is common to observe *DEPDC5*, *NPRL2*, or *NPRL3* downregulation during cancer progression, and all three genes have been classified as tumor suppressors.

Recurrent loss-of-function mutations at the *DEPDC5* locus have been identified in glioblastoma, ovarian cancer, and gastrointestinal stromal tumors (GIS) (Bar-Peled et al. 2013; Pang et al. 2019). *NPRL2* is subject to frequent and early inactivation through aberrant splicing or through overlapping homozygous deletions in lung cancer and breast cancer (Jayachandran et al. 2010; Lerman et al. 2000).

One feature of the frequency with which GATOR1 components are downregulated in cancer has been their potential to serve as biomarkers in early detection and prognosis. For example, *NPRL2* expression was found to be an independent prognostic factor in hepatocellular carcinoma and osteosarcoma, with expression levels decreasing in step with cancer progression (Gao et al. 2013; Otani et al. 2009). In colorectal tumors, where disease progression is also positively correlated with *NPRL2* downregulation, *NPRL2* mRNA in tumor tissue was found to correlate with blood *NPRL2* mRNA levels, making the latter a convenient marker for early-stage colorectal cancer detection (Liu et al. 2014). Recently, an eight-gene Tregs-related risk signature identified *NPRL3*, and no other GATOR1 or mTORC1 components, as an immune-mediated prognostic marker for drug-response in endometrial carcinoma (Liu et al. 2021).

In general, normal or overexpression of GATOR1 components is associated with a positive prognosis and sensitivity to chemotherapies, while underexpression promotes proliferation and drives tumor drug resistance (Solanki et al. 2023). One explanation is that mTORC1 hyperactivation may robustly suppress cellular autophagy while driving the proliferation and drug-induced selective adaptation of cancer cells. However, this is not always the case, especially with regard to *NPRL2*. On the one hand, *NPRL2* overexpression inhibits non-small-cell lung tumor growth and increases chemosensitivity to cisplatin (Jayachandran et al. 2010). On the other hand, *NPRL2* overexpression was positively associated with hyperproliferation of castration-resistant prostate cancer and reduced sensitivity to everolimus (Chen et al. 2019), suggesting the complicated nature of the rewiring of the signaling network related to GATOR1 when its subunit level changes.

Conclusion

By allowing cells to put the brakes on growth and proliferation when nutrients are limited, GATOR1 has emerged as one of the important negative regulators of the mTORC1 signaling pathway. Structural models of GATOR1 and its key interactors have come a long way in elucidating the mechanism by which it carries out its critical function. So far, we have learned that (1) GATOR1 is a specific GAP for the Raga GTPase, and it uses a conserved arginine finger to stimulate its GTP hydrolysis; (2) it interacts with the Rag GTPases in at least two functionally opposing and biologically relevant configurations; (3) it requires the large KICSTOR complex for

its lysosomal localization; (4) it is antagonized by the GATOR2 complex and its associated leucine and arginine sensors, Sestrin2 and Castor1; (5) its GAP activity is promoted by a nutrient sensor, SAMTOR; and (6) the consequence of its dysregulation has the potential to permeate key physiological processes. However, there remain many more questions than answers. For instance, what is the precise role of the GATOR1 inhibitory mode? What is the nature of its interaction with GATOR2? Does it receive input from additional sensors? By what mechanisms does it contribute to human disease? These and other outstanding questions make this an exciting time to study GATOR1, and we look forward to new structural, biochemical, and biophysical insights to help fill the gap in our understanding of GATOR1 function.

Acknowledgments We thank all members of the Shen Laboratory for helpful insights. This work was supported by grants from the American Heart Association (928304) and the National Institute of General Medical Sciences of the National Institute of Health (R35 GM146824) to K.S.

References

- Ahmadian MR, Stege P, Scheffzek K, Wittinghofer A (1997) Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. *Nat Struct Biol* 4:686–689. <https://doi.org/10.1038/nsb0997-686>
- Back JW, de Jong L, Muijsers AO, de Koster CG (2003) Chemical cross-linking and mass spectrometry for protein structural modeling. *J Mol Biol* 331:303–313. [https://doi.org/10.1016/S0022-2836\(03\)00721-6](https://doi.org/10.1016/S0022-2836(03)00721-6)
- Bacq A, Roussel D, Bonduelle T, Zagaglia S, Maletic M, Ribierre T, Adle-Biasette H, Marchal C, Jennesson M, An I, Picard F, Navarro V, Sisodiya SM, Baulac S (2022) Cardiac investigations in sudden unexpected death in DEPDC5-related epilepsy. *Ann Neurol* 91:101–116. <https://doi.org/10.1002/ana.26256>
- Bagnall RD, Crompton DE, Petrovski S, Lam L, Cutmore C, Garry SI, Sadleir LG, Dibbens LM, Cairns A, Kivity S, Afawi Z, Regan BM, Duflou J, Berkovic SF, Scheffer IE, Semsarian C (2016) Exome-based analysis of cardiac arrhythmia, respiratory control, and epilepsy genes in sudden unexpected death in epilepsy. *Ann Neurol* 79:522–534. <https://doi.org/10.1002/ana.24596>
- Baldassari S, Licchetta L, Tinuper P, Bisulli F, Pippucci T (2016) GATOR1 complex: the common genetic actor in focal epilepsies. *J Med Genet* 53:503–510. <https://doi.org/10.1136/jmedgenet-2016-103883>
- Baldassari S, Picard F, Verbeek NE, van Kempen M, Brilstra EH, Lesca G, Conti V, Guerrini R, Bisulli F, Licchetta L, Pippucci T, Tinuper P, Hirsch E, de Saint Martin A, Chelly J, Rudolf G, Chipaux M, Ferrand-Sorbets S, Dorfmueller G, Sisodiya S, Balestrini S, Schoeler N, Hernandez-Hernandez L, Krithika S, Oegema R, Hagebeuk E, Gunning B, Deckers C, Berghuis B, Wegner I, Niks E, Jansen FE, Braun K, de Jong D, Rubboli G, Talvik I, Sander V, Uldall P, Jacquemont M-L, Nava C, Leguern E, Julia S, Gambardella A, d'Orsi G, Crichtiutti G, Faivre L, Darmency V, Benova B, Krsek P, Biraben A, Lebre A-S, Jennesson M, Sattar S, Marchal C, Nordli DR, Lindstrom K, Striano P, Lomax LB, Kiss C, Bartolomei F, Lepine AF, Schoonjans A-S, Stouffs K, Jansen A, Panagiotakaki E, Ricard-Mousnier B, Thevenon J, de Bellescize J, Catenioix H, Dorn T, Zenker M, Müller-Schlüter K, Brandt C, Krey I, Polster T, Wolff M, Balci M, Rostasy K, Achaz G, Zacher P, Becher T, Cloppenborg T, Yuskaitis CJ, Weckhuysen S, Poduri A, Lemke JR, Möller RS, Baulac S (2019) The landscape of epilepsy-related GATOR1 variants. *Genet Med* 21:398–408. <https://doi.org/10.1038/s41436-018-0060-2>

- Baple EL, Maroofian R, Chioza BA, Izadi M, Cross HE, Al-Turki S, Barwick K, Skrzypiec A, Pawlak R, Wagner K, Coblenz R, Zainy T, Patton MA, Mansour S, Rich P, Qualmann B, Hurles ME, Kessels MM, Crosby AH (2014) Mutations in KPTN cause macrocephaly, neurodevelopmental delay, and seizures. *Am J Hum Genet* 94:87–94. <https://doi.org/10.1016/j.ajhg.2013.10.001>
- Bar-Peled L, Schweitzer LD, Zoncu R, Sabatini DM (2012) Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell* 150:1196–1208. <https://doi.org/10.1016/j.cell.2012.07.032>
- Bar-Peled L, Chantranupong L, Cherniack AD, Chen WW, Ottina KA, Grabiner BC, Spear ED, Carter SL, Meyerson M, Sabatini DM (2013) A tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. *Science* 340:1100–1106. <https://doi.org/10.1126/science.1232044>
- Basel-Vanagaite L, Hershkovitz T, Heyman E, Raspall-Chaure M, Kakar N, Smirin-Yosef P, Vila-Pueyo M, Kornreich L, Thiele H, Bode H, Lagovsky I, Dahary D, Haviv A, Hubshman MW, Pasmanik-Chor M, Nürnberg P, Gothelf D, Kubisch C, Shohat M, Macaya A, Borck G (2013) Biallelic SZT2 mutations cause infantile encephalopathy with epilepsy and dysmorphic corpus callosum. *Am J Hum Genet* 93:524–529. <https://doi.org/10.1016/j.ajhg.2013.07.005>
- Baulac S (2016) Chapter 3 - mTOR signaling pathway genes in focal epilepsies. In: Rossignol E, Carmant L, Lacaille J-C (eds) *Progress in brain research, neurobiology of epilepsy*. Elsevier, pp 61–79. <https://doi.org/10.1016/bs.pbr.2016.04.013>
- Brown EJ, Albers MW, Bum Shin T, Ichikawa K, Keith CT, Lane WS, Schreiber SL (1994) A mammalian protein targeted by G1-arresting rapamycin–receptor complex. *Nature* 369:756–758. <https://doi.org/10.1038/369756a0>
- Burchett SA (2000) Regulators of G protein signaling: a bestiary of modular protein binding domains. *J Neurochem* 74:1335–1351. <https://doi.org/10.1046/j.1471-4159.2000.0751335.x>
- Carroll B, Nelson G, Rabanal-Ruiz Y, Kucheryavenko O, Dunhill-Turner NA, Chesterman CC, Zahari Q, Zhang T, Conduit SE, Mitchell CA, Maddocks ODK, Lovat P, von Zglinicki T, Korolchuk VI (2017) Persistent mTORC1 signaling in cell senescence results from defects in amino acid and growth factor sensing. *J Cell Biol* 216:1949–1957. <https://doi.org/10.1083/jcb.201610113>
- Chantranupong L, Wolfson RL, Orozco JM, Saxton RA, Scaria SM, Bar-Peled L, Spooner E, Isasa M, Gygi SP, Sabatini DM (2014) The Sestrins interact with GATOR2 to negatively regulate the amino-acid-sensing pathway upstream of mTORC1. *Cell Rep* 9:1–8. <https://doi.org/10.1016/j.celrep.2014.09.014>
- Chantranupong L, Scaria SM, Saxton RA, Gygi MP, Shen K, Wyant GA, Wang T, Harper JW, Gygi SP, Sabatini DM (2016) The CASTOR proteins are arginine sensors for the mTORC1 pathway. *Cell* 165:153–164. <https://doi.org/10.1016/j.cell.2016.02.035>
- Chen Z, Jiang Q, Zhu P, Chen Y, Xie X, Du Z, Jiang L, Tang W (2019) NPRL2 enhances autophagy and the resistance to Everolimus in castration-resistant prostate cancer. *Prostate* 79:44–53. <https://doi.org/10.1002/pros.23709>
- Chichili VPR, Kumar V, Sivaraman J (2013) A method to trap transient and weak interacting protein complexes for structural studies. *Intrinsic Disord Proteins* 1:e25464. <https://doi.org/10.4161/idp.25464>
- Coleman DE, Berghuis AM, Lee E, Linder ME, Gilman AG, Sprang SR (1994) Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. *Science* 265:1405–1412. <https://doi.org/10.1126/science.8073283>
- Condon KJ, Sabatini DM (2019) Nutrient regulation of mTORC1 at a glance. *J Cell Sci* 132:jcs222570. <https://doi.org/10.1242/jcs.222570>
- Crino PB (2015) mTOR signaling in epilepsy: insights from malformations of cortical development. *Cold Spring Harb Perspect Med* 5:a022442. <https://doi.org/10.1101/cshperspect.a022442>
- Daumke O, Weyand M, Chakrabarti PP, Vetter IR, Wittinghofer A (2004) The GTPase-activating protein Rap1GAP uses a catalytic asparagine. *Nature* 429:197–201. <https://doi.org/10.1038/nature02505>

- De Franceschi N, Wild K, Schlacht A, Dacks JB, Sinning I, Filippini F (2014) Longin and GAF domains: structural evolution and adaptation to the subcellular trafficking machinery. *Traffic* 15:104–121. <https://doi.org/10.1111/tra.12124>
- Devinsky O, Hesdorffer DC, Thurman DJ, Lhatoo S, Richerson G (2016) Sudden unexpected death in epilepsy: epidemiology, mechanisms, and prevention. *Lancet Neurol* 15:1075–1088. [https://doi.org/10.1016/S1474-4422\(16\)30158-2](https://doi.org/10.1016/S1474-4422(16)30158-2)
- Dibbens LM, de Vries B, Donatello S, Heron SE, Hodgson BL, Chintawar S, Crompton DE, Hughes JN, Bellows ST, Klein KM, Callenbach PMC, Corbett MA, Gardner AE, Kivity S, Iona X, Regan BM, Weller CM, Crimmins D, O'Brien TJ, Guerrero-López R, Mulley JC, Dubeau F, Licchetta L, Bisulli F, Cossette P, Thomas PQ, Gecz J, Serratosa J, Brouwer OF, Andermann F, Andermann E, van den Maagdenberg AMJM, Pandolfo M, Berkovic SF, Scheffer IE (2013) Mutations in DEPDC5 cause familial focal epilepsy with variable foci. *Nat Genet* 45:546–551. <https://doi.org/10.1038/ng.2599>
- Dibble CC, Manning BD (2013) Signal integration by mTORC1 coordinates nutrient input with biosynthetic output. *Nat Cell Biol* 15:555–564. <https://doi.org/10.1038/ncb2763>
- Dokudovskaya S, Waharte F, Schlessinger A, Pieper U, Devos DP, Cristea IM, Williams R, Salamero J, Chait BT, Sali A, Field MC, Rout MP, Dargemont C (2011) A conserved coatomer-related complex containing sec13 and seh1 dynamically associates with the vacuole in *Saccharomyces cerevisiae*. *Mol Cell Proteomics* 10:M110.006478. <https://doi.org/10.1074/mcp.M110.006478>
- Doxsey DD, Tettoni SD, Egri SB, Shen K (2023) Redundant electrostatic interactions between GATOR1 and the Rag GTPase heterodimer drive efficient amino acid sensing in human cells. *J Biol Chem* 299:104880. <https://doi.org/10.1016/j.jbc.2023.104880>
- Efeyan A, Zoncu R, Chang S, Gumper I, Snitkin H, Wolfson RL, Kirak O, Sabatini DD, Sabatini DM (2013) Rag GTPase-mediated regulation of mTORC1 by nutrients is necessary for neonatal autophagy and survival. *Nature* 493:679–683. <https://doi.org/10.1038/nature11745>
- Egri SB, Ouch C, Chou H-T, Yu Z, Song K, Xu C, Shen K (2022) Cryo-EM structures of the human GATOR1-Rag-Ragulator complex reveal a spatial-constraint regulated GAP mechanism. *Mol Cell* 82:1836–1849.e5. <https://doi.org/10.1016/j.molcel.2022.03.002>
- Faingold CL (2012) Brainstem networks: reticulo-cortical synchronization in generalized convulsive seizures. In: Noebels JL, Avoli M, Rogawski MA, Olsen RW, Delgado-Escueta AV (eds) *Jasper's basic mechanisms of the epilepsies*. National Center for Biotechnology Information (US), Bethesda, MD
- Gao Y, Wang J, Fan G (2013) NPRL2 is an independent prognostic factor of osteosarcoma. *Cancer Biomark* 12:31–36. <https://doi.org/10.3233/CBM-120290>
- Gong R, Li L, Liu Y, Wang P, Yang H, Wang L, Cheng J, Guan K-L, Xu Y (2011) Crystal structure of the Gtr1p–Gtr2p complex reveals new insights into the amino acid-induced TORC1 activation. *Genes Dev* 25:1668–1673. <https://doi.org/10.1101/gad.16968011>
- Grant GA (2006) The ACT domain: a small molecule binding domain and its role as a common regulatory element*. *J Biol Chem* 281:33825–33829. <https://doi.org/10.1074/jbc.R600024200>
- Gu X, Orozco JM, Saxton RA, Condon KJ, Liu GY, Krawczyk PA, Scaria SM, Harper JW, Gygi SP, Sabatini DM (2017) SAMTOR is an S-adenosylmethionine sensor for the mTORC1 pathway. *Science* 358:813–818. <https://doi.org/10.1126/science.aao3265>
- Hassan AM, Keene DL, Whiting SE, Jacob PJ, Champagne JR, Humphreys P (1999) Ketogenic diet in the treatment of refractory epilepsy in childhood. *Pediatr Neurol* 21:548–552. [https://doi.org/10.1016/S0887-8994\(99\)00045-4](https://doi.org/10.1016/S0887-8994(99)00045-4)
- Iffland PH, Carson V, Bordey A, Crino PB (2019) GATORopathies: the role of amino acid regulatory gene mutations in epilepsy and cortical malformations. *Epilepsia* 60:2163–2173. <https://doi.org/10.1111/epi.16370>
- Ingelman M, Bianchi V, Eklund H (1997) The three-dimensional structure of flavodoxin reductase from *Escherichia coli* at 1.7 Å resolution. *J Mol Biol* 268:147–157. <https://doi.org/10.1006/jmbi.1997.0957>
- Inoki K, Li Y, Xu T, Guan K-L (2003) Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. *Genes Dev* 17:1829–1834. <https://doi.org/10.1101/gad.1110003>

- Insel TR (2014) Brain somatic mutations: the dark matter of psychiatric genetics? *Mol Psychiatry* 19:156–158. <https://doi.org/10.1038/mp.2013.168>
- Jayachandran G, Ueda K, Wang B, Roth JA, Ji L (2010) NPRL2 sensitizes human non-small cell lung cancer (NSCLC) cells to cisplatin treatment by regulating key components in the DNA repair pathway. *PLoS One* 5:e11994. <https://doi.org/10.1371/journal.pone.0011994>
- Jeong J-H, Lee K-H, Kim Y-M, Kim D-H, Oh B-H, Kim Y-G (2012) Crystal structure of the Gtr1p(GTP)-Gtr2p(GDP) protein complex reveals large structural rearrangements triggered by GTP-to-GDP conversion. *J Biol Chem* 287:29648–29653. <https://doi.org/10.1074/jbc.C112.384420>
- Jiang C, Dai X, He S, Zhou H, Fang L, Guo J, Liu S, Zhang T, Pan W, Yu H, Fu T, Li D, Inuzuka H, Wang P, Xiao J, Wei W (2023) Ring domains are essential for GATOR2-dependent mTORC1 activation. *Mol Cell* 83:74–89.e9. <https://doi.org/10.1016/j.molcel.2022.11.021>
- Jung J, Genau HM, Behrends C (2015) Amino acid-dependent mTORC1 regulation by the lysosomal membrane protein SLC38A9. *Mol Cell Biol* 35:2479–2494. <https://doi.org/10.1128/MCB.00125-15>
- Kim E, Goraksha-Hicks P, Li L, Neufeld TP, Guan K-L (2008) Regulation of TORC1 by Rag GTPases in nutrient response. *Nat Cell Biol* 10:935–945. <https://doi.org/10.1038/ncb1753>
- Kim JS, Ro S-H, Kim M, Park H-W, Semple IA, Park H, Cho U-S, Wang W, Guan K-L, Karin M, Lee JH (2015) Sestrin2 inhibits mTORC1 through modulation of GATOR complexes. *Sci Rep* 5:9502. <https://doi.org/10.1038/srep09502>
- Kinsman SL, Vining EPG, Quaskey SA, Mellits D, Freeman JM (1992) Efficacy of the ketogenic diet for intractable seizure disorders: review of 58 cases. *Epilepsia* 33:1132–1136. <https://doi.org/10.1111/j.1528-1157.1992.tb01770.x>
- Lang EJ, Cross PJ, Mittelstädt G, Jameson GB, Parker EJ (2014) Allosteric ACTion: the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr Opin Struct Biol* 29:102–111. <https://doi.org/10.1016/j.sbi.2014.10.007>
- Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11:204–220. <https://doi.org/10.1038/nrg2719>
- Lawrence RE, Fromm SA, Fu Y, Yokom AL, Kim DJ, Thelen AM, Young LN, Lim C-Y, Samelson AJ, Hurlley JH, Zoncu R (2019) Structural mechanism of a Rag GTPase activation checkpoint by the lysosomal folliculin complex. *Science* 366:971–977. <https://doi.org/10.1126/science.aax0364>
- Lee PL, Tang Y, Li H, Guertin DA (2016) Raptor/mTORC1 loss in adipocytes causes progressive lipodystrophy and fatty liver disease. *Mol Metab* 5:422–432. <https://doi.org/10.1016/j.molmet.2016.04.001>
- Lerman, M.I., Minna, J.D., for The International Lung Cancer Chromosome 3p21.3 Tumor Suppressor Gene Consortium (2000) The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: identification and evaluation of the resident candidate tumor suppressor genes1. *Cancer Res* 60:6116–6133
- Levine TP, Daniels RD, Wong LH, Gatta AT, Gerondopoulos A, Barr FA (2013) Discovery of new Longin and Roadblock domains that form platforms for small GTPases in Ragulator and TRAPP-II. *Small GTPases* 4:62–69. <https://doi.org/10.4161/sgtp.24262>
- Lipton JO, Sahin M (2014) The neurology of mTOR. *Neuron* 84:275–291. <https://doi.org/10.1016/j.neuron.2014.09.034>
- Liu A, Liu D, Du Y, Pei F, Yang G, Liu B, Zhang H, Wang X, Fan Y, Chen Y, Jiang Y, Chen J (2014) Relationship between tumor and peripheral blood NPRL2 mRNA levels in patients with colorectal adenoma and colorectal cancer. *Cancer Biol Ther* 15:489–495. <https://doi.org/10.4161/cbt.28016>
- Liu J, Geng R, Yang S, Shao F, Zhong Z, Yang M, Ni S, Cai L, Bai J (2021) Development and clinical validation of novel 8-gene prognostic signature associated with the proportion of regulatory T cells by weighted gene co-expression network analysis in uterine corpus endometrial carcinoma. *Front Immunol* 12:788431. <https://doi.org/10.3389/fimmu.2021.788431>

- Long X, Lin Y, Ortiz-Vega S, Yonezawa K, Avruch J (2005) Rheb binds and regulates the mTOR kinase. *Curr Biol* 15:702–713. <https://doi.org/10.1016/j.cub.2005.02.053>
- Ma XM, Blenis J (2009) Molecular mechanisms of mTOR-mediated translational control. *Nat Rev Mol Cell Biol* 10:307–318. <https://doi.org/10.1038/nrm2672>
- Mannick JB, Morris M, Hockey H-UP, Roma G, Beibel M, Kulmatycki K, Watkins M, Shavlakadze T, Zhou W, Quinn D, Glass DJ, Klickstein LB (2018) TORC1 inhibition enhances immune function and reduces infections in the elderly. *Sci Transl Med* 10:eaq1564. <https://doi.org/10.1126/scitranslmed.aq1564>
- Mattson G, Conklin E, Desai S, Niellander G, Savage MD, Morgensen S (1993) A practical approach to crosslinking. *Mol Biol Rep* 17:167–183. <https://doi.org/10.1007/BF00986726>
- Maydell BV, Wyllie E, Akhtar N, Kotagal P, Powaski K, Cook K, Weinstock A, Rothner AD (2001) Efficacy of the ketogenic diet in focal versus generalized seizures. *Pediatr Neurol* 25:208–212. [https://doi.org/10.1016/S0887-8994\(01\)00310-1](https://doi.org/10.1016/S0887-8994(01)00310-1)
- McDaniel SS, Rensing NR, Thio LL, Yamada KA, Wong M (2011) The ketogenic diet inhibits the mammalian target of rapamycin (mTOR) pathway. *Epilepsia* 52:e7–e11. <https://doi.org/10.1111/j.1528-1167.2011.02981.x>
- Menon S, Manning BD (2008) Common corruption of the mTOR signaling network in human tumors. *Oncogene* 27:S43–S51. <https://doi.org/10.1038/onc.2009.352>
- Mentch SJ, Mehrmohamadi M, Huang L, Liu X, Gupta D, Mattocks D, Gómez Padilla P, Ables G, Bamman MM, Thalacker-Mercer AE, Nichenametla SN, Locasale JW (2015) Histone methylation dynamics and gene regulation occur through the sensing of one-carbon metabolism. *Cell Metab* 22:861–873. <https://doi.org/10.1016/j.cmet.2015.08.024>
- Miertzschke M, Koerner C, Vetter IR, Keilberg D, Hot E, Leonardy S, Søgaaard-Andersen L, Wittinghofer A (2011) Structural analysis of the Ras-like G protein MglA and its cognate GAP MglB and implications for bacterial polarity. *EMBO J* 30:4185–4197. <https://doi.org/10.1038/emboj.2011.291>
- Moloney PB, Cavalleri GL, Delanty N (2021) Epilepsy in the mTORopathies: opportunities for precision medicine. *Brain Commun* 3:fcab222. <https://doi.org/10.1093/braincomms/fcab222>
- Nagy GN, Suardi  az R, Lopata A, Ozohanics O, V  key K, Brooks BR, Leveles I, T  th J, V  rtessy BG, Rosta E (2016) Structural characterization of arginine fingers: identification of an arginine finger for the pyrophosphatase dUTPases. *J Am Chem Soc* 138:15035–15045. <https://doi.org/10.1021/jacs.6b09012>
- Nakashima N, Noguchi E, Nishimoto T (1999) *Saccharomyces cerevisiae* putative G protein, Gtr1p, which forms complexes with itself and a novel protein designated as Gtr2p, negatively regulates the Ran/Gsp1p G protein cycle through Gtr2p. *Genetics* 152:853–867
- Neklesa TK, Davis RW (2009) A genome-wide screen for regulators of TORC1 in response to amino acid starvation reveals a conserved Npr2/3 complex. *PLoS Genet* 5:e1000515. <https://doi.org/10.1371/journal.pgen.1000515>
- N  rpel J, Cavadini S, Schenk AD, Graff-Meyer A, Hess D, Seebacher J, Chao JA, Bhaskar V (2021) Structure of the human C9orf72-SMCR8 complex reveals a multivalent protein interaction architecture. *PLoS Biol* 19:e3001344. <https://doi.org/10.1371/journal.pbio.3001344>
- Otani S, Takeda S, Yamada S, Sakakima Y, Sugimoto H, Nomoto S, Kasuya H, Kanazumi N, Nagasaka T, Nakao A (2009) The tumor suppressor NPRL2 in hepatocellular carcinoma plays an important role in progression and can be served as an independent prognostic factor. *J Surg Oncol* 100:358–363. <https://doi.org/10.1002/jso.21241>
- Pajusalu S, Reimand T,   unap K (2015) Novel homozygous mutation in KPTN gene causing a familial intellectual disability-macrocephaly syndrome. *Am J Med Genet A* 167:1913–1915. <https://doi.org/10.1002/ajmg.a.37105>
- Panchaud N, P  li-Gulli M-P, De Virgilio C (2013a) Amino acid deprivation inhibits TORC1 through a GTPase-activating protein complex for the Rag family GTPase Gtr1. *Sci Signal* 6:ra42. <https://doi.org/10.1126/scisignal.2004112>
- Panchaud N, P  li-Gulli M-P, De Virgilio C (2013b) SEACing the GAP that nEGOCiates TORC1 activation. *Cell Cycle* 12:2948–2952. <https://doi.org/10.4161/cc.26000>

- Pang Y, Xie F, Cao H, Wang C, Zhu M, Liu X, Lu X, Huang T, Shen Y, Li K, Jia X, Li Z, Zheng X, Wang S, He Y, Wang L, Fletcher JA, Wang Y (2019) Mutational inactivation of mTORC1 repressor gene DEPDC5 in human gastrointestinal stromal tumors. *Proc Natl Acad Sci U S A* 116:22746–22753. <https://doi.org/10.1073/pnas.1914542116>
- Parmigiani A, Nourbakhsh A, Ding B, Wang W, Kim YC, Akopiants K, Guan K-L, Karin M, Budanov AV (2014) Sestrins inhibit mTORC1 kinase activation through the GATOR complex. *Cell Rep* 9:1281–1291. <https://doi.org/10.1016/j.celrep.2014.10.019>
- Peng M, Yin N, Li MO (2017) SZT2 dictates GATOR control of mTORC1 signaling. *Nature* 543:433–437. <https://doi.org/10.1038/nature21378>
- Petit CS, Roczniaik-Ferguson A, Ferguson SM (2013) Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. *J Cell Biol* 202:1107–1122. <https://doi.org/10.1083/jcb.201307084>
- Privé GG, Milburn MV, Tong L, de Vos AM, Yamaizumi Z, Nishimura S, Kim SH (1992) X-ray crystal structures of transforming p21 ras mutants suggest a transition-state stabilization mechanism for GTP hydrolysis. *Proc Natl Acad Sci U S A* 89:3649–3653
- Qu A, Leahy DJ (1995) Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin. *Proc Natl Acad Sci U S A* 92:10277–10281. <https://doi.org/10.1073/pnas.92.22.10277>
- Rabinowitz JD, White E (2010) Autophagy and metabolism. *Science* 330:1344–1348. <https://doi.org/10.1126/science.1193497>
- Rho JM (2017) How does the ketogenic diet induce anti-seizure effects? *Neurosci Lett. Metabolism and Neural Function* - Charles Roe (Honor Special Issue) 637:4–10. doi:<https://doi.org/10.1016/j.neulet.2015.07.034>
- Ribierre T, Deleuze C, Bacq A, Baldassari S, Marsan E, Chipaux M, Muraca G, Roussel D, Navarro V, Leguern E, Miles R, Baulac S (2018) Second-hit mosaic mutation in mTORC1 repressor DEPDC5 causes focal cortical dysplasia-associated epilepsy. *J Clin Invest* 128:2452–2458. <https://doi.org/10.1172/JCI99384>
- Ricos MG, Hodgson BL, Pippucci T, Saidin A, Ong YS, Heron SE, Licchetta L, Bisulli F, Bayly MA, Hughes J, Baldassari S, Palombo F, Epilepsy Electroclinical Study Group, Santucci M, Meletti S, Berkovic SF, Rubboli G, Thomas PQ, Scheffer IE, Tinuper P, Geoghegan J, Schreiber AW, Dibbens LM (2016) Mutations in the mammalian target of rapamycin pathway regulators NPRL2 and NPRL3 cause focal epilepsy. *Ann Neurol* 79:120–131. <https://doi.org/10.1002/ana.24547>
- Sabatini DM, Erdjument-Bromage H, Lui M, Tempst P, Snyder SH (1994) RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. *Cell* 78:35–43. [https://doi.org/10.1016/0092-8674\(94\)90570-3](https://doi.org/10.1016/0092-8674(94)90570-3)
- Sabers CJ, Martin MM, Brunn GJ, Williams JM, Dumont FJ, Wiederrecht G, Abraham RT (1995) Isolation of a protein target of the FKBP12-rapamycin complex in mammalian cells (*). *J Biol Chem* 270:815–822. <https://doi.org/10.1074/jbc.270.2.815>
- Sancak Y, Peterson TR, Shaul YD, Lindquist RA, Thoreen CC, Bar-Peled L, Sabatini DM (2008) The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* 320:1496–1501. <https://doi.org/10.1126/science.1157535>
- Sancak Y, Bar-Peled L, Zoncu R, Markhard AL, Nada S, Sabatini DM (2010) Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* 141:290–303. <https://doi.org/10.1016/j.cell.2010.02.024>
- Saxton RA, Chantranupong L, Knockenbauer KE, Schwartz TU, Sabatini DM (2016a) Mechanism of arginine sensing by CASTOR1 upstream of mTORC1. *Nature* 536:229–233. <https://doi.org/10.1038/nature19079>
- Saxton RA, Knockenbauer KE, Wolfson RL, Chantranupong L, Pacold ME, Wang T, Schwartz TU, Sabatini DM (2016b) Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway. *Science* 351:53–58. <https://doi.org/10.1126/science.aad2087>
- Scheffzek K, Ahmadian MR, Kabsch W, Wiesmüller L, Lautwein A, Schmitz F, Wittinghofer A (1997) The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* 277:333–338. <https://doi.org/10.1126/science.277.5324.333>

- Schlenker O, Hendricks A, Sinning I, Wild K (2006) The structure of the mammalian signal recognition particle (SRP) receptor as prototype for the interaction of small GTPases with Longin domains. *J Biol Chem* 281:8898–8906. <https://doi.org/10.1074/jbc.M512415200>
- Schröter G, Mann D, Kötting C, Gerwert K (2015) Integration of Fourier transform infrared spectroscopy, fluorescence spectroscopy, steady-state kinetics and molecular dynamics simulations of Gαi1 distinguishes between the GTP hydrolysis and GDP release mechanism*. *J Biol Chem* 290:17085–17095. <https://doi.org/10.1074/jbc.M115.651190>
- Schürmann A, Brauers A, Maßmann S, Becker W, Joost H-G (1995) Cloning of a novel family of mammalian GTP-binding proteins (RagA, RagBs, RagB1) with remote similarity to the Ras-related GTPases*. *J Biol Chem* 270:28982–28988. <https://doi.org/10.1074/jbc.270.48.28982>
- Sekiguchi T, Hirose E, Nakashima N, Ii M, Nishimoto T (2001) Novel G proteins, Rag C and Rag D, interact with GTP-binding proteins, Rag A and Rag B*. *J Biol Chem* 276:7246–7257. <https://doi.org/10.1074/jbc.M004389200>
- Sengupta S, Peterson TR, Laplante M, Oh S, Sabatini DM (2010) mTORC1 controls fasting-induced ketogenesis and its modulation by ageing. *Nature* 468:1100–1104. <https://doi.org/10.1038/nature09584>
- Shen K, Sabatini DM (2018) Ragulator and SLC38A9 activate the Rag GTPases through noncanonical GEF mechanisms. *Proc Natl Acad Sci U S A* 115:9545–9550. <https://doi.org/10.1073/pnas.1811727115>
- Shen K, Choe A, Sabatini DM (2017) Intersubunit crosstalk in the Rag GTPase heterodimer enables mTORC1 to respond rapidly to amino acid availability. *Mol Cell* 68:552–565.e8. <https://doi.org/10.1016/j.molcel.2017.09.026>
- Shen K, Huang RK, Brignole EJ, Condon KJ, Valenstein ML, Chantranupong L, Bomaliyamu A, Choe A, Hong C, Yu Z, Sabatini DM (2018) Architecture of the human GATOR1 and GATOR1-Rag complexes. *Nature* 556:64–69. <https://doi.org/10.1038/nature26158>
- Shen K, Rogala KB, Chou H-T, Huang RK, Yu Z, Sabatini DM (2019a) Cryo-EM structure of the human FLCN-FNIP2-Rag-Ragulator complex. *Cell* 179:1319–1329.e8. <https://doi.org/10.1016/j.cell.2019.10.036>
- Shen K, Valenstein ML, Gu X, Sabatini DM (2019b) Arg-78 of Nprl2 catalyzes GATOR1-stimulated GTP hydrolysis by the Rag GTPases. *J Biol Chem* 294:2970–2975. <https://doi.org/10.1074/jbc.AC119.007382>
- Shiozawa K, Maita N, Tomii K, Seto A, Goda N, Akiyama Y, Shimizu T, Shirakawa M, Hiroaki H (2004) Structure of the N-terminal domain of PEX1 AAA-ATPase: characterization of a putative adaptor-binding domain*. *J Biol Chem* 279:50060–50068. <https://doi.org/10.1074/jbc.M407837200>
- Sim JC, Scerri T, Fanjul-Fernández M, Riseley JR, Gillies G, Pope K, van Roozendaal H, Heng JJ, Mandelstam SA, McGillivray G, MacGregor D, Kannan L, Maixner W, Harvey AS, Amor DJ, Delatycki MB, Crino PB, Bahlo M, Lockhart PJ, Leventer RJ (2016) Familial cortical dysplasia caused by mutation in the mammalian target of rapamycin regulator NPRL3. *Ann Neurol* 79:132–137. <https://doi.org/10.1002/ana.24502>
- Solanki S, Sanchez K, Ponnusamy V, Kota V, Bell HN, Cho C-S, Kowalsky AH, Green M, Lee JH, Shah YM (2023) Dysregulated amino acid sensing drives colorectal cancer growth and metabolic reprogramming leading to chemoresistance. *Gastroenterology* 164:376–391.e13. <https://doi.org/10.1053/j.gastro.2022.11.014>
- Tafur L, Hinterdorfer K, Gabus C, Lamanna C, Bergmann A, Sadian Y, Hamdi F, Kyriakis FL, Kastiris PL, Loewith R (2022) Cryo-EM structure of the SEA complex. *Nature* 611:399–404. <https://doi.org/10.1038/s41586-022-05370-0>
- Takei N, Inamura N, Kawamura M, Namba H, Hara K, Yonezawa K, Nawa H (2004) Brain-derived neurotrophic factor induces mammalian target of rapamycin-dependent local activation of translation machinery and protein synthesis in neuronal dendrites. *J Neurosci* 24:9760–9769. <https://doi.org/10.1523/JNEUROSCI.1427-04.2004>
- Tang X, Zhang Y, Wang G, Zhang C, Wang F, Shi J, Zhang T, Ding J (2022) Molecular mechanism of S-adenosylmethionine sensing by SAMTOR in mTORC1 signaling. *Sci Adv* 8:eabn3868. <https://doi.org/10.1126/sciadv.abn3868>

- Tochio H, Tsui MM, Banfield DK, Zhang M (2001) An autoinhibitory mechanism for nonsyntaxin SNARE proteins revealed by the structure of Ykt6p. *Science* 293:698–702. <https://doi.org/10.1126/science.1062950>
- Tsun Z-Y, Bar-Peled L, Chantranupong L, Zoncu R, Wang T, Kim C, Spooner E, Sabatini DM (2013) The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol Cell* 52:495–505. <https://doi.org/10.1016/j.molcel.2013.09.016>
- Valenstein ML, Rogala KB, Lalgudi PV, Brignole EJ, Gu X, Saxton RA, Chantranupong L, Kolibius J, Quast J-P, Sabatini DM (2022) Structure of the nutrient-sensing hub GATOR2. *Nature* 607:610–616. <https://doi.org/10.1038/s41586-022-04939-z>
- Venkatesan C, Angle B, Millichap JJ (2016) Early-life epileptic encephalopathy secondary to *SZT2* pathogenic recessive variants. *Epileptic Disord* 18:195–200. <https://doi.org/10.1684/epd.2016.0828>
- Wang S, Tsun Z-Y, Wolfson R, Shen K, Wyant GA, Plovianich ME, Yuan ED, Jones TD, Chantranupong L, Comb W, Wang T, Bar-Peled L, Zoncu R, Straub C, Kim C, Park J, Sabatini BL, Sabatini DM (2015) The amino acid transporter SLC38A9 is a key component of a lysosomal membrane complex that signals arginine sufficiency to mTORC1. *Science* 347:188–194. <https://doi.org/10.1126/science.1257132>
- Weckhuysen S, Marsan E, Lambrecq V, Marchal C, Morin-Brureau M, An-Gourfinkel I, Baulac M, Fohlen M, Kallay Zetchi C, Seeck M, de la Grange P, Dermaut B, Meurs A, Thomas P, Chassoux F, Leguern E, Picard F, Baulac S (2016) Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. *Epilepsia* 57:994–1003. <https://doi.org/10.1111/epi.13391>
- Wei Y, Reveal B, Reich J, Laursen WJ, Senger S, Akbar T, Iida-Jones T, Cai W, Jarnik M, Lilly MA (2014) TORC1 regulators Iml1/GATOR1 and GATOR2 control meiotic entry and oocyte development in *Drosophila*. *Proc Natl Acad Sci U S A* 111:E5670–E5677. <https://doi.org/10.1073/pnas.1419156112>
- Wolfson RL, Chantranupong L, Saxton RA, Shen K, Scaria SM, Cantor JR, Sabatini DM (2016) Sestrin2 is a leucine sensor for the mTORC1 pathway. *Science* 351:43–48. <https://doi.org/10.1126/science.aab2674>
- Wolfson RL, Chantranupong L, Wyant GA, Gu X, Orozco JM, Shen K, Condon KJ, Petri S, Kedir J, Scaria SM, Abu-Remaileh M, Frankel WN, Sabatini DM (2017) KICSTOR recruits GATOR1 to the lysosome and is necessary for nutrients to regulate mTORC1. *Nature* 543:438–442. <https://doi.org/10.1038/nature21423>
- Wong HC, Mao J, Nguyen JT, Srinivas S, Zhang W, Liu B, Li L, Wu D, Zheng J (2000) Structural basis of the recognition of the Dishevelled DEP domain in the Wnt signaling pathway. *Nat Struct Biol* 7:1178–1184. <https://doi.org/10.1038/82047>
- Wu X, Tu BP (2011) Selective regulation of autophagy by the Iml1-Npr2-Npr3 complex in the absence of nitrogen starvation. *Mol Biol Cell* 22:4124–4133. <https://doi.org/10.1091/mbc.E11-06-0525>
- Yao Y, Jones E, Inoki K (2017) Lysosomal regulation of mTORC1 by amino acids in mammalian cells. *Biomol Ther* 7:51. <https://doi.org/10.3390/biom7030051>
- Ye C, Sutter BM, Wang Y, Kuang Z, Tu BP (2017) A metabolic function for phospholipid and histone methylation. *Mol Cell* 66:180–193.e8. <https://doi.org/10.1016/j.molcel.2017.02.026>
- Zhang Y, Gao X, Saucedo LJ, Ru B, Edgar BA, Pan D (2003) Rheb is a direct target of the tuberous sclerosis tumour suppressor proteins. *Nat Cell Biol* 5:578–581. <https://doi.org/10.1038/ncb999>
- Zhang T, Wang R, Wang Z, Wang X, Wang F, Ding J (2017) Structural basis for Ragulator functioning as a scaffold in membrane-anchoring of Rag GTPases and mTORC1. *Nat Commun* 8:1394. <https://doi.org/10.1038/s41467-017-01567-4>

Chapter 13

Mitochondrial Alpha-Keto Acid Dehydrogenase Complexes: Recent Developments on Structure and Function in Health and Disease



Eszter Szabo, Balint Nagy, Andras Czajlik, Timea Komlodi,
Oliver Ozohanics, Laszlo Tretter, and Attila Ambrus

Abstract The present work delves into the enigmatic world of mitochondrial alpha-keto acid dehydrogenase complexes discussing their metabolic significance, enzymatic operation, moonlighting activities, and pathological relevance with links to underlying structural features. This ubiquitous family of related but diverse multienzyme complexes is involved in carbohydrate metabolism (pyruvate dehydrogenase complex), the citric acid cycle (α -ketoglutarate dehydrogenase complex), and amino acid catabolism (branched-chain α -keto acid dehydrogenase complex, α -ketoadipate dehydrogenase complex); the complexes all function at strategic points and also participate in regulation in these metabolic pathways. These systems are among the largest multienzyme complexes with at times more than 100 protein chains and weights ranging up to ~10 million Daltons. Our chapter offers a wealth of up-to-date information on these multienzyme complexes for a comprehensive understanding of their significance in health and disease.

Keywords α -keto acid dehydrogenase complexes · Mitochondrion · ROS · Metabolic disease · Pathogenic mutation · Pyruvate dehydrogenase complex · Alpha-ketoglutarate dehydrogenase complex · Alpha-ketoadipate dehydrogenase complex · Branched-chain α -keto acid dehydrogenase complex

Eszter Szabo, Balint Nagy and Andras Czajlik contributed equally to this work.

E. Szabo · B. Nagy · A. Czajlik · T. Komlodi · O. Ozohanics · L. Tretter · A. Ambrus (✉)
Department of Biochemistry, Semmelweis University, Budapest, Hungary
e-mail: szabo.eszter1@semmelweis.hu; nagy.balint@semmelweis.hu; czajlik.andras@semmelweis.hu; komlodi.timea@semmelweis.hu; ozohanics.oliver@semmelweis.hu; tretter.laszlo@semmelweis.hu; ambrus.attila@semmelweis.hu

Abbreviations

KGDHc	Alpha-ketoglutarate (also known as 2-oxoglutarate) dehydrogenase complex
PDHc	Pyruvate dehydrogenase complex
BCKDHc	Branched-chain α -keto acid dehydrogenase complex
KADHc	Alpha-ketoadipate dehydrogenase complex
CoA	Coenzyme A
E1k/p/a/b	α -ketoglutarate/pyruvate/ α -ketoadipate/branched-chain α -keto acid dehydrogenase
E2k/p/b	[dihydro]lipoamide/lipoyl succinyl/acetyl/acyltransferase (or transsuccinylase/acetylase/acylase)
E3 or LADH	[dihydro]lipoamide/lipoyl dehydrogenase
DHTKD1	Dehydrogenase E1 and transketolase domain-containing 1 (the E1a subunit)
h	Human origin
FAD	Flavin adenine dinucleotide
NAD ⁺ /NADH	Nicotinamide adenine dinucleotide (oxidized/reduced forms)
ThDP	Thiamine di(/pyro)phosphate
ec	<i>E. coli</i> origin
LD	Lipoyl domain
E3BP	E3-binding protein (Protein X)
SBD/PSBD	Peripheral subunit-binding domain (in E2)
CD	Catalytic (core) domain (in E2)
cryo-EM	Cryogenic electron microscopy
BDK	BCKDH kinase
BDP	BCKDH phosphatase
PDK	PDH kinase
PDP	PDH phosphatase
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
HDX-MS	Hydrogen-deuterium exchange mass spectrometry
ROS	Reactive oxygen species
<i>DLST</i>	Gene coding for human dihydrolipoamide succinyltransferase
<i>DLD</i>	Gene coding for human dihydrolipoamide dehydrogenase
KG	α -ketoglutarate
KA	α -ketoadipate
pyr	Pyruvate
LA/DHLA	Lipoic acid/dihydrolipoic acid
LAM/DHLAM	Lipoamide/dihydrolipoamide
KGDHL/OGDHL	KGDH-like (E1k-like) subunit

Metabolic Roles of the Mitochondrial α -Keto Acid Dehydrogenase Complexes

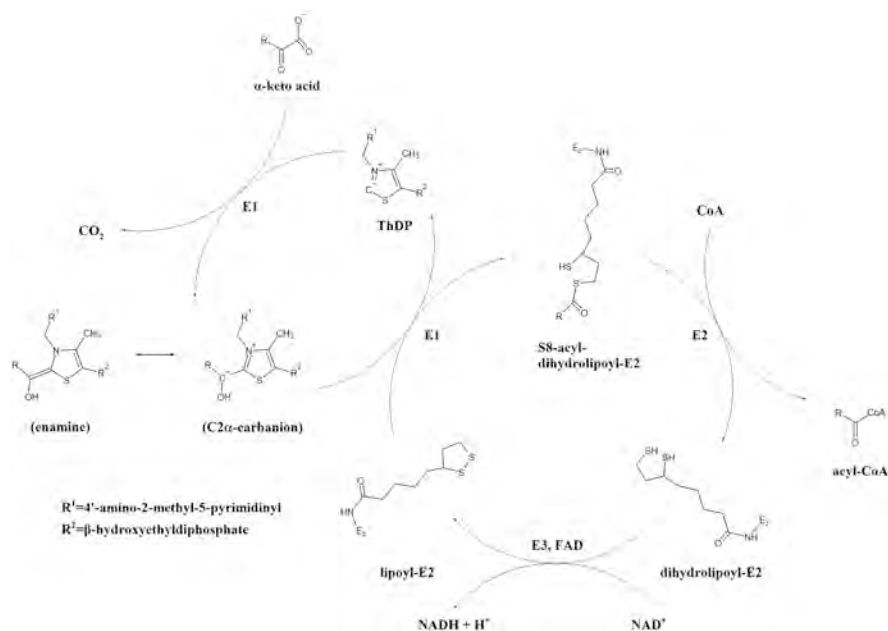
The mitochondrial α -keto acid (or 2-oxo acid) dehydrogenase complexes are multi-subunit enzymes with crucial metabolic roles. The enzyme family comprises the pyruvate (PDHc), α -ketoglutarate (KGDHc), branched-chain α -keto acid (BCKDHc), and α -ketoadipate (KADHc) dehydrogenase complexes that catalyze the oxidative decarboxylation of their specific α -keto acid substrates while generating the respective acyl-CoA products and reducing NAD^+ to $\text{NADH} (+\text{H}^+)$ (Reed 1974; Yeaman 1989); α -ketoadipate and α -ketoglutarate are α -keto acids with ω -carboxylic end groups, while pyruvate and the implicated branched-chain α -keto acids (see below) are α -keto acids with aliphatic end groups.

By producing acetyl-CoA from pyruvate, the PDHc connects glycolysis to the citric acid/(Szent-Györgyi-)Krebs/tricarboxylic acid (TCA) cycle as well as to the biosynthetic pathways of certain fatty acids and steroids. The BCKDHc and KADHc participate in the catabolism of selected amino acids channeling eventually acetyl-CoA [BCKDHc and KADHc] and succinyl-CoA [BCKDHc] into the Krebs cycle; the substrates of the BCKDHc are the oxoacids α -ketoisocaproate, α -keto- β -methylvalerate and α -ketoisovalerate [direct products: isovaleryl-CoA, α -methylbutyryl-CoA, and isobutyryl-CoA, respectively] that form from the branched-chain amino acids Leu, Ile, and Val, respectively, by transamination, while the KADHc converts α -ketoadipate [direct product: glutaryl-CoA], a common metabolic intermediate in the Lys, hydroxy-Lys, and Trp degradation pathways. The BCKDHc also accepts α -ketobutyrate as substrate and converts it to propionyl-CoA (that eventually leads to succinyl-CoA); α -ketobutyrate is a degradation product of Met (and in prokaryotes Thr). The fate of acetyl-CoA produced by the above three enzyme complexes [PDHc, KADHc, BCKDHc], e.g., mitochondrial oxidation or feeding lipid synthesis pathways in the cytosol (see more details below), depends on the activities of the KGDHc and isocitrate dehydrogenase, rate-limiting enzymes in the Krebs cycle; the KGDHc generates succinyl-CoA from α -ketoglutarate in the Krebs cycle. Further metabolic roles of succinyl-CoA include being the precursor for porphyrin synthesis (see Scheme 13.1).

A hybrid complex could also form between the hKGDHc (h for human) and hKADHc both in vitro (Nemeria et al. 2018) and in vivo (Leandro et al. 2020a), which suggests a link between the Krebs cycle and the Lys/Trp catabolic pathways (see further details below).



Scheme 13.2 General reaction catalyzed by the α -keto acid dehydrogenase complexes



Scheme 13.3 Catalytic roles of the subunits/components in the α -keto acid dehydrogenase complexes

lipoate prosthetic group is reoxidized by the FAD-dependent E3 component ([dihydro]lipoamide/lipoyl dehydrogenase, [DH]LADH) producing $NADH (+H^+)$ (Massey 1960a; Ozohanics et al. 2023).

In the above-discussed series of reactions, the rate-limiting steps differ among the cognate complexes. For instance, based on kinetic studies using recombinant *E. coli* enzymes, the rate-limiting step is the formation of the first covalently attached pre-decarboxylation intermediate on the E1p (lactyl-ThDP) for the PDHc (Balakrishnan et al. 2012; Berg et al. 1998; Cate et al. 1980; Patel et al. 2014), while it is the E2k-catalyzed succinyl transfer to CoA for the KGDHc (Chakraborty et al. 2018). For the hKADHc and hBCKDHc, the reductive acylations of the respective hE2-lipoates by the corresponding hE1 components are the rate-limiting steps (Nemeria et al. 2018; Li et al. 2004).

Substrate specificity is primarily provided by the E1 and to a lesser degree by the E2 components, while the E3 component with the regenerative function is shared

among all the complexes. The E1b component has evidently the broadest substrate specificity as it accepts all the implicated branched-chain α -keto acids and also α -ketobutyrate, among a few other compounds, as substrates (see above), e.g., the bovine E1b also converts pyruvate (Pettit et al. 1978) and 4-methylthio- α -ketobutyrate, an alternative degradation product of Met (Jones and Yeaman 1986). Observed maximal reaction rates were 70, 50, 40, 30, and 20% for α -ketoisocaproate, α -ketobutyrate, α -keto- β -methylvalerate, 4-methylthio- α -ketobutyrate, and pyruvate, respectively, relative to α -ketoisovalerate (Pettit et al. 1978; Jones and Yeaman 1986). In terms of substrate specificity, there are overlaps, although to much smaller degrees, in the other complexes, as well. For instance, hE1a is also capable of converting α -ketoglutarate (Nemeria et al. 2018) and α -ketopimelate (Nemeria et al. 2022), hE1k can use α -ketoadipate (Nemeria et al. 2018), and ecE1k can also operate with pyruvate and α -ketovalerate (Shim et al. 2011). Acceptance of non-traditional/physiological substrates in an in vitro subunit-specific enzyme assay likely plays little significance in vivo, since subsequently, the relevant E2 component also has to be able to accept the emerging intermediates as substrates. For example, although ecE1k could use pyruvate or α -ketovalerate (see above), no overall ecKGDHc activity could be observed with these substrates, implying discrimination at the E2k level (Shim et al. 2011). Similarly, benzoylformate could readily be decarboxylated by hE1b; however, the resulting intermediate could not be further processed by hE2b, leading to the (irreversible [suicidal]) inhibition of the hBCKDHc (Machius et al. 2006). The hE2k component, however, possesses comparable succinyl- and glutaryltransferase activities (see above that hE1k can also use α -ketoadipate and that there also exists a functional hybrid hKADHc/hKGDHc complex) (Nemeria et al. 2018). Interestingly, the lipoate-carrying (or lipoyl) domains (LDs) of the ecE2p and ecE2k subunits could only serve as proper substrates for the ecE1p and ecE1k components, respectively (Graham et al. 1989) (see below).

Overall Structures of the Complexes

The α -keto acid dehydrogenase complexes possess fundamentally similar structural architectures inasmuch as they all comprise multiple copies of three catalytically active components (E1, E2, E3); the respective subunits display high structural similarity over different species and cognate complexes. The E2 components are of particular structural importance since these also form unique scaffolds to which the relevant E1 and E3 subunits can directly or indirectly and usually selectively tether as peripheral components. The active sites of the three catalytic components are interconnected by the flexible, lipoylated arms of the E2 core structure. In certain members of the family, these “mandatory” components are supplemented with additional structurally important (E3-binding protein, E3BP, in most eukaryotic PDHc) or regulatory (kinases and phosphatases in most eukaryotic PDHc and BCKDHc) proteins/enzymes. Stability and activity of the BCKDHc isolated from rat liver proved also to depend on the presence of K^+ (Shimomura et al. 1988).

There are a few exceptions to the above rather general structural architecture (individual/unique E1-3 subunits/components) in nature. For instance, catalytic domains of the E1k and E2k subunits are expressed as a single protein chain with no characteristic lipoyl prosthetic group in *Corynebacteriales* (actinobacteria), therefore, hybrid complexes form where an E2p subunit provides the lipoate moiety (Hoffelder et al. 2010; Wagner et al. 2011); in *Mycobacterium smegmatis* and *Corynebacterium glutamicum*, this two-in-one protein (OdhA) assembles into a 800-kDa homohexamer that adopts a three-blade propeller shape (Yang et al. 2023). Or the *Pseudomonas putida* (Gram-negative soil bacterium) expresses two different E3 isoforms, and E3-Val is specifically expressed for the assembly of the BCKDHc, while E3-Glc is part of the PDHc and KGDHc (Sokatch et al. 1981). In the diplomid *Paradiplonema papillatum*, for the PDHc activity, E1p is replaced by the archaeal protein AceE that partners with the E2 and E3 subunits from the BCKDHc and/or KGDHc (Zahonova et al. 2023). Furthermore, as was already mentioned above, the hKGDHc and hKADHc form a hybrid complex, where they in fact also share the hE2k component, while the hE1a and hE1k components exhibit activities of the same order of magnitude [in the muscle, heart, and brain (Artiukhov et al. 2020)]; interestingly, no E1a protein has been yet identified in prokaryotes.

Mitochondrial supercomplexes were shown to exist in in vitro studies of mitochondria (Dudkina et al. 2010; Javadov et al. 2021). These clearly included the electron transport system protein complexes (CI, CIII, CIV or CII, CIII, CIV), which were the focus of most studies, but unknown associated proteins were also visible on cryo-electron tomography images. Earlier however, Lyubarev and Kurganov (Lyubarev and Kurganov 1989), based on enzyme interaction data, proposed a supercomplex-like organization of the TCA cycle enzymes with succinate dehydrogenase being one of the anchor points to the mitochondrial membrane. Novel results from Plokhikh et al., where presumably E1/E3 associates were visible near the CI and also in close proximity to the E2 core, point toward a broad inclusion of proteins in the mitochondrial supercomplexes, with PDHc and KGDHc being anchored near the membrane and CI (Plokhikh et al. 2023; Sumegi and Srere 1984).

Structural Characteristics of the Individual Components

Structures of the E1 Components

The E1 component catalyzes the first two steps in the catalytic cycle using a thiamine diphosphate (ThDP, or thiamine pyrophosphate, TPP) prosthetic group and Mg^{2+} . The E1 proteins in the cognate complexes show surprisingly low sequence identity (<20%), yet they all adopt the characteristic α -keto acid dehydrogenase fold (Fig. 13.1). The thiamine-containing active sites reside at the subunit interfaces (Frank et al. 2007); this was first described in the related transketolase enzyme (Lindqvist et al. 1992). Even though these enzymes exist in either dimeric or

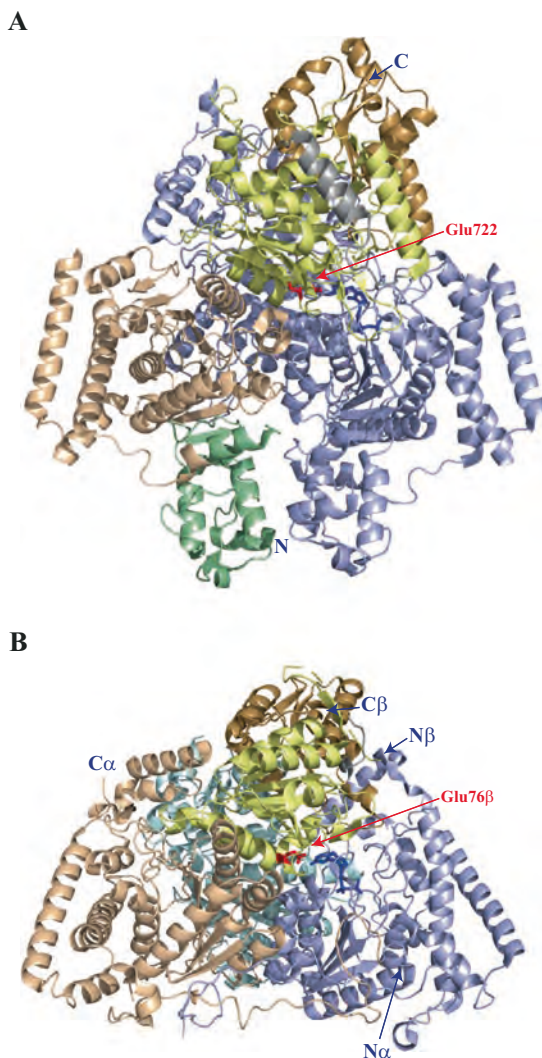


Fig. 13.1 Representative structural models of human E1 components. **(a)** Cryo-EM structure of the homodimeric hE1k component at 2.92 Å resolution (PDB ID: 7WGR; Zhong et al. 2022). Domains in one of the monomers are depicted by different colors (pale green for the N-terminal, wheat for the $\alpha\beta 1$, pale yellow for the $\alpha\beta 2$, and sand for the $\alpha\beta 3$ domain; the linker regions and the adjacent monomer are colored in gray and light blue, respectively). The active site residue and the corresponding ThDP cofactor are represented by red and blue sticks, respectively. **(b)** Crystal structure of the heterotetrameric hE1b component at 1.39 Å resolution (PDB ID: 2BFD; AEvarsson et al. 2000; Wynn et al. 2001). Domains in one of the α - β heterodimers, which are functionally equivalent to a single hE1k monomer, are depicted by different colors (wheat for the $\alpha\beta$ domain *plus* the C-terminal stretch in the α subunit, while pale yellow and sand for the N- and C-terminal domains, respectively, in the β subunit; the linker region is in gray). The active site residue and the corresponding ThDP cofactor are represented as in panel **a**. The α and β subunits in the adjacent heterodimer are colored in light blue and pale cyan, respectively

tetrameric form (see Fig. 13.1), they always function as “dimers of active sites”; the hE1a (Bezerra et al. 2020; Leandro et al. 2020b), E1k in all species (Frank et al. 2007; Zhong et al. 2022), E1b in the Gram-positive *Lactococcus lactis* bacteria (Berthold et al. 2007), and E1p in Gram-negative bacteria form homodimers (Arjunan et al. 2002), whereas the E1b in most species (Pettit et al. 1978; Wexler et al. 1991; Nakai et al. 2004) and E1p in Gram-positive bacteria and eukaryotes (Perham 1991; Ciszak et al. 2003) form heterotetrameric assemblies ($\alpha_2\beta_2$) encapsulating two active sites [here an $\alpha\beta$ heterodimer is functionally equivalent to an E1k chain (Frank et al. 2007)]. Each protein chain in the homodimeric E1p component comprises a flexible N-terminal stretch and three domains; the three domains—namely, the N-terminal, middle, and C-terminal domains—all adopt α/β -class folds (Arjunan et al. 2002; Park et al. 2004). The pyrophosphate moiety in ThDP is coordinated by the N-terminal domain in one subunit, whereas the aminopyrimidine moiety is tethered by both the same N-terminal domain and the middle domain in the adjacent monomer (Arjunan et al. 2002); exact locations of the thiamine-binding residues/patches and the conformation of the prosthetic group may vary among species and/or ThDP-dependent enzymes. The E1k and hE1a components are more closely related to the heterotetrameric E1p and E1b components in terms of intrinsic structure while also adopting the above-described domain structure (Frank et al. 2007; Zhong et al. 2022; Bezerra et al. 2020; Leandro et al. 2020b; AEvarsson et al. 2000; Ciszak et al. 2003).

The N-terminal stretches in E1k in most species and E1p in Gram-negative bacteria participate in the assembly of the related complexes *via* interacting with the corresponding E2k catalytic core domain (CD) and E3 component (Frank et al. 2007; McCartney et al. 1998; Zhou et al. 2018; Rice et al. 1992; Nakano et al. 1994) and the peripheral subunit-binding domain (PSBD) of E2p (Park et al. 2004; Song et al. 2010; Patel et al. 2014), respectively. E1k has an extended N-terminal region with varying lengths in different species (Frank et al. 2007) that usually carries two E2k-binding motifs (Zhou et al. 2018), not present in hE1a (Bezerra et al. 2020). The full-length N-terminal regions in ecE1p and hE1k could structurally not be investigated due to their high flexibility (Arjunan et al. 2002; Zhong et al. 2022). This same phenomenon even prevented ecE1k from crystallization; hence, only a truncated protein could be crystallized and structurally analyzed (Frank et al. 2007; Bezerra et al. 2020; Leandro et al. 2020b).

In the heterotetrameric E1p (Perham 1991; Ciszak et al. 2003) and E1b (AEvarsson et al. 2000) components, the above-described domains are somewhat redistributed among the α and β subunits; each α subunit comprises a domain that resembles the homodimeric N-terminal domain, while each β subunit contains a homodimeric middle domain-like domain (here as an N-terminal domain) and a homodimeric C-terminal domain-like domain (here also as a C-terminal domain). The above domain in the α subunit and the N-terminal domain in the β subunit coordinate the ThDP and stabilize the heterotetramer mostly by hydrophobic interactions (AEvarsson et al. 2000; Ciszak et al. 2003). The flexible N-terminal stretches present in E1k, E1p in Gram-negative bacteria, and hE1a are completely absent in the heterotetrameric E1p and E1b, owing likely to the different overall structural

arrangements in the BCKDHC and eukaryotic PDHC. In the latter complexes, the C-terminal domains in the E1- β subunits associate with the respective E2-PSBDs (Wynn et al. 1992; Rahmatullah et al. 1989); see more on complex assembly below. The hE1p component structure is tetrahedral, where the identical subunits are related by twofold symmetry axes (Ciszak et al. 2003).

Near the active sites, which are buried in the subunit interfaces, the respective substrate-binding pockets as well as ThDP, adopting a characteristic “V” shape, can be found (Machius et al. 2006; Arjunan et al. 2002). The prosthetic group is generally activated by a conserved Glu or His residue serving as a catalytic base (Nakai et al. 2004; Ciszak et al. 2003; Wynn et al. 2001). The active sites in the hE1k and hE1a are highly similar, but only some of the substrate-binding residues are conserved. Three loops are conserved in the active site regions of the two proteins, while a fourth one is somewhat shorter in hE1k leading to a substrate-guiding channel with less positive surface polarity. The longer fourth loop in hE1a facilitates the binding of the longer α -ketoadipate substrate (Zhong et al. 2022). HDX-MS (Nemeria et al. 2021) and molecular modeling (Ozohanics et al. 2023) results show that the hE1a N-terminal region is not a focal point in binding the hE2k component; although it might be needed for guiding the assembly of the overall complex, the majority of the binding interactions rather originate from the interdomain linker and the middle domain in hE1a and lead toward the core domain in hE2k. In hE1a, a His residue (His435) proved to play key roles in the interaction with hE2k and the acyl transfer reaction (Nemeria et al. 2021). HDX-MS was also used to study the hE1k-hE2k interaction, where the data pinpointed the ThDP binding domain as the focus of interactions (Zhou et al. 2018). hE2k utilizes identical regions in its core domain to interact with either the hE1a or hE1k (Ozohanics et al. 2023). The signatures of the Ca^{2+} -binding sites that can be found in the Ca^{2+} -activated hE1k are not present in hE1a (Bunik and Degtyarev 2008).

Multimeric ThDP-dependent enzymes were all proposed to function *via* alternating active site mechanisms (also referred to as half-the-sites-reactivity) (Frank et al. 2004; Khailova et al. 1991; Kovina and Kochetov 1998; Sergienko and Jordan 2002). Catalytic events in the two spatially separated active sites are presumably synchronized *via* reversible proton transfer (a “proton wire”) between the two ThDP prosthetic groups (Frank et al. 2004; Schröder-Tittmann et al. 2013). Such proton wires were described in the crystal structures of the *Bacillus stearothermophilus* (Frank et al. 2004) and *E. coli* (Nemeria et al. 2010) E1p components, where mutating the wire-forming acidic residues resulted in decreased catalytic activities. The proton wire is also present in the ecE1k component; however, here mutagenesis studies have not yet been performed (Frank et al. 2007). In concert with these, it was reported that in eukaryotic and Gram-positive bacterial E1p components, one catalytic center decarboxylates the pyruvate moiety, while the other center concurrently reductively acetylates the lipoamide prosthetic group (Sergienko and Jordan 2002); roles are alternating in a so-called “flip-flop” mechanism (Frank et al. 2004; Khailova et al. 1991; Kovina and Kochetov 1998; Sergienko and Jordan 2002). Unlike the eukaryotic E1p active sites, the hE1b active sites operate independently; hence, generality of the alternating active site mechanism of the E1 components

(and of the ThDP-dependent enzymes) was disputed (Li et al. 2007). The mechanism of the acyl transfer to the E2-LD by E1 was most extensively studied on ecP-DHc; however, the findings should also translate well to the cognate complexes, given the extensive homology of the subunits. The reaction is a reductive acyl transfer between the enamine form of the E1-bound acyl-ThDP, which provides a nucleophile, and the disulfide bond in the dithiolane ring of the lipoate cofactor (Chiu et al. 1996). The subsequent formation of the acyl-dihydrolipoate derivative leads to the liberation of the E1-ThDP (see Scheme 13.3); two key His residues facilitate these concerted actions (Song and Jordan 2012; Chakraborty et al. 2018; Tittmann 2009).

In Gram-negative bacteria, 12 homodimers of E1 participate in building up the quaternary structures of the PDH complexes, compared to eukaryotes and Gram-positive bacteria with 12–40 heterotetramers. In the KGDH complexes, 6–12 E1 homodimers exist. In the hBCKDHc, 12 E1 heterotetramers are present (see Appendix 3).

Structures of the E2 Cores

The E2 subunits assemble to various symmetric oligomers and as such form the structural and mechanistic cores in the α -keto acid dehydrogenase complexes (Perham 1991). They (1) form the scaffolds to which the respective E1 and E3 subunits associate as peripheral components, (2) exhibit the dihydrolipoyl transacylase activities, and (3) shuffle reaction intermediates within the complexes, from one active site to the other. The multifaceted functions are supported by multidomain structures; an individual E2 subunit may comprise a catalytic core domain (CD), a peripheral subunit-binding domain (PSBD) [except for the mammalian E2k (Nakano et al. 1991; Zhou et al. 2018)], and one to three lipoyl domains (LD). The up to three consecutive N-terminal LDs each bear a covalently bound lipoate cofactor, and all of them together act as a swinging arm that interconnects all types of the active sites (E1–3) in the overall complexes (Perham 1991; Reed 1974). Domains are separated by unstructured Pro-, Ala-, and charged amino acid-rich linker regions (Perham 1991). Owing to these highly flexible linker regions, high-resolution structures of full-length E2 components have hitherto not been published; however, individual E2 domains were investigated in various complexes.

In the core structures, E2 subunits form trimeric components usually *via* their CDs; the monomers are related by a threefold rotational axis (Mattevi et al. 1992a; Knapp et al. 1998; Nagy et al. 2021; Liu et al. 2022; Kato et al. 2006). Twenty (in E2p cores in eukaryotes and Gram-positive bacteria) and eight (in all other E2 cores) trimers then assemble into higher-order structures forming 60-meric pentagonal dodecahedrons with icosahedral symmetry (Jiang et al. 2018; Prajapati et al. 2019; Vijayakrishnan et al. 2011; Yu et al. 2008; Liu et al. 2022) and 24-meric cubic structures with octahedral symmetry (Mattevi et al. 1992a; Knapp et al. 1998; Nagy et al. 2021; Liu et al. 2022; Kato et al. 2006; Wagenknecht et al. 1990; Skerlova et al. 2021), respectively (Fig. 13.2) (see a few exceptions in Appendix 3). Its

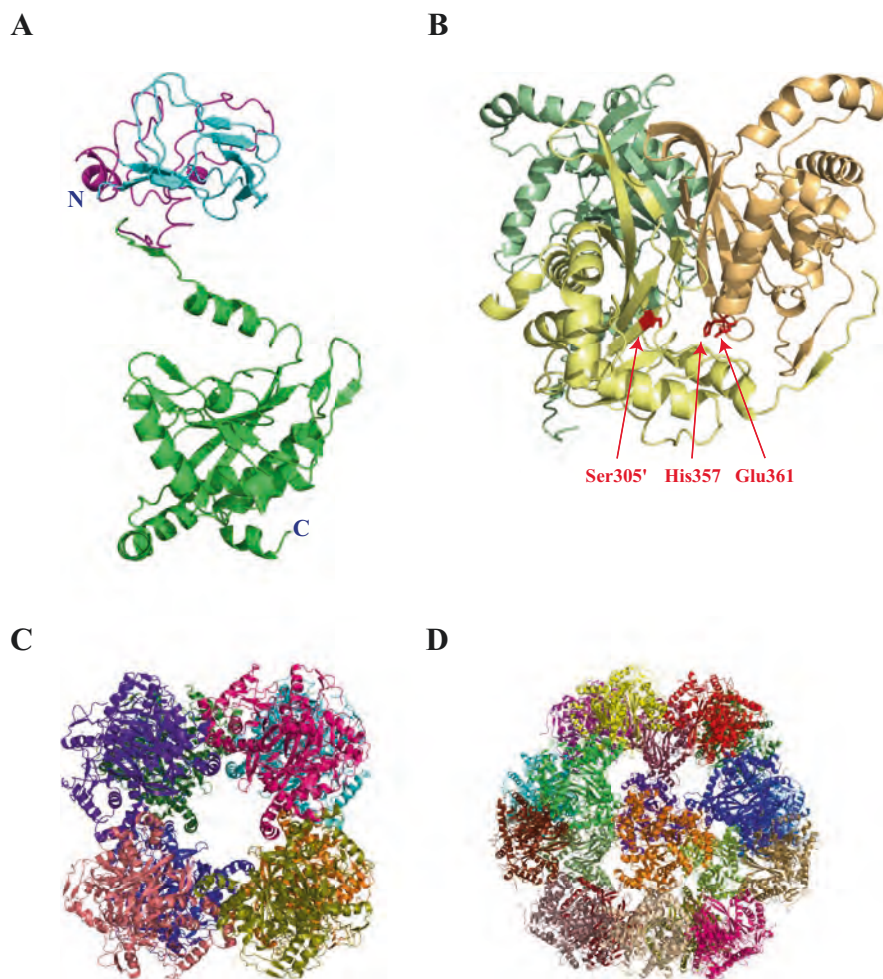


Fig. 13.2 Representative structural models of human E2 assemblies. **(a)** Overall conformation of the intact mature hE2k monomer as defined by cryo-EM (core domain in green) and chemical cross-linking-constrained homology modeling (lipoyl domain in cyan *plus* PSBD-like middle linker in magenta). The N- and C-termini are labeled by capital letters in blue. **(b)** Cryo-EM structural model of the hE2k homotrimer. The three monomers are illustrated by different colors. The active site residues are represented as red sticks. The residue with prime belongs to an adjacent monomer. **(c)** 24-meric model structure of hE2k at 2.90 Å resolution (PDB ID: 6H05; Nagy et al. 2021). This assembly adopts an 8x3 symmetry, where the hE2k homotrimers occupy each vertex of a cube. Each trimer is depicted by a different color. **(d)** 60-meric model structure of hE2p at 3.10 Å resolution (PDB ID: 6CT0; Jiang et al. 2018). The structural assembly possesses a 20x3 symmetry, where hE2p homotrimers occupy each vertex of a pentagonal dodecahedron. Each trimer is depicted by a different color

structural topology and the fact that it also recruits structurally similar but non-catalytic E3-binding protein (E3BP) subunits make the eukaryotic PDHc cores

unique in this family of enzymes; according to the different proposed models, 12 (Hiromasa et al. 2004) or 20 (Brautigam et al. 2009) E3BP chains replace, or 12 E3BP chains supplement (Smolle et al. 2006), the E2p chains. The E3BP comprises a single LD, an E3-binding domain (E3BD), and a (catalytically inactive) C-terminal domain; due to its single LD, E3BP may only be capable of interacting with certain E1–3 active sites.

Several cryo-EM structures are now available (Jiang et al. 2018; Prajapati et al. 2019; Vijayakrishnan et al. 2011; Yu et al. 2008; Liu et al. 2022). More recent studies on mammalian PDH complexes (see Figs. 13.2 and 13.3), together with biophysical experiments, established the existence of the substitution model, although it is still not entirely clear which stoichiometry is valid (48E2:12E3BP model (Jiang et al. 2018; Vijayakrishnan et al. 2011), 40E2:20E3BP model (Prajapati et al. 2019), supported also by their ITC analysis); interpretation of related experimental data at times is obscure due to occasional heterogeneity in the E3BP distribution (Prajapati et al. 2019). Interestingly, in fungi (*N. crassa*), an additional model was concluded (Forsberg et al. 2020). Differences were also found between recombinant hE2p cores expressed in the presence of excess E3BP and native bovine E2p cores (Vijayakrishnan et al. 2011). Another potential explanation for the observed differences is that the average E2:E3BP composition may differ among tissues. Nevertheless, it became clear that the E3(dimer)/E3BP binding stoichiometry is 1:1 (Prajapati et al. 2019), hence, the number of the E3BP chains determines the maximal number of the E3 chains tethered; however, the former is not unambiguously established: 6–12 E3 dimers are generally present in the above PDHcs.

To some surprise, the distribution of the tethered E1 and E3 components at times also proved to be uneven and local clusters could be identified for both components (Prajapati et al. 2019); the presence of E3 patches also suggests that the E2/E3BP subunits may not always be present in an E2₂-E3BP₁ trimeric form, as it was hypothesized earlier. In fact, E2₃ and E2₁-E3BP₂ trimers were indeed identified by native mass spectrometry (Prajapati et al. 2019). Interestingly, in the ecE2p homotrimeric component (note that *E. coli* is a Gram-negative bacterium), among the monomers, a task-sharing inter-chain acetyl transfer mechanism was identified that suggests different roles for each chain in a trimer (Song and Jordan 2012); a similar “division-of-labor”-type mechanism was also proposed for the local E2-E1 and E3BP-E3 clusters in the hPDHc (see above) (Prajapati et al. 2019).

The E2 subunits/components (and the E3BP) are structurally highly similar in the cognate complexes; however, the specific intra- and intertrimeric interfaces are still capable of securing adequately selective self-assembly in the mitochondrion (Liu et al. 2022). Trimers internally primarily stabilize *via* H-bonding, whereas the neighboring trimers interact *via* dedicated hydrophobic pockets in the E2/E3BP core domains (“knob-to-socket”-type interactions) (Kato et al. 2006). The intratrimeric interactions are significantly stronger than the intertrimeric ones; hence, the E2 cores are rather flexible, capable of “breathing,” which could be important for catalysis (Byron and Lindsay 2017). The unique overall symmetry in the eukaryotic (and Gram-positive bacterial) E2p cores is due to distinctive sets of intertrimeric interactions (Liu et al. 2022). Oligomerization of E2 is also essential for catalytic

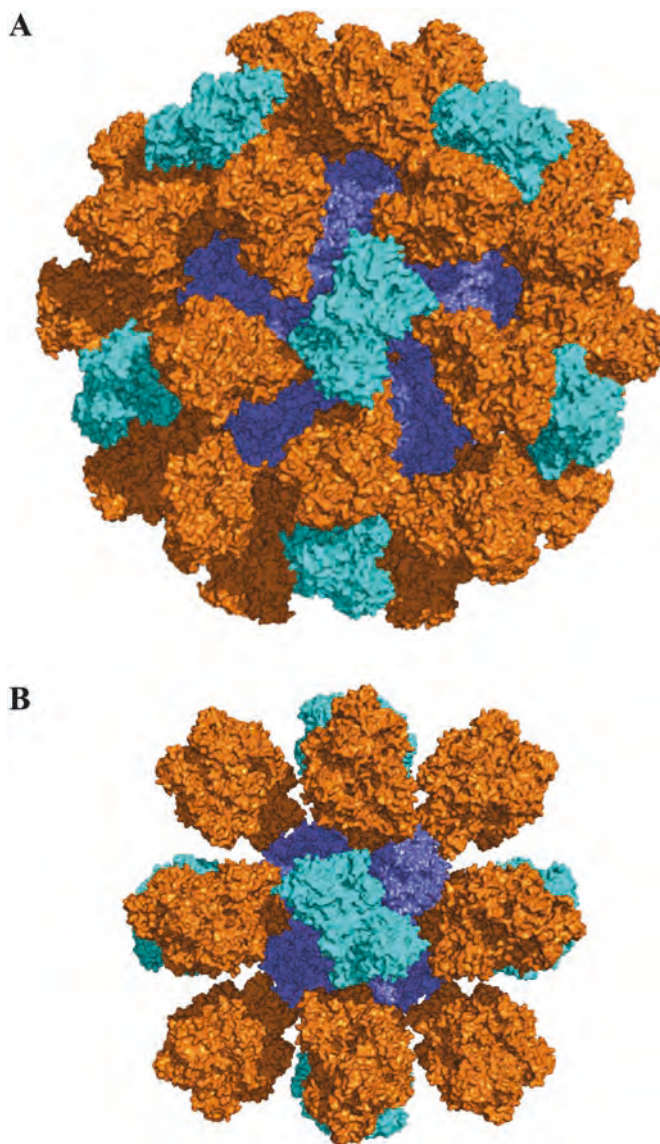


Fig. 13.3 Overall structural topologies of the hPDHc and hKGDHc. **(a)** Illustrative structural model of the overall hPDHc. The corresponding components are the core-forming hE2p (PDB ID: 6CT0, Jiang et al. 2018), the peripheral hE1p (PDB ID: 3EXE, Kato et al. 2008), and also peripheral hE3 (see below, PDB ID, 6I4Q, Szabo et al. 2019) colored in slate, orange, and cyan, respectively. **(b)** Illustrative structural model of the overall hKGDHc. The corresponding components are the core-forming hE2k (PDB ID, 6H05, (Nagy et al. 2021)), the peripheral hE1k (PDB ID: 7WGR, (Zhong et al. 2022)), and also peripheral hE3 (PDB ID: 6I4Q, (Szabo et al. 2019)). Colors are as in panel **a**. The protein structures are represented as solvent accessible surfaces. The rather flexible N-terminal regions for the core-forming components and hE3BP (for panel **a**, see Jiang et al., 2018) are not included. Construction of the structural models is based on Liu et al., 2022 (Fig. 1B therein)

activity: the active sites reside within solvent accessible channels (~30 Å in length with two entries) at the interfaces of two adjacent monomers (Mattevi et al. 1992a, 1993a; Knapp et al. 1998; Nagy et al. 2021; Liu et al. 2022; Kato et al. 2006). The *A. vinelandii* E2p structure (Mattevi et al. 1993a), and later also the bovine E2b structure (Kato et al. 2006) showed that the CoA approaches the active site through the channel leg that opens to the central cavity, while the LD-bound acyl-DHLA substrate enters from the outer surface of the cube-shaped core structure; molecular modeling proposed similar substrate-binding patterns for ecE2k and hE2k as well (Knapp et al. 1998; Nagy et al. 2021). The crystal structures of the *A. vinelandii* E2p-CoA binary and E2p-CoA-DHLA ternary complexes revealed two distinct binding modes for the CoA (extended [“in”] and compact/helical [“out”] conformations) (Mattevi et al. 1993a). A series of crystal structures of the bovine E2b under various substrate conditions revealed similar states in CoA binding and also a unique gating mechanism in which CoA binding is responsible for the opening of the acyl-DHLA binding site (Kato et al. 2006). The binding of CoA to hE2p altered the structures of the core region and the overall hPDHc (Prajapati et al. 2019).

The E2 active sites are highly conserved consisting of a **DH(R/Q)XX(D/N)G** consensus sequence in the C-terminal domain (the bold residues are highly conserved, while the ones in parentheses are likely the most frequent residues in that position); a His, an Asp, or Asn and a Ser or Thr (not part of the above sequence) play key roles in the catalysis as proved by mutagenesis studies (Meng and Chuang 1994; Chakraborty et al. 2018; Russell and Guest 1990; Yu et al. 2008). The corresponding His and Asp/Asn residues apparently either stabilize the tetrahedral oxyanionic intermediate (ecE2k) (Chakraborty et al. 2018) or participate in the relevant acid-base-type catalytic steps [Asp may trigger the effective tautomer of His for deprotonation (Mattevi et al. 1993a; Nagy et al. 2021)]. The role of the Ser/Thr residue has not been entirely resolved; it either also stabilizes the tetrahedral intermediate (Mattevi et al. 1993a; Meng and Chuang 1994) or simply forms a H-bond, *via* its peptide N atom, with the pantetheine tail of CoA (Kato et al. 2006). The catalytic mechanism in this step involves a nucleophilic attack by the CoA-thiolate onto the S8-acyldihydrolipoamide *via* the formation of a tetrahedral oxyanionic intermediate; mutagenesis ruled out the possibility of a nucleophilic His attack (Chakraborty et al. 2018).

As was mentioned above, E2 components contribute to the substrate specificities of the cognate complexes. The residues along the substrate-binding channels and the active sites are relatively well conserved among the various E2 components, with a few exceptions that presumably account for the differing substrate specificities.

The E2k and E2b monomers carry a single LD in most species, whereas E2p comprises one to three LDs depending on the species; the consecutive LDs proved to be functionally redundant (Guest et al. 1985). The most plausible reason for having multiple LDs is to help the lipoate coenzyme efficiently reach all the E1–3 active sites. Overall ecPDHc and ecKGDHc activities could partially be restored *via* reconstituting the ecE1p/E1k and ecE3 components with the respective ecE2 core domains and LD/LD-PSBD constructs indicating that the relevant reaction

intermediates could still be transferred among the ecE1–3 active sites with no covalent bonding between the LD and the core domain in ecE2 (Chakraborty et al. 2018; Wang et al. 2014a). LDs adopt β -barrel-type folds created by two four-stranded β -sheets around hydrophobic interior cores (Dardel et al. 1993; Green et al. 1995; Chang et al. 2002; Ricaud et al. 1996; Berg et al. 1996). The lipoate cofactor binds to the ϵ -amino group of a lysine residue that resides at the tip of a type 1 β -turn between strands 4 and 5, while a prominent loop between strands 1 and 2 (loop L1) plays important roles in the specificity of the E1–E2 interaction (Jones and Perham 2008; Berg et al. 1998; Jones et al. 2000; Wallis and Perham 1994; Chang et al. 2002). As was mentioned above, E2p-LD and E2k-LD were exclusively recognized by the Elp and Elk components, respectively, in *E. coli* (Graham et al. 1989).

For the scaffolding function, most E2 components carry a single PSBD that binds the E1 and/or E3 components; the mammalian E2k represents an exception as it has no PSBD (Rice et al. 1992; Nakano et al. 1994) (see more on assembly later).

Structure of the Common E3 Component

The common E3 component is a FAD-dependent obligate homodimer in all species, according to the crystal structures published to date (Schierbeek et al. 1989; Mattevi et al. 1991, 1992b, 1993b; Glasser et al. 2017; Faure et al. 2000; Mande et al. 1996; Li de la Sierra et al. 1997; Toyoda et al. 1998; Rajashankar et al. 2005; Nakai et al. 2008; Chandrasekhar et al. 2013; Bryk et al. 2013; Brautigam et al. 2005, 2006, 2011; Ciszak et al. 2006) (Fig. 13.4; Szabo et al. 2019). The two active sites are located at the homodimer interface where both two monomers provide catalytically important residues; one of the monomers provides the majority of the active site residues, including the redox-active disulfide cysteines, the FAD prosthetic group, and the stabilizing amino acids, as well as the NAD⁺/NADH binding site, while the other monomer provides the catalytic base (a His residue) (Mattevi et al. 1991; Brautigam et al. 2005; Szabo et al. 2019). The isoalloxazine ring of FAD plays an organizing role in the active site: the redox-active disulfide, the catalytic base His, and the dihydrolipoate binding site all reside at the *si* face of this ring system, while the NAD⁺/NADH binding site is located at the *re* face. This also means that the reductive and oxidative half-reactions, comprehensively governed by a ping-pong bi-bi mechanism, are spatially separated by the isoalloxazine ring system of FAD (Mattevi et al. 1991; Brautigam et al. 2005).

In the active site, positioning of the redox-active Cys residues clearly supports the respective catalytic roles: the “distal” Cys, which resides in the substrate binding cavity farther away from the FAD, forms a mixed disulfide bond with the substrate DHLA, while the “proximal” Cys, located closer to the FAD, gets converted to a charge-transfer complex with the FAD, later in the catalytic mechanism (Williams Jr 1992; Mattevi et al. 1991; Brautigam et al. 2005; Szabo et al. 2019; de Kok and van Berkel 1996). The catalytic base His, which was identified on the basis of the crystal structures and was later also confirmed by site-directed mutagenesis

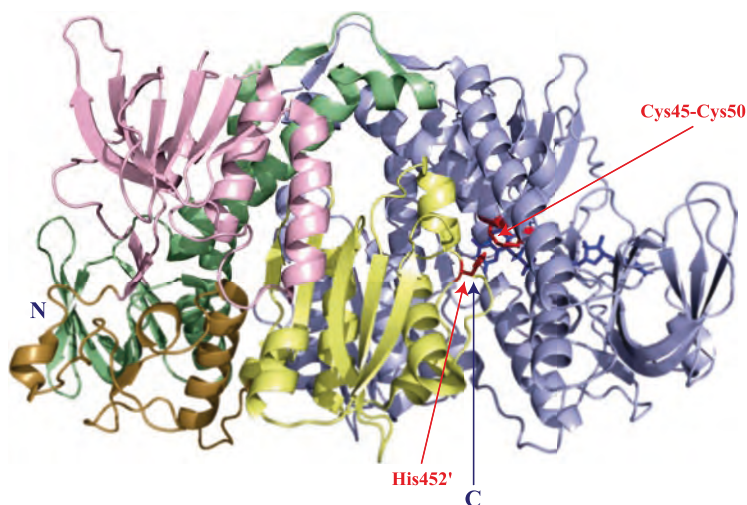


Fig. 13.4 Crystal structure of the human E3 homodimer at 1.75 Å resolution (PDB ID: 6I4Q, Szabo et al. 2019). Domains in one of the monomers are illustrated by different colors (pale green for the FAD-binding, light pink for the NAD⁺/NADH-binding, sand for the central, and pale yellow for the interface domain). The redox-active disulfide bridge in the active site and the corresponding FAD prosthetic group are represented by red and blue sticks, respectively. The adjacent monomer is colored in light blue. The N- and C-termini are labeled by capital letters in blue

(Williams Jr. et al. 1989; Benen et al. 1992; Kim and Patel 1992), protrudes toward the distal Cys in the active site. The *cis* peptide bond with the sequentially adjacent Pro and the H-bond formed with the side-chain carboxylic group of a key Glu are of great importance in the proper orientation of this His (Mattevi et al. 1991; Brautigam et al. 2005; Szabo et al. 2019). Additionally, the electron-withdrawing effect of the above Glu is likely indispensable in properly adjusting the side-chain pK_a value of this His, which secures the efficient deprotonation of the substrate DHLA (Benen et al. 1992; Kim and Patel 1992).

The dihydrolipoylysine moieties attached to the LD arms of the E2(-E3BP) components are proposed to approach the E3 active site through a ~10 Å long, rather hydrophobic channel (de Kok et al. 1994). This channel continues, passing the active site, in a perpendicular leg of rather hydrophilic character (>20 Å in length) (de Kok et al. 1994; de Kok and van Berkel 1996; Szabo et al. 2018, 2019). This channel leg could not permit the dihydrolipoylysine moiety (~14 Å) (Reed 1974) to reach to active site; instead, it was proposed that it serves as a physical outlet for H₂O (de Kok et al. 1994) and/or—the catalytically liberated—H⁺ (H₃O⁺) (de Kok and van Berkel 1996). Recently, a 1.75 Å resolution crystal structure of the hE3 revealed two key residues, a Glu and an Arg (from the adjacent monomer), in the above mentioned “H₂O/H⁺-channel” that are clearly capable of each adopting two alternative conformations. The synchronous movements of the respective side chains appear to dynamically modulate the geometry, polarity, and extended H₂O

network of the $\text{H}_2\text{O}/\text{H}^+$ -channel, which might play a role in $\text{H}^+/\text{H}_3\text{O}^+$ transduction (Szabo et al. 2019).

Differences in E3-binding strategies are shown among the cognate complexes and, in the case of PDHc, even among species. E3 in principle may compete with the E1p and E1b components for the E2p-PSBD in bacteria and E2b-PSBD in most species, respectively: the binding sites in e.g., the *E. coli* E2p-PSBD or *Bacillus stearothermophilus* (Gram-positive bacterium) E2b-PSBD for the respective E1 and E3 components largely overlap (Chandrasekhar et al. 2013; Jung et al. 2003; Wynn et al. 1992; Brautigam et al. 2011); nevertheless, according to a newer model, the differential utilization of the three ecE2p chains (see also above) eliminates the need for competition between ecE1p and ecE3 (Song and Jordan 2012). According to HDX-MS experiments, in Gram-negative PDHcs, the E3 component interacts with the E2-PSBD *via* its FAD-binding and interface domains (Chandrasekhar et al. 2013). In contrast, the E3 subunit does not interact directly with E2 but rather with the E3BP-E3BD in eukaryotic PDHcs (Brautigam et al. 2006; Ciszak et al. 2006). As for the hE3-hE3BP interaction, the hE3BP binds a single hE3 dimer (see also above) (Prajapati et al. 2019); the binding site for the hE3BD proved to be along the twofold symmetry axis of the hE3 dimer, hence, both hE3 monomers participate, with in fact no apparent conformational change, in this interaction (Brautigam et al. 2006; Ciszak et al. 2006). According to X-ray crystallographic results, hE3 utilizes the same surface to assemble into the hBCKDHc (Brautigam et al. 2011). In the mammalian KGDHc (and hKADHc, unpublished data), E3 is recruited only after the E1 component binds to the E2 core (Rice et al. 1992; Nakano et al. 1994). The affinities of the E3 components for the cognate complexes in general are usually much smaller when compared to the respective E1 components (Ambrus et al. 2015b) (and references therein); however, the KGDH/KADH/BCKDH complexes are still functional with no dedicated E3 binding proteins present. Nevertheless, recently, Hevler et al. (2023) suggested that the MRPS36 protein could serve as a structural tether between the E2 core and the E3 component in the KGDHc.

Regardless of the differences in the subunit-subunit interactions, the core architectures of the cognate complexes appear to determine the chain stoichiometries as well as the spatial arrangements of the peripheral components within the complexes. Complexes with 24-meric cubic E2-core architectures [Gram negative PDHc (Mattevi et al. 1992a; Yang et al. 1985, 1986; Ambrus et al. 2015b), BCKDHc (Reed and Hackert 1990; AEvarsson et al. 2000), KGDHc (Wagenknecht et al. 1983, 1987; Nagy et al. 2021; Ambrus et al. 2015b), and KADHc (Nemeria et al. 2018)] usually tether 12 E1 components (dimers or heterotetramers) along the core-edges and six E3 dimers on the faces (see Appendix 3). Chain stoichiometry is not reported for the hKADHc, but it is likely identical to the one in hKGDHc, since these two complexes form hybrid complexes sharing hE2k and hE3 (see above). PDHcs in Gram-positive bacteria and eukaryotes operate with 60-meric pentagonal dodecahedron-shaped core structures (see above), with E3BP integrated into the structures (only for eukaryotes); the overall structures of the PDHcs here comprise 12-40 E1 heterotetramers and 6-20 E3 dimers, likely at the edges and faces, respectively (see Appendix 3). The differences in the conditions *in vitro* and *in vivo* may

also explain the mismatches in the measured data and the most likely structural assemblies (AEvarsson et al. 2000; Vijaykrishnan et al. 2011). There are also some interesting exceptions to these general stoichiometries (see Appendix 3 and the “Overall Structures of the Complexes” section; detailed discussion of these assemblies is beyond the scope of this chapter).

Structures of the Regulatory Components

Most eukaryotic/mammalian PDH and BCKDH complexes are regulated by phosphorylation (Reed et al. 1985). The post-translational modification takes place at the respective phosphorylation sites in the E1 α subunits. Phosphorylation by the BCKDH/PDH kinases (BDK/PDK) and dephosphorylation by the BCKDH/PDH phosphatases (BDP/PDP) lead to the inactivation and reactivation of the overall complexes, respectively (AEvarsson et al. 2000; Spriet and Heigenhauser 2002).

The high-resolution structures of the human and rat BDKs and their complexes with various ligands were recently determined (Machius et al. 2001; Tso et al. 2013, 2014; Liu et al. 2023; Roth Flach et al. 2023). The three-dimensional structures of all four PDKs and their complexes with ADP, ATP, and even lipoyl domain 2 (LD2) have also been determined by X-ray diffraction (Steussy et al. 2001; Kato et al. 2005; Tso et al. 2006; Devedjiev et al. 2007; Kato et al. 2007; Green et al. 2008; Wynn et al. 2008). The folded BDK and PDK chains comprise two domains each (regulatory and kinase domains) connected by flexible linker regions (Machius et al. 2001; Steussy et al. 2001). The two domains interact *via* two interfaces, one of which forms the active site (Liu et al. 2023; Tuganova et al. 2001). The nucleotide binding sites, characteristic to the GHKL superfamily (Dutta and Inouye 2000), reside in the kinase domains and face the active site clefts (Liu et al. 2023). In general, the PDK kinase domains are highly conserved, while the regulatory domains are not that much. The overall structures of the different PDK isoforms (PDK 1-4) are highly homologous (RMSD<1.5 Å); a noticeable difference exists between the PDK1 and the other PDK isoforms in the overall buildup of the ATP-binding pocket.

Both the BDK and PDK enzymes are present in dimeric forms; dimerization occurs *via* the respective kinase domains (Machius et al. 2001; Steussy et al. 2001). Dimerization triggers autophosphorylation at a Ser residue in BDK, the role of which is still rather obscure (Davie et al. 1995; Wynn et al. 2000). The BDK dimer possesses a tendency for further oligomer formation, whose functional role is presently unknown (Machius et al. 2001; Wynn et al. 2000). Ligand (ADP and pyruvate) induced tetramerization was also observed for PDK, which was assumed to contribute to the inhibition of the enzyme (Hiromasa et al. 2006). Nucleotide binding requires Mg²⁺ and K⁺ in BDK; the Mg²⁺ is coordinated by an Asn residue and the (nucleotide-)phosphate moieties, whereas the K⁺ is tethered by the α -(nucleotide-)phosphate as well as Val, Asp, Phe, and Gly residues (Machius et al. 2001). In contrast, while Mg²⁺ plays the same role in PDK, and its binding pocket is also rather similar to the one in BDK (Steussy et al. 2001), the K⁺ appears to have two binding

sites, one associated with the nucleotide and another one with the LD (Green et al. 2008).

The BDK enzyme also exhibits an ATPase activity (Wynn et al. 2000). The ADP, which is liberated by either activity of the BDK, inhibits the enzyme preventing constitutive ATP hydrolysis; inhibition by ADP is overwritten in the presence of the BCKDHc (Machius et al. 2001; Liu et al. 2023).

The hBDP is also known as the PP2Cm phosphatase (protein phosphatase 2C of mitochondrial (matrix) residence). Its wild-type and N94K mutant crystal structures were recently determined (Almo et al. 2007; Dolatabad et al. 2019). The active site is encapsulated in a cleft that primarily contains negatively charged residues and two metal ions (M1 and M2); in the absence of competing divalent metal ions, Mn^{2+} occupies both the positions. There is also a third metal ion position, which is occupied by Mg^{2+} . All these metal ions, directly or indirectly (*via* H_2O), are coordinated by Asp, Glu, Gly, and His residues (Wynn et al. 2012; Dolatabad et al. 2019; Almo et al. 2007). An Arg residue appears to be the catalytically most important residue as its guanidinium moiety can position the phosphoryl group of the substrate for dephosphorylation (Wynn et al. 2012).

The PDPs are members of the PPM family of serine/threonine phosphatases. They are Mg^{2+} -dependent enzymes and the activity of PDP1 is also heavily enhanced by Ca^{2+} because it is needed for the efficient binding to the inner LD in E2p (see below). The PDP1 comprises a regulatory (PDP1r) and a catalytic (PDP1c) subunit, while the PDP2 has only a catalytic subunit. PDP1r is a flavoprotein with a still not fully characterized regulatory role. Crystal structures of the catalytic domain in mammalian (rat and bovine) PDP1 were reported (Vassilyev and Symersky 2007; Guo et al. 2020). The catalytic subunits in the PDP1 and PDP2 are sequentially highly homologous and their catalytic sites are typical for the protein family (the PDP1c active site contains two metal ions (Mn^{2+}/Mg^{2+}), acidic (mainly Asp) residues and several specific/structural water molecules). There is a long insertion in the PDP1c sequence, which forms a highly hydrophobic pocket that most likely interacts with the LAM cofactor; it is considered to be the most important specificity feature of the enzyme.

Regulation

The succinyl- and glutaryl-CoA end products both inhibit the hKGDHc and hKADHc, which is consonant with the hybrid hKGDH-hKADH complex formation (Nemeria et al. 2018); glutaryl-CoA formation may also lead to inhibitory glutarylation at hE1a (Boyko et al. 2022). A delayed impact of KADHc inhibition on rat brain metabolism was also ascribed to protein glutarylation (Boyko et al. 2022). The respective acyl-CoA end products also inhibit the PDHc and BCKDHc in most species (Roche et al. 2001; Boyer and Odessey 1991; Parker and Randle 1978).

Vertebrate KGDH complexes are activated by Ca^{2+} ; the sensitivity of this phenomenon is enhanced upon low ATP/ADP ratios (Denton et al. 2016) ($>100 \mu\text{M}$ Ca^{2+} inhibits the rat brain KGDHc (Lai and Cooper 1986)). On the contrary, since Ca^{2+} -binding signatures are missing in the hE1a structure, the hKADHc is likely insensitive to Ca^{2+} (Nemeria et al. 2018). Ca^{2+} also activates the BCKDHc in rat liver mitochondria (Patel and Olson 1982; Hauschildt et al. 1982).

The high NADH/NAD⁺ ratio catalytically inactivates the E3 component *via* the four-electron reduced state of E3 that lacks LADH activity (Massey and Veeger 1961; Kikuchi 1973); NAD⁺ destabilizes the four-electron reduced state and reactivates E3. Regulation *via* the NADH/NAD⁺ ratio is indispensable where regulation *via* phosphorylation/dephosphorylation is not evolved (Kikuchi 1973; Wilkinson and Williams 1981; Moxley et al. 2014); this form of regulation is practically ubiquitous among all the α -keto acid dehydrogenase complexes.

Selected long-chain fatty acids are capable of inactivating the mammalian PDHc and BCKDHc (Roche et al. 2001; Tan et al. 2018; Supruniuk et al. 2023). ADP, but not AMP, stimulates the eukaryotic KGDHc, by lowering the K_m for KG, whereas ATP/GTP inhibits it (Sheu and Blass 1999; Strumilo 2005). It was also demonstrated that the hepatic rat BCKDHc is stimulated *via* insulin signaling (Shin et al. 2014).

In most eukaryotes, there is a phosphorylation-based regulatory mechanism for PDHc and BCKDHc; for instance, three and two specific Ser phosphorylation sites are present in highly conserved regions in the hE1 α and hE1 β subunits, respectively (Ciszak et al. 2003; Kato et al. 2008; AEvarsson et al. 2000; Li et al. 2004). Inactivating phosphorylation is governed by four (PDK1-4) and a single (BDK) kinase(s), as well as two (PDP1-2) and a single (BDP) phosphatase(s) in the above complexes, respectively; the different isoforms possess distinct site and tissue specificities, which leads to tissue-specific differences in the phosphorylation statuses of the respective E1 components. Kinase and phosphatase activities are the main factors behind the fine adjustment of the corresponding activities in a given tissue (Neinast et al. 2019; Bao et al. 2004; Kato et al. 2008; Kolobova et al. 2001; Korotchikina and Patel 1995; Suryawan et al. 1998; Zhou et al. 2012; Patel and Korotchikina 2001). In different E1 components, the phosphorylation order of the corresponding sites may also differ (Green et al. 2008); selected phosphorylation sites may even stay “silent” (Zhao et al. 1994).

The regulatory components generally bind to the highly flexible N-terminal arms that protrude from the complex cores. The interaction partners proved to be the inner LD of E2p, single LD of E2b, and E3BP for the PDK1-3, BDK, and PDK4, respectively (Kato et al. 2005; Roche et al. 2003; Machius et al. 2001); the hBDK and hBDP bind to the linker region between the LD and PSBD in hE2b, in close spatial vicinity to the phosphorylation loop in hE1b, according to (Chuang et al. 2002; Wynn et al. 2012). The LA cofactor also appeared to be generally important for the binding. In another study, LD2 and LD3, but not LD1, strongly interacted with PDK1 and PDK2 (Wang et al. 2015). Given that the overall folds of all the LDs are structurally highly similar, the source of specificity of the PDKs for the LDs is not obvious (perhaps only steric). PDP1 also binds to the inner LD of E2p (Damuni

and Reed 1987). The consequences of the phosphorylation were most thoroughly investigated in the BCKDHc (Li et al. 2004; Wynn et al. 2004).

Indirect regulation occurs *via* the regulatory components. ATP (substrate), acetyl-CoA, and NADH activate the PDKs, while pyruvate (in liver), ADP (product), CoA, and NAD⁺ inhibit them. Also, low ATP levels activate the PDPs (Huang et al. 2002; Kwon et al. 2004; Patel and Korotchikina 2006). PDP1 is only functional in the presence of Mg²⁺, while Ca²⁺ increases its affinity for the inner LD of E2p (Damuni and Reed 1987). In contrast, PDP2 is insensitive to Ca²⁺ (Karpova et al. 2003). Similarly to the connection between pyruvate and PDKs, branched-chain α -keto acids also inhibit BDK (Tso et al. 2013). Ca²⁺-stimulated inhibition of BDK by ThDP, the coenzyme for the E1 subunits, is also known (Noguchi et al. 2018). Insulin stimulates the PDP in rat white and brown adipose tissue mitochondria (Cooney et al. 1993).

PDHc is also regulated by SIRT4, a NAD⁺-dependent mammalian deacetylase enzyme, which is also involved in several other biological processes; SIRT4 removes the LA prosthetic group from the E2p component, which obviously leads to a dysfunctional PDHc. This action indirectly also affects the PDK and PDP, since their interactions with the enzyme complex become weaker (Mathias et al. 2014).

The PDHc is also regulated by acetylation at its E1 subunit and PDP (Fan et al. 2014).

Interestingly, oxaloacetate competitively inhibits ecE1k suggesting a “cross-talk” within the Krebs cycle (Frank et al. 2007).

Moonlighting Functions

The α -keto acid dehydrogenase complexes and their subcomplexes/components/subunits (especially dihydrolipoamide dehydrogenase) also have non-standard, so-called moonlighting functions. Many, but not all, of them are exclusively manifested in pathological states. They are also common in bacteria, where some of the intact α -keto acid dehydrogenase complexes are not even present or have less vital roles. The most important moonlighting functions are summarized in Appendix 4.

α -Keto Acid Dehydrogenase Complexes in the Crossroads of Cell Metabolism

The α -keto acid dehydrogenase complexes are at the crossroads of oxidative cell metabolism catalyzing oxidative decarboxylation of α -keto acids involved in many cellular processes (see above). However, substrates and products of these dehydrogenases are not only involved in metabolism but also in cell signaling, epigenetic regulation, and redox homeostasis (see below).

The KGDHc catalyzes a rate-limiting step in the TCA cycle providing NADH for the mitochondrial electron transfer system (ETS), hence, supporting mitochondrial oxidative phosphorylation and ATP production. Upon α -KG oxidation, succinyl-CoA is synthesized, which serves as a substrate for mitochondrial substrate-level phosphorylation (SLP) catalyzed by succinyl-CoA ligase (SCL); the latter enzyme has essential roles upon the impairment of the ETS (Heldt and Schwalbach 1967). Irreversible, non-enzymatic succination of selected Cys residues in E2k by fumarate hampered the overall KGDHc activity in a Complex I deficiency model; this observed phenomenon exacerbated the mitochondrial ATP deficit in this model system (Piroli et al. 2023). An ETS damage is associated with decreased ATP production leading to utilization of cellular ATP *via* reversal of the adenine nucleotide translocase (ANT) and ATP synthase in order to maintain the mitochondrial membrane potential (Chinopoulos 2011b). SCL, however, is capable of inhibiting the reversal of ANT *via* mitochondrial SLP (Chinopoulos 2011a; Chinopoulos and Seyfried 2018). This has special importance in cancer cells like glioblastomas with inhibited PDHc and pyruvate carboxylase (PC) activities; glioblastomas are usually exposed to hypoxia and rely on enhanced glutaminolysis (Chinopoulos and Seyfried 2018). This latter process supports succinyl-CoA formation *via* the glutamine-glutamate-KG axis contributing to cancer cell survival. A substantially decreased provision of succinyl-CoA, due to a low KGDHc activity, leads to a compromised SLP contributing to mitochondrial impairments in diseases with low KGDHc activities (Kiss et al. 2013). KG is involved in 42 mitochondrial and 72 cytosolic reactions in humans according to the Metabolic Atlas (<https://metabolicatlas.org/explore/Mouse-GEM/gem-browser/metabolite/MAM01306m>; <https://metabolicatlas.org/explore/Mouse-GEM/gem-browser/metabolite/MAM01306c>). These reactions include the oxidative decarboxylation of α -keto acids, transamination of amino acids, amino acid metabolism, and transport processes, among others. KG is also a co-substrate of the KG-dependent dioxygenases, which are associated with epigenetic regulations like DNA and histone methylation, and adaptation to hypoxia (Losman et al. 2020). Furthermore, succinyl-CoA is a source of protein, DNA, and histone succinylation (for post-translational modifications, see below).

The PDHc is a link between glycolysis and the mitochondrial oxidative metabolism as providing acetyl-CoA to the TCA cycle. Besides this, acetyl-CoA is also utilized for (i) lipid (fatty acid and cholesterol) syntheses after leaving the mitochondria in the form of citrate, (ii) mitochondrial ketone body formation, and (iii) protein/histone/DNA acetylation (in the mitochondria, cytosol, and nucleus). Pyruvate is involved in 16 mitochondrial and 42 cytosolic reactions in humans according to the Metabolic Atlas (<https://metabolicatlas.org/explore/Human-GEM/gem-browser/metabolite/MAM02819m>; <https://metabolicatlas.org/explore/Human-GEM/gem-browser/metabolite/MAM02819c>).

Branched-chain amino acids (BCAAs) Val, Leu, and Ile represent ~20% of the human dietary amino acids (Brosnan and Brosnan 2006) and represent ~30–35% of the essential amino acid intake. During their catabolism, valine is entirely glucogenic, isoleucine is partially glucogenic and ketogenic, whereas leucine is entirely

ketogenic. The first step in BCAA catabolism is transamination, followed by decarboxylation; the latter reaction is the rate-limiting step in BCAA catabolism (Xu et al. 2020). The relevant physiological aspects are covered in a recent review (Neinast et al. 2019).

BCAAs are involved in key cellular reactions like protein turnover regulation, autophagy signaling, mitochondrial function, and neurotransmitter metabolism. Leu has a primary role in sensing amino acid availability for protein synthesis. Leu can activate the mTORC1 pathway *via* binding and inactivating Sestrin2, an inhibitor of mTOR activity, indicating that the amino acid pool is available for protein synthesis. In the absence of Leu, mTORC1 is inactive, and protein synthesis is slow (Wolfson et al. 2016).

BCAAs and their catabolic products play roles in regulating autophagy. The low intracellular levels of amino acids, especially that of Leu, and the elevated concentrations of uncharged tRNAs activate GCN2, a serine-threonine protein kinase. The activated GCN2 phosphorylates the eukaryotic initiation factor eIF2 α . Inhibition of eIF2 α leads to the general repression of protein synthesis, but the translation of the ATF4 protein stimulates the transcription of genes responsible for autophagy (Moro et al. 2016).

During physical exercise, increased rates of BCAA oxidation could be observed. Considering that the BCKDHC is the rate-limiting step in the catabolism of these amino acids, it seemed plausible that the overall BCKDHC activity is somehow upregulated. The short-term upregulation could eventually be connected to the dephosphorylation/phosphorylation of the enzyme (Shimomura et al. 1995), while the long-term upregulation (of the transaminases and the BCKDHC) could be attributed to the PGC-1 α -stimulated transcription enhancement of the respective genes (Hatazawa et al. 2014).

A compromised activity of the BCKDHC leads to the accumulation of the respective BCAAs and α -keto acids (see above) and consequently also to a disease referred to as maple syrup urine disease (MSUD; see below); MSUD is associated with neurological dysfunction and brain damage (Dancis et al. 1959). The heterogeneous tissue distribution of the BCKDHC activity permits an interorgan metabolism of the proteinogenic branched-chain amino acids, which has broad implications in the pathophysiology of various diseases (Suryawan et al. 1998; Patrick et al. 2022; Xu et al. 2023).

KADHC is involved in the catabolism of L-Trp, L-Lys, and L-hydroxylysine and is responsible for the conversion of α -ketoadipate to glutaryl-CoA generating NADH. Silencing of the *DHTKD1* gene that encodes the hE1a protein augmented ROS formation, mitochondrial content, mtDNA content, and mitochondrial membrane potential, while it impaired mitochondrial biogenesis and the mitochondrial cristae structure leading to damaged cell growth and enhanced apoptosis (Wang et al. 2021; Bunik and Degtyarev 2008; Danhauser et al. 2012; Xu et al. 2013). These findings highlight the critical role of KADHC in mitochondrial energy production. Inhibition of hE1a by phosphonate analogues resulted in significant differences in sugar metabolism metabolite levels (Artiukhov et al. 2020). Mutations of hE1a are also frequently observed in various cancers (Bell et al. 2011).

Tissue Distribution of the α -Keto Acid Dehydrogenase Complexes

The KGDHc is a rate-limiting enzyme in the TCA cycle; therefore, it is ubiquitously expressed in almost every tissue containing mitochondria. However, there are tissue-specific differences in enzyme activities. The distribution of the KGDHc in different brain regions and cell types was recently reviewed in detail by Hansen and Gibson (2022). Interestingly, human astrocytes did not express the hE1k and hE2k subunits or simply showed very low immunoreactivity for them, also reflected in a low overall KGDHc activity relative to the rodent astrocytes (Dobolyi et al. 2020; Ko et al. 2001). In accord with this, the KGDHc activity was three times higher in astrocytes than in neurons in rodents (Park et al. 2000).

The PDHc is ubiquitously expressed in every tissue; however, its activity displays a heterogeneous pattern due to the tissue-specific distributions of the PDK isoforms (Klyuyeva et al. 2019). PDK 1 and 2 were ubiquitously expressed in mouse tissues with the highest protein levels and activities in the heart. PDK3 was expressed in the heart, brain, and testes, while PDK4, at the protein level, was observed in the heart and diaphragm. PDK3 was primarily responsible for PDH regulation in the kidneys and brain. The protein levels of PDK4 were low in the well-fed state. Due to the tissue-specific inhibition of PDH *via* PDK, the PDHc activity was low in astrocytes, rather favoring the pyruvate-lactate conversion that supports the neuron-glial lactate shuttle (Halim et al. 2010; Bouzier-Sore et al. 2006).

Sequential enzymes can associate to form supramolecular complexes, often called metabolons. Metabolon formation by the mitochondrial branched-chain aminotransferase (BCATm) and the BCKDHc has been demonstrated in the liver. The mitochondrial transaminase is associated with the BCKDHc-E1 α subunit, and this association increased the k_{cat} in the BCKDHc-E1 component by a factor of 12 (Islam et al. 2007). In other organs, the two steps (transamination and oxidative decarboxylation) can be separated, and this shows a remarkable tissue specificity. The highest transaminase activity was found in the striated muscles, kidneys, brain, and heart (Harper et al. 1984; Hutson et al. 2005). The oxidation of the formed keto acids is the most intensive in the liver, intermediate level activity of the BCKDHc is displayed in the kidneys and heart, and low activity is found in the adipose tissue, muscle, and brain (Harper et al. 1984; Dimou et al. 2022; Harris et al. 1997; Sweatt et al. 2004; Suryawan et al. 1998). Transaminases in the brain show cell-specific compartmentation with cytosolic expression in neurons and mitochondrial localization in the astrocytes (Xing et al. 2018). Controversial reports were published on the cell-specific expression of the BCKDHc in the brain. According to one view, the BCKDHc is found in neuronal mitochondria; therefore, the keto acids produced in astrocytes have to be transported into the neurons for completing their breakdown (Cole et al. 2012). However, earlier studies on astroglia-rich primary cultures showed complete breakdown of Val, Ile, and Leu (Bixel et al. 2001; Murin et al. 2009a, b).

Role of the α -Keto Acid Dehydrogenase Complexes in Redox Homeostasis

The reactive oxygen species (ROS) are inevitable side products of aerobic respiration. Imbalance between the production and elimination of ROS leads to oxidative stress, which is associated with protein, nucleotide, and lipid damage. However, ROS also play important roles in cellular signaling in health and disease, as was recently reviewed by (Sies and Jones 2020) and (Brand 2016). Mitochondria are one of the major sources of the cellular ROS (Gorlach et al. 2015). In this cellular compartment, several ROS-producing sites can be distinguished, like the α -keto acid dehydrogenase complexes (Tretter and Adam-Vizi 2004; Starkov et al. 2004b; Andreyev et al. 2005, 2015; Brand 2016; Fisher-Wellman et al. 2013). The isolated, intact KGDHc is capable of producing superoxide and, by its spontaneous dismutation, H_2O_2 in both the physiological and reverse directions of the catalytic reaction (Tretter and Adam-Vizi 2004; Ambrus et al. 2015b). NAD^+ inhibits the ROS production in both directions and switches the enzyme back to the physiological catalytic mode (Tretter and Adam-Vizi 2004; Ambrus et al. 2015b). In the reverse catalytic direction, the isolated KGDHc generated ROS after NADH addition in the absence of ThDP and CoA. In the presence of substrates and cofactors, the rate of ROS production augmented with the $NADH/NAD^+$ ratio for the KGDHc, BCKDHc, and PDHc (Tretter and Adam-Vizi 2004; Starkov et al. 2004a; Quinlan et al. 2014); ROS production in these complexes was ascribed to the common E3 subunit (Tretter and Adam-Vizi 2004; Quinlan et al. 2014; Tahara et al. 2007; Ambrus and Adam-Vizi 2013; Ambrus et al. 2009). Notably, in isolated mitochondria, ROS production by the above complexes is also biased by the Complex I-initiated ROS formation; therefore, specific inhibitors must be utilized to measure enzyme-selective ROS. Under conditions favoring the highest rate of ROS generation by each enzyme, the KGDHc doubled the rate by PDHc and increased four times the rate by the BCKDHc in isolated skeletal muscle mitochondria (Quinlan et al. 2014). ROS-production property of the flavin moiety of the KADHc has been identified in rat skeletal muscle mitochondria (Goncalves et al. 2016); ROS formation augmented with the $NADH/NAD^+$ ratio in the presence of α -ketoadipate in the forward catalytic direction. Importantly, the KGDHc and PDHc are also targets of excess ROS, and their inhibition may lead to metabolic deficiency in mitochondria (Tretter and Adam-Vizi 2005). ROS inhibits the E1 subunits in these enzymes leading to overall declines in the enzymatic activities (Bunik and Sievers 2002). Importantly, thiol radicals produced at the E3 subunits of the KGDHc and PDHc are more reactive than superoxide/ H_2O_2 and able to inhibit the E1 subunits (thioredoxin can alleviate this effect) (Bunik and Sievers 2002). These complexes can also be inactivated by other oxidants such as the lipid peroxidation product 4-hydroxy-2-nonenal, which reacts with the lipoic acid cofactor (Humphries and Szweda 1998). A H_2O_2 -mediated reversible glutathionylation of the LA cofactor in the KGDHc protects the enzyme from oxidant-mediated irreversible inactivation; this procedure appears to be an antioxidant mechanism protecting the enzyme. After the removal of oxidants,

glutaredoxin removes glutathione from the enzyme leading to its reactivation (Humphries and Szweda 1998; Nulton-Persson et al. 2003; McLain et al. 2011).

Regulation of the α -Keto Acid Dehydrogenase Complexes by Calcium

Calcium is a general activator of the mitochondrial dehydrogenases; therefore, it has key roles in controlling the oxidative metabolism including ROS production (Denton 2009; Gherardi et al. 2020). Direct Ca^{2+} -induced activation of the KGDHc and indirect Ca^{2+} -dependent stimulation of the PDHc *via* PDP have been described in many organs (Tretter and Adam-Vizi 2004; Adam-Vizi and Starkov 2010; Chinopoulos and Adam-Vizi 2006; Starkov et al. 2004a). The Ca^{2+} -transfer from the endoplasmic reticulum (ER) to mitochondria supports the KGDHc activity, thus, inhibition of this process leads to reduction in the KGDHc activity (Cardenas et al. 2020); the latter process leads to an increased NAD^+ level that promotes autophagy mediated by NAD^+ -dependent sirtuins in normal and OXPHOS-deficient cells (Cardenas et al. 2020). Therefore, inhibition of the ER-mitochondrial Ca^{2+} -transport might prove to be a valid target in cancer therapy. In the skeletal muscle, Ca^{2+} increases the sensitivity of BDK for ThDP leading to the inhibition of the enzyme and in turn activation of the BCKDHc, which is important under conditions when enhanced BCAA oxidation is required, e.g., during physical exercise (Noguchi et al. 2018).

KGDHc and PDHc in the Nucleus

Translocation of the mitochondrial dehydrogenases, except for succinate dehydrogenase (Liu et al. 2021), to the nucleus is a special form of the mitochondrion-to-nucleus communication (Kafkia et al. 2022; Liu et al. 2021; Nagaraj et al. 2017; Srivastava et al. 2023; Zervopoulos et al. 2022; Ng and Tang 2014); these processes may facilitate the production of intermediates for epigenetic regulation of transcription, which can lead, e.g., to tumor formation (Morrison 2022; Tran et al. 2017). The nuclear KGDHc binds to lysine acetyltransferase 2A (KAT2A) that also has a succinyltransferase activity and a binding site for succinyl-CoA; this complex acts as a histone H3 succinyltransferase (Wang et al. 2017) (see Appendix 4). The nuclear PDHc facilitates histone acetylation according to the cell cycle, in response to epidermal growth factor or mitochondrial dysfunction (Sutendra et al. 2014); translocation to the nucleus lowers the mitochondrial PDHc activity.

Roles of the α -Keto Acid Dehydrogenase Complexes in Various Diseases

Role of the α -Keto Acid Dehydrogenase Complexes in Neurodegenerative Diseases

In Alzheimer's disease (AD), both the PDHc and KGDHc activities decline (Mastrogiamco et al. 1993; Gibson et al. 1998; Blass and Gibson 1991; Bubber et al. 2005; Sorbi et al. 1983; Gibson et al. 2012) (see Appendix 1); these activity reductions were attributed to thiamine loss, enzyme mutations, or downregulation (Gibson et al. 1998; Mastrogiamco et al. 1993; Jia et al. 2023). In the presence of thiamine, increased KGDHc activities could be measured in postmortem brain slices from AD patients (Mastrogiamco et al. 1993). Overexpression of the KGDH-like/E1k-like (KGDHL/OGDHL) subunit in 3xTg-AD mice ameliorated the learning and memory deficits as well as reduced the neuroinflammation, A β deposition, and tau phosphorylation *via* activation of Wnt/ β -catenin signaling (Yao et al. 2022). The KGDHL/OGDHL protein is a genetic variant of E1k found mostly in the brain and liver (Bunik and Degtyarev 2008; Bunik et al. 2008); it might be a tumor suppressor, since its downregulation led to enhanced tumor formation (Chang et al. 2022). Recently, decreased KGDHc and succinyl-CoA ligase activities were detected in AD mutant human-induced pluripotent stem cell-derived cerebrocortical neurons; decreased activity of KGDHc was explained by protein S-nitrosylation of the E1 and E3 subunits in the KGDHc in these samples and in postmortem human AD brains (Andreyev et al. 2023). Decreased KGDHc protein levels that correlated with the neuronal degeneration were reported in Parkinson's disease (Mizuno et al. 1994). In a recent study, a decreased level of the E1p α protein was observed in the putamen and substantia nigra of patients with idiopathic Parkinson's disease (Miki et al. 2017). Low PDHc and KGDHc activities in patient fibroblasts were reported in Friedrich's ataxia (FA) (Blass et al. 1976; Melancon et al. 1979); reduced E3 protein levels (Mastrogiamco et al. 1996) or activities (Kark and Rodriguez-Budelli 1979; Filla et al. 1978) were also observed in this disease. Recently, decreased mRNA levels for E1p were also reported in peripheral blood mononuclear cells from FA patients (Pathak et al. 2019). Compromised KGDHc and PDHc activities were also reported in Huntington's disease (Sorbi et al. 1983; Butterworth et al. 1985; Naseri et al. 2015); it has been described that histone deacetylase inhibitors also inhibit PDK expression leading to activation of the PDHc, a phenomenon that might be therapeutically valuable in Huntington's disease (Naia et al. 2017). Reduced KGDHc activities were also reported in other diseases affecting the central nervous system, like head trauma after skull, scalp, or brain injury (Pandya et al. 2023), spinal cord injury (Boyko et al. 2021), and progressive supranuclear palsy (Park et al. 2001; Albers et al. 2000). PDHc was also implicated in epilepsy (Bhandary and Aguan 2015). Suboptimal KGDHc or PDHc/PDP activities often get accompanied by lactic acidosis (Guffon

et al. 1993; Bonnefont et al. 1992; Brown 1992; Robinson and Sherwood 1975); if pyruvate cannot enter the TCA cycle, the pyruvate-lactate conversion may result in extreme high lactate concentrations.

KGDHc in Cancer

Various metabolic pathways converge in the mitochondria in which KGDHc plays a pivotal role following the reprogramming signals of cancer cells and modulating the epigenetic landscape in cancer (Vatrinet et al. 2017).

E2k, and a functional TCA cycle, were implicated in MYC-induced leukemogenesis showing the metabolic dependence of T-lymphoblasts on the TCA cycle (Anderson et al. 2016); these results point to potential new drug targets in cancer therapy.

α KG accumulation, due either to KGDHc suppression or exogenous KG supplementation, gave rise to aspartate depletion and dysregulation of the malate-aspartate shuttle in phosphatidylinositol-3-kinase-deficient cancer cells (Ilic et al. 2017); this leads to a dysregulation in glucose metabolism and inability to sustain the NAD^+ /NADH redox balance in these tumors. It also proves that suppression of the KGDHc increases the vulnerability of selected cancer cells and perhaps suppresses tumor growth (Chang et al. 2022).

In several types of cancers, including neuroblastoma and triple-negative cancer cells, downregulation of E2k suppressed tumor growth (Anderson et al. 2016, 2021; Shen et al. 2021). The suppression of E3 attenuated glutaminolysis-induced ferroptosis *via* reduction of ROS production in head and neck cancer (Shin et al. 2020), while accumulation of KG induced ROS generation leading to ferroptosis in diffuse large B-cell lymphoma, which led to suppression of tumor growth (Cai et al. 2023). According to the literature, increased KGDHc activity is rather oncogenic, whereas KGDHL acts as a tumor suppressor (Chang et al. 2022). Therefore, the KGDHc could potentially be a valid drug target in selected KGDHc-sensitive cancer types. It is important to note that inhibition of the KGDHc in tumors bearing isocitrate dehydrogenase 1/2, malate dehydrogenase, or fumarate hydratase mutations may lead to the accumulation of certain oncometabolites, like fumarate or 2-hydroxyglutarate, thus targeting the KGDHc could only be useful in well-selected tumors, which require further investigation.

PDHc in Cancer

The PDHc and its regulatory enzymes, PDK and PDP, are gatekeeper enzymes between glycolysis and the mitochondrial oxidative phosphorylation, which constitutes them as potential drug targets in cancer therapy (Stacpoole 2017; Schoenmann et al. 2023; Anwar et al. 2021). Stimulation of the PDHc activity *via* inhibition of PDK could in principle divert metabolism toward oxidative phosphorylation, which

would disfavor the proliferation of glycolytic cancer cells; also, upregulation/induction of PDKs were found in selected cancers (as well as in starvation and diabetes) (Papandreou et al. 2006; Kim et al. 2006; Kaplon et al. 2013; Michelakis et al. 2010; Lu et al. 2008; Hitosugi et al. 2011). Enhanced PDHc activity may also induce cell death *via* augmented ROS production (Stacpoole 2017; Sradhanjali and Reddy 2018; Woolbright et al. 2019). Additionally, nuclear PDHc governs protein/histone acetylations in cancer cell signaling pathways as well (Guertin and Wellen 2023). It was also recently reported that hindrance of the PDHc activity—*via* inhibition of PDP—hampered histone acetylation which suspended the HIF-1 α pathway and gene expression; thus, inhibition of PDP might also be a feasible target in cancer therapy (Karagiota et al. 2023).

BCAAs and Diabetes

BCAA metabolism also plays a role in the insulin-glucose homeostasis. The BCKDHc is stimulated *via* insulin-mediated signal transduction. Under normal (non-diabetic) conditions, glucose administration decreases the blood concentration of BCAA (Felig et al. 1969). This effect is attributed to the insulin-dependent stimulation of the BCKDHc activity. Similarly, BCKDHc deficiency results in insulin resistance (Yu et al. 2023). In diabetic patients, increased BCAA concentrations can be measured (Holecek 2023). In accord with this observation, elevated BCAA concentrations have a predictive value for future insulin resistance (McCormack et al. 2013). Similarly, deprivation of individual BCAAs improved insulin sensitivity in mice (Xiao et al. 2014). Inhibition of BDK improved insulin sensitivity and reversed the high-fat-induced steatosis. The authors claimed that reduced branched-chain α -keto acid (BCKA) levels can be interpreted as “fasting signals” (Bollinger et al. 2022).

BCAAs and Cancer

Considering that BCAAs (Leu) stimulate protein synthesis *via* activating mTORC1, stimulating the catabolism of BCAAs could be useful in cancer chemotherapy. In ovarian and breast cancer cell lines, BDK inhibitors increased the sensitivity of the tumor cells to chemotherapy (Ibrahim et al. 2023).

BCAAs and Heart Failure

Accumulation of BCAAs can result in cardiac remodeling and consequent heart failure (Voronova et al. 2022). Elevation of the BCKA concentrations can induce resistance to insulin in the heart combined with heart failure (Uddin et al. 2019). A BDK

inhibitor successfully lowered the BCKA concentrations in the heart and plasma (Tso et al. 2013) and improved the cardiac metabolism (Roth Flach et al. 2023).

Mice with Deficient KGDHc

Application of genetically modified mice strains with reduced KGDHc activities is a powerful approach to study the role of the KGDHc in mitochondrial or cellular metabolism and disease models. Mice having a heterozygous mutation in the gene for E2k (*DLST*^{+/-}) or the gene for E3 (*DLD*^{+/-}) have been utilized to study the role of the KGDHc in mitochondrial bioenergetics, oxygen consumption, and ROS generation (Starkov et al. 2004a; Yang et al. 2009; Horvath et al. 2022); the respective homozygous knockout animals are nonviable, which provided in vivo evidence for the requirement of the mitochondrial oxidative pathway during embryonal development (Johnson et al. 1997; Yang et al. 2009). It was reported that the *DLD*^{+/-} mice exhibited a ~30% reduction in the KGDHc activity and a ~20% reduction in the PDHc activity (Klivenyi et al. 2004). Similarly, the total brain of *DLST*^{+/-} mice displayed a ~40% decline in the KGDHc activity, whereas the activities for the PDHc and other mitochondrial enzymes, like malate dehydrogenase, were not altered (Yang et al. 2009). In mitochondria isolated from brain of mice having *DLST*^{+/-}, the O₂ consumption decreased compared to the wild-type mice when using KG or glutamate *plus* malate (GM) as respiratory fuel substrates, which enter the TCA cycle before the step catalyzed by the KGDHc (Starkov et al. 2004b; Horvath et al. 2022). Additionally, mitochondria isolated from mice brain with *DLD*^{+/-} showed lower O₂ consumption not only with KG and GM but also with pyruvate *plus* malate, demonstrating that DLD is a component of the PDHc (Starkov et al. 2004b). Furthermore, the mitochondrial O₂ consumption supported with succinate did not change in either the *DLST*^{+/-} or the *DLD*^{+/-} animal; succinate fuels the TCA cycle after the step catalyzed by the KGDHc. In accord with this, the mitochondrial membrane potential also decreased in *DLD*^{+/-} mice brain mitochondria respiring on KG. Although bioenergetic impairment was observed in situ in mitochondria isolated from either the *DLST*^{+/-} or *DLD*^{+/-} mice, both strains possessed normal phenotypes (Johnson et al. 1997; Yang et al. 2009). Furthermore, extensive histological study of the *DLST*^{+/-} mice brains showed normal brain morphology and development. Importantly, the *DLST*^{+/-} animals were more vulnerable to mitochondrial neurotoxins like 3-nitropropionic acid, which was used in Parkinson's disease models or malonate and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), which were utilized to mimic Huntington's disease (Klivenyi et al. 2004; Yang et al. 2009). These findings suggest that reduction in the KGDHc activity might increase the susceptibility to neurodegenerative diseases or exacerbate the symptoms/pathogenesises of neurological disorders.

It is well-known that KGDHc, among several other mitochondrial enzymes, is an important source of ROS under specific conditions when the NADH/NAD⁺ ratio is high (Starkov et al. 2004b; Tretter and Adam-Vizi 2004) (see above).

Studies using mitochondria isolated from brains of $DLD^{+/-}$ or $DLST^{+/-}$ mice demonstrated a decreased H_2O_2 production rate relative to control when using KG as respiratory substrate (Starkov et al. 2004b; Horvath et al. 2022). This finding suggests that all the three subunits of the KGDHc contribute to ROS generation in the normal/physiological direction of the overall catalytic reaction. Moreover, the succinate- or α -glycerophosphate (α -GP)-evoked H_2O_2 formation also decelerated in brain mitochondria from $DLST^{+/-}$ and $DLD^{+/-}$ mice under conditions favoring reverse electron transfer (RET), relative to control (Horvath et al. 2022); RET is a special mechanism for ROS production: in the course of succinate or α -GP oxidation, the electrons are channeled backwards from Complex II or α -glycerophosphate dehydrogenase (GPDH), *via* the Q-junction, to Complex I, where the FMN subunit might be responsible for ROS generation (Liu et al. 2002) (this process is characterized with a high rate of ROS formation and stimulated NADH production). Based on earlier studies, the KGDHc is potentially capable of directly binding to Complex I; hence, the E3 subunit might be able to instantaneously oxidize the NADH generated at Complex I upon RET (Sumegi and Srere 1984). It has been suggested that RET contributes to the ischemia-reperfusion injury (Chouchani et al. 2014; Zhang et al. 2018). Ischemia is associated with accumulation of succinate due to the reversal of succinate dehydrogenase or because of the canonical operation of the Krebs cycle. During reperfusion, the accumulated succinate is oxidized and contributes to a high ROS production rate *via* RET leading to oxidative stress. According to Ben-Yoseph et al. (1993) and Nguyen et al. (2007), α -GP might also accumulate in the ischemic rat brain, contributing to mitochondrial and cellular injury *via* RET-induced ROS production. The fact that the RET-evoked ROS generation proved to be sensitive to the heterozygous knock-out of *DLST* or *DLD* suggests that the KGDHc might play roles in ischemia-reperfusion injury (see Fig. 13.5).

α -Keto Acid Dehydrogenase Complex Deficiencies

As was seen above, the α -keto acid dehydrogenase complexes are crucial in human physiology; therefore, their deficiencies generally lead to severe clinical consequences. Isolated deficiency of an α -keto acid dehydrogenase complex develops upon the dysfunction of one of the two specific components (E1 or E2). Since E3 is a common subunit, its malfunction simultaneously, but not necessarily equally, affects all the complexes; the glycine cleavage system also contains a dihydrolipoamide dehydrogenase subunit, however, the functionality of this complex is rarely affected in human E3 deficiency (Quinonez and Thoene 2014; Ambrus and Adam-Vizi 2018). Besides the loss of function, dissociation of the complexes—due to selected pathogenic mutants—might also give rise to moonlighting catalytic activities, like ROS generation (see above), which can potentially also exacerbate

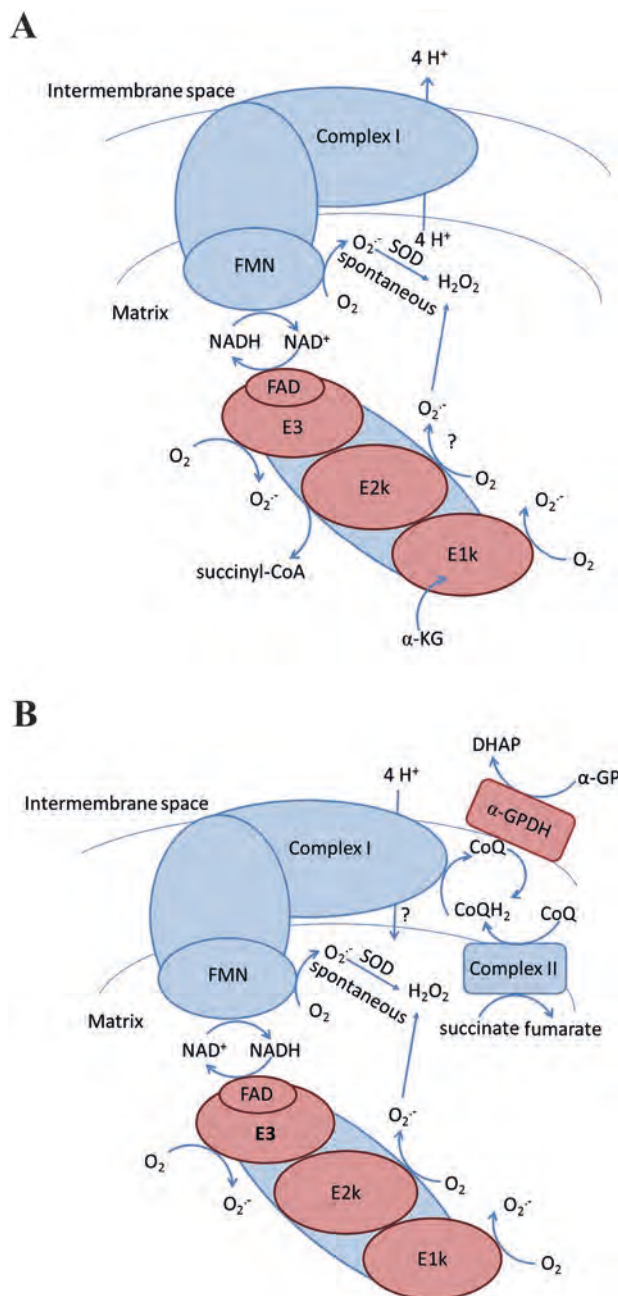


Fig. 13.5 Role of the KGDHc in ROS production observed during forward or reverse electron transfer in ischemia-reperfusion injury. **(a)** Forward electron flow with α -ketoglutarate. **(b)** Reverse electron flow with succinate or α -glycerophosphate. α -KG α -ketoglutarate, $E1\alpha$ α -ketoglutarate dehydrogenase, $E2\alpha$ lipoamide succinyltransferase, $E3$ lipoamide dehydrogenase, FAD flavin adenine dinucleotide, FMN flavin adenine mononucleotide, SOD superoxide dismutase, CoQ oxidized Coenzyme Q, $CoQH2$ reduced Coenzyme Q, α -GP α -glycerophosphate, α -GPDH α -glycerophosphate dehydrogenase, $DHAP$ dihydroxyacetone phosphate

the clinical outcomes. Also, mutant complexes and the respective subunits often degrade faster.

Below, we summarize the most important aspects of the α -keto acid dehydrogenase complex deficiencies.

Pyruvate Dehydrogenase Complex Deficiency

There are two major types of PDHc deficiencies: primary and secondary. Primary PDHc deficiencies are generally caused by mutations, whereas the secondary deficiencies are the results of vitamin deficiencies, which lead to insufficient cofactor levels; the latter cases most often imply thiamine deficiency, which may arise due to lack of intake, transport dysfunctions, genetic conditions, alcoholism (because of the vitamin-free “empty calories” of distilled spirits that dominate calorie intake), or chronic diarrhea (see also the so-called beriberi disease). The most common primary PDHc deficiency is the result of mutations in the *PDHA1* gene that encodes the E1 α subunit; this gene is localized on the X chromosome (Sperl et al. 2015) and missense/nonsense mutations, deletions, and insertions have all been reported for it (see below). Other relevant genes prone to have disease-causing mutations include the *PDHB*, *DLAT*, *DLD*, and *PDHX* that encode the E1 β , E2p, E3, and E3BP subunits, respectively (Sperl et al. 2015).

Since this enzyme complex catalyzes one of most important metabolic steps, mutations causing total loss of function often lead to embryonic fatality. Still the levels of clinical severity in the PDHc deficiency cases can be quite distinct, which is also the basis of classification—neonatal, infantile, and benign forms exist (Brown et al. 1994; Robinson et al. 2001; Wexler et al. 1997). Since the brain preferentially satisfies its high energy demand *via* the aerobic oxidation pathway, neurological symptoms and developmental abnormalities dominate the clinical phenotypes in all types of PDHc deficiency. These include severe neonatal lactic acidosis, progressive neurological and neuromuscular degradation, hypotonia, seizures, ataxia, respiratory distress, facial dysmorphism, microcephaly, ventricular enlargement, periventricular leukomalacia, peripheral neuropathy, and brain abnormalities like Leigh syndrome (Patel et al. 2012). The metabolic derangements can also give rise to ROS generation that may cause oxidative stress; PDKs, PDK2 in particular, is readily inactivated by ROS further modulating the PDHc activity (Hurd et al. 2012). Since the above clinical phenotypes often overlap with those observed in many other genetic mitochondrial disorders, the early and differential diagnosis of PDHc deficiency is usually difficult (Brown et al. 1994). High plasma pyruvate and lactate levels—with the latter in much domination often leading to metabolic acidosis—are the most important and typical signs for this disease state.

Below, we discuss a few representative cases for *PDHA1* gene alterations. A 4- and a 7-bp deletion were documented by (Endo et al. 1989) and (Dahl et al. 1990), respectively; clinical severity in female patients was largely dependent on

the X-chromosome inactivation patterns in the brain (see also below). Twenty mutations (deletions, insertions, point mutations) have been documented by Dahl et al. (1992); 12 of these mutations resided in exons 10 and 11 (this gene contains 11 exons). Young male probands presenting as acute Leigh syndrome patients bore the D258A and R263G mutations; the latter substitution lowered the affinity for ThDP (Matthews et al. 1993; Naito et al. 1997). The sex ratio in hE1 α deficiency proved to be ~1:1; however, more often than not, the missense mutations affected males, while deletions and insertions were found in females; it was hypothesized that insertions/deletions may cause intrauterine death in hemizygous males (Matthews et al. 1994). In monozygous female twins with differing disease severities, heterozygous mutations for the same H367L substitution were observed; X-inactivation analysis showed ~75:25 and ~50:50 inactivations, while the overall PDHc activity and the hE1 α protein level were both reduced, relative to the controls, in the fibroblast samples of the two patients (Horga et al. 2019). It was concluded here again that the X inactivation pattern affects the phenotypic expression in heterozygous females.

α -Ketoglutarate Dehydrogenase Complex Deficiency

hE1k deficiency is a disorder with autosomal recessive inheritance. It clinically features infantile- and pediatric-onset basal ganglia-associated movement disorders, hypotonia, developmental delays, ataxia, metabolic acidosis, microcephaly, and seizures (Kohlschütter et al. 1982; Guffon et al. 1993; Yap et al. 2021; Whittle et al. 2023). In its etiology, four homozygous mutations in the oxoglutarate (α -ketoglutarate) dehydrogenase gene (*OGDH*) on chromosome 7p13 have been reported. One of the three resultant amino acid substitutions (N320S) affected a conserved site, which is required for ThDP binding in the ThDP-binding domain; reduced hE1k protein expression and activity were also observed (Yap et al. 2021). The S297Y and P189L variants proved to be unstable protein products (Whittle et al. 2023). The c.935G-A transition resulted in a splicing defect with in-frame skipping of exon 7 (Whittle et al. 2023).

For hE2k deficiency (gene: *DLST*), an autosomal dominant inheritance was observed. In four unrelated probands with pheochromocytoma/paraganglioma syndrome-7 (PPGL7), a germline heterozygous missense mutation (G374E, c.1121G-A transition in exon 14) was identified; the G374E variant displayed compromised catalytic activity relative to control in vitro and resulted in abnormal accumulation of the oncometabolite 2-hydroxyglutaric acid (Remacha et al. 2019). Further in vivo studies revealed that *DLST* can apparently also act as a tumor suppressor gene; not unexpectedly, TCA cycle abnormalities occurred when playing its roles in the pathogenesis of paragangliomas. Additional such patients carried the heterozygous mutations responsible for the R231Q, D304N, and Y422C protein variants, whose in vitro functional analyses showed no alterations from the control; it was not

excluded, however, that these variants could still exert other unknown effects. Another patient carried a heterozygous splice site mutation (not studied in detail).

BCKDHC Deficiency (Maple Syrup Urine Disease, MSUD)

The first clinical description of MSUD dates back to 1954 (Menkes et al. 1954). The disease was named after the distinctive sweet odor of urine, sweat, and even earwax of the patients; the origin of the burnt sugar smell was identified in isoleucine metabolites (Hassan and Gupta 2023). Unfortunately, the consequences of the dysfunctional BCKDHC are not confined to smelly by-products. Accumulation of the respective branched-chain α -keto acids and the corresponding amino acids leads to impairments in the central nervous system, skeletal muscle, and immune system (Blackburn et al. 2017).

This genetic disease is governed by autosomal recessive inheritance and classified to several subtypes according to genetic origin and clinical features. MSUD types IA, IB, II, and III trace back to hE1b α , hE1b β , hE2b, and hE3 deficiencies (Le and Burrage 2021). Uncategorized disease forms are characterized by increased or decreased activities of the regulatory components (Oyarzabal et al. 2013; Maguolo et al. 2022; Ozcelik et al. 2023). Inborn errors in the first three main subunits (hE1b α , hE1b β , hE2b) are responsible for the great majority of the clinical cases (in a 45:35:20 ratio for hE1b α , hE1b β , and hE2b, respectively) (Strauss et al. 2020).

The patient population, which includes statistically every ~185,000th live birth worldwide (Chuang et al. 2019), is divided into five main groups (classic, intermediate, intermittent, thiamine-responsive, and E3 deficiency types) (Hassan and Gupta 2023). The most common (classic) MSUD is responsible for 50–75% of the clinical cases (orpha.net). The consequences of the 0–2% residual BCKDHC activities are elevated branched-chain amino acid, alloisoleucine, and branched-chain α -keto acid levels in plasma and urine, often even in the neonatal age (Hassan and Gupta 2023). Unmanaged neonates exhibit irritability, poor feeding, lethargy, intermittent apnea, and opisthotonus with “bicycling” movements, which develops to nausea, anorexia, dystonia, ataxia, cognitive impairment, hyperactivity, sleep disturbances, hallucinations, and choreoathetosis later (Blackburn et al. 2017). Patients of the less severe intermediate form, which accounts for ~30% of the cases (orpha.net), show 3–30% residual overall activities (Hassan and Gupta 2023). The higher activities lead to variable onsets and milder symptoms including feeding problems, poor growth, and developmental delay (Blackburn et al. 2017). The intermittent form is generally associated with normal growth/neurological development and stress-induced classic disease episodes (Blackburn et al. 2017); usually 5–20% residual overall activities characterize these cases (Hassan and Gupta 2023). The thiamine-responsive form resembles the intermediate form in terms of clinical severity (Blackburn et al. 2017); here 2–40% residual overall activities are generally observed (Hassan and Gupta 2023). Although the thiamine-responsive form ameliorates upon thiamine administration, these patients also require

branched-chain amino acid restriction (Strauss et al. 2020). The fifth MSUD subtype collects all the hE3 deficiency cases (Le and Burrage 2021).

Direct correlations between the genotypic and phenotypic presentations have not unequivocally been established as yet. Slightly more severe cases are registered for the hE1 α and hE1 β subunits as compared to hE2b in the Human Gene Mutation Database.

High prevalence—1 in 176 live births (Marshall and DiGeorge 1981)—in the Mennonite community highlights the significance of the Y393N-hE1 α mutation, which compromises proper assembling in the hE1b component (AEvarsson et al. 2000).

The thiamine-responsive form is almost exclusively associated with mutations in the hE2b component. The explanation for this phenomenon was based on the mutation-initiated degradation of the hE2b protein. The consequences of the latter are not limited to the lack of the second subunit and reaction but also of the inactivation by BDK, which eventually leads to the hyperactivation of the hE1b component. Since the first half-reaction at hE1b (see Scheme 13.3) is still possible, thiamine supplementation provides constant cofactor provision and also the possibility for cofactor exchange at the hE1b component for the prolonged (partial) conversion of the relevant branched-chain α -keto acids in the hE1b-specific reaction (Chuang et al. 2004). Nevertheless, thiamine responsiveness of the V27A-hE1 β pathogenic mutation somewhat tinges the above hypothesis (Yang et al. 2012).

The newly revealed regulatory element mutations all led to mild MSUD forms (Oyarzabal et al. 2013; Ozcelik et al. 2023; Maguolo et al. 2022). Interestingly, the single pathogenic variant structure elucidated belongs to the N94K-hBDP (Dolatabad et al. 2019). Structure determination of other relevant disease-causing variants in the future will hopefully further facilitate the understanding of the molecular pathomechanisms of the MSUD and other diseases connected to the BCKDHC.

KADHc Deficiency

Recent findings implicate hE1a in the etiology of several metabolic disorders, like α -amino adipic and α -keto adipic aciduria associated with neurological symptoms, Charcot-Marie-Tooth disease type 2, obesity, type 2 diabetes (Danhauser et al. 2012; Stiles et al. 2016; Hagen et al. 2015; Leandro and Houten 2020; Plubell et al. 2018; Timmons et al. 2018; Lim et al. 2014; Xu et al. 2012, 2018; Luan et al. 2020), and eosinophilic esophagitis, a chronic allergic disorder (Sherrill et al. 2018). The *DHTRD1* gene was mapped to chromosome 10p14 (Bunik et al. 2008). α -amino adipic and α -keto adipic aciduria were linked to two pathogenic mutations by Danhauser et al. (2012); an A-to-G transition altering the first codon was one of the mutations, while the second mutation led to the G729R substitution that affected a conserved residue. The symptoms were caused by the reduced overall activity of the KADHc, and they were milder than those of glutaryl-CoA dehydrogenase deficiency. In mice, this aciduria led to elevated insulin secretion and myelin damage *via* dysregulation of the myelin protein zero transcription. Bezerra

and colleagues identified seven mutations at the protein level (Bezerra et al. 2020). Three substitutions (L234G, Q305H, R455Q) reside in the N-terminal domain, while the other four (R715C, G729R (see also above), P773L, S777P) are present near the middle domain. Study of orthologues showed that residue 715 is always an Arg, while residues 305 and 777 are also highly conserved (82% and 93%, respectively). None of these residues directly affect the conserved catalytic machinery of the 2-oxo acid dehydrogenase family. These mutations were shown to introduce instability and thermal lability into the protein structure without affecting the active site; the R715C variant is somewhat a more special case since here a salt-bridge network also is destabilized, and this results in reduced binding and acyl transfer to hE2k (Bezerra et al. 2020). Another study by the Houten group (Leandro et al. 2020b) described the impacts of ten pathogenic hE1a mutations concluding that the mutations all compromised to different degrees the thermal stability and the catalytic activity of the hE1a component. Three of the mutants (L234G, R455Q, R715C) displayed high propensity to form insoluble protein and were all virtually catalytically inactive.

Variant G729R was also studied by Zhang et al. (2020), where enzyme kinetics measurements, chemical cross-linking, and HDX-MS were used to characterize the reaction rate and subunit interactions of the reassembled mutant KADHc. They observed a 37 times lower rate for the glutaryl transfer reaction and a rearrangement in the interaction of the protein with hE2k, relative to the controls; the strength of the interaction was not significantly affected, only the position of the substrate transfer channel.

Charcot-Marie-Tooth disease type 2Q (CMT2Q) affected individuals from 13 to 71 years old in the Chinese population of Shandong, as described by Xu et al. (2012). All of these individuals had feet with abnormally high plantar longitudinal arches, and other upper- and lower-extremities-related muscle problems were also observed: atrophy, angulated fibers, and fat around the muscle bundle. Also, vacuoles were seen in their muscle mitochondria. All these symptoms were linked to a nonsense mutation in the *DHTKDI* gene (c.1455T>G, Tyr485TER), which resulted in a fast decay of the mutant mRNA. CMT2Q could also be observed in animals and was characterized by peripheral neuropathy with impaired functions due to nerve degeneration; muscle atrophy was generally also present (Xu et al. 2018).

E3 Deficiency

E3 deficiency is an autosomal recessive disorder that develops due to mutations to the *DLD* gene encoding hE3 [reviewed in (Szabo and Ambrus 2023; Cameron et al. 2006; Quinonez et al. 2013; Quinonez and Thoene 2014; Ambrus and Adam-Vizi 2018)]. Early-onset E3 deficiency often leads to premature death in connection with the characteristic severe metabolic decompensations (lactic acidosis, hypoglycemia, hyperammonemia), neurological abnormalities (lethargy, low muscle tone, convulsive seizures, psychomotor disorders, ataxia, microcephaly, encephalopathy,

delayed development, Leigh syndrome, etc.), or sometimes even hypertrophic cardiomyopathy (Cameron et al. 2006; Quinonez et al. 2013; Quinonez and Thoene 2014; Grafakou et al. 2003; Shany et al. 1999; Shaag et al. 1999; Odievre et al. 2005; Hong et al. 1996, 1997; Sakaguchi et al. 1986; Cerna et al. 2001). Another typical clinical presentation is episodic liver failure. The metabolic decompensation and hepatic encephalopathy that occur during acute liver failure can lead to the death of patients; however, in this form of E3 deficiency, the life expectancy of patients is longer, and symptoms often appear only in the later years of life (Barak et al. 1998; Shaag et al. 1999; Quinonez and Thoene 2014; Brassier et al. 2013; Hong et al. 2003; Cameron et al. 2006). Recently, the myopathic phenotype of E3 deficiency has also been reported in adult patients (Carrozzo et al. 2014; Quintana et al. 2010).

The versatile phenotypic presentation of E3 deficiency is a consequence of the complex pathomechanism. Mutations that result in premature protein synthesis termination, impaired splicing, or amino acid substitution/deletion have been reported to date (Cameron et al. 2006; Quinonez et al. 2013; Quinonez and Thoene 2014; Liu et al. 1993; Shany et al. 1999; Hong et al. 1996, 1997, 2003; Cerna et al. 2001; Grafakou et al. 2003; ; Odievre et al. 2005; Shaag et al. 1999; Brassier et al. 2013; Quintana et al. 2010; Carrozzo et al. 2014). According to the analyses of clinical samples (Cameron et al. 2006; Quinonez et al. 2013; Quinonez and Thoene 2014; Liu et al. 1993; Shany et al. 1999; Hong et al. 1996, 1997, 2003; Cerna et al. 2001; Grafakou et al. 2003; Odievre et al. 2005; Shaag et al. 1999; Brassier et al. 2013; Quintana et al. 2010; Carrozzo et al. 2014) and recombinant disease-causing hE3 variants (Ambrus et al. 2011; Szabo et al. 2023; Vaubel et al. 2011; Brautigam et al. 2006; Liu et al. 1995; Yuan and Kim 2010, 2012; Yuan et al. 2008, 2009; Patel et al. 2009; Kim 2012), the underlying molecular pathomechanisms primarily may include (i) decreased protein provision, (ii) impaired catalytic activity, and (iii) reduced affinities for the respective multienzyme complexes. It was concluded that disease severity correlates relatively well with the residual overall hKGDHc activity; it is surprising in a way that the residual LADH activities themselves have not displayed such good correlations (Shany et al. 1999; Cameron et al. 2006). The complex pathomechanism may further be aggravated by the fact that certain disease-causing hE3 variants (G194C-, E340K-, D444V-, and P453L-hE3) (Ambrus et al. 2011) and the hE1k-hE2k subcomplex (emerging due to impaired complex assembly) (Ambrus et al. 2015b) are also capable of stimulating/enhancing ROS production.

To better understand the corresponding molecular pathomechanisms, thorough *in vitro* biochemical and structure-related analyses have recently been carried out in our laboratory applying primarily calibrated gel-filtration, molecular dynamics (MD) simulation, HDX-MS, CD spectropolarimetry, and X-ray crystallography, as well as biochemical assays for LADH and overall complex activities, FAD loss, and ROS generation (Ambrus et al. 2011, 2015a, 2016; Szabo et al. 2018, 2019, 2023; Ambrus and Adam-Vizi 2013). Two pathogenic mutations, the G101del and M326V, were found to induce the dissociation of the homodimer, which most directly results in the complete loss of the LADH catalytic activity (see above) (Szabo et al. 2023).

Substitution of a Leu for the *cis*-Pro453 residue, adjacent to the catalytic base His452, led to significant direct and relayed structural alterations in the active site and the FAD-binding site (Szabo et al. 2019), accounting not only for the virtually lost LADH activity but likely also for the significantly increased ROS-producing activity associated with the disease-causing P453L-hE3 variant (Ambrus et al. 2011). Minor structural alterations were also detected in the active sites of the I358T- and I318T-hE3 disease-causing variants (Szabo et al. 2023). A group of pathogenic homodimer interface substitutions (D444V, I445M, R447G, R460G) was associated with altered properties of the H⁺/H₂O channel (Szabo et al. 2019). The G194C and G426E pathogenic substitutions both introduced a polar side chain and enhanced the local dynamics near the NAD⁺/NADH binding site (Szabo et al. 2019). Crystallization of the K37E-, E340K-, I12T-, G101del-, and M326V-hE3 pathogenic variants has hitherto been unsuccessful.

Therapeutic Considerations

Current clinical management strategies for the various α -keto acid dehydrogenase complex deficiencies are discussed in detail in the respective GeneReviews chapters (<https://www.ncbi.nlm.nih.gov/books/NBK1116/>). Current challenges and potential future strategies in the therapy/management of inherited metabolic/mitochondrial diseases were recently reviewed by us in (Szabo and Ambrus 2023).

Acknowledgments This work was supported by the Hungarian Scientific Research Fund (OTKA grant 143627, to A.A.) and National Research, Development and Innovation Fund (TKP2021-EGA-25 grant, to A.A.). Project no. TKP2021-EGA-25 has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-EGA funding scheme.

Appendix 1: Malfunctions of the α -Keto Acid Dehydrogenase Complexes in Diseases

Disease	α -Keto acid dehydrogenase complex malfunction	References
KGDHc for review, see: Hansen and Gibson (2022)		
Alzheimer’s disease	Decreased overall activity	Mastrogiacono et al. (1993), Gibson et al. (1988, 1998, 2012), Blass and Gibson (1991), Bubber et al. (2005), Jia et al. (2023), Yao et al. (2022)

Disease	α -Keto acid dehydrogenase complex malfunction	References
Parkinson's disease	Decreased overall activity due to E2 dysfunction	Mizuno et al. (1994), Kobayashi et al. (1998), Miki et al. (2017)
Friedrich's ataxia	Decreased overall activity due to E3 dysfunction and decreased immunoreactive levels of E3	Blass et al. (1976), Melancon et al. (1979), Costantini et al. (2016), Mastrogiacomio et al. (1996)
Spinocerebellar ataxia type 1	Decreased immunoreactive levels of E2 and E3	Mastrogiacomio et al. (1996)
Huntington's disease	Decreased overall activity	Naseri et al. (2015)
Cancer	Dysregulation of overall activity	Chang et al. (2022), Vatrinet et al. (2017), Anderson et al. (2016, 2021), Ilic et al. (2017), Shen et al. (2021), Shin et al. (2020), Cai et al. (2023)
Head trauma/traumatic brain injury	Decreased overall activity	Pandya et al. (2023), Mkrtchyan et al. (2018)
Spinal cord injury	Decreased overall activity	Boyko et al. (2021)
Progressive supranuclear palsy	Decreased overall activity in cerebellum and superior frontal cortex	Park et al. (2001), Albers et al. (2000)
Wernicke-Korsakoff syndrome	Decreased overall activity	Butterworth et al. (1993)
Lactic acidosis in childhood	Low overall activity due to E2 dysfunction	Guffon et al. (1993), Bonnefont et al. (1992)
Dihydrolipoamide dehydrogenase (E3) deficiency	Decreased overall activity and (with selected mutations) enhanced ROS generation due to <i>DLD</i> gene mutations	Quinonez and Thoene (1993), Wongkittichote et al. (2023), Ambrus and Adam-Vizi (2018), Alfarsi et al. (2021), Staretz-Chacham et al. (2021)
Atypical KGDHc deficiency with no PDHc and BCKDHc deficiencies	Decreased overall activity due to <i>DLD</i> gene mutation	Odievre et al. (2005)
(Isolated) KGDHc-E1 deficiency	Decreased overall activity due to <i>OGDH</i> gene mutations	Kohlschutter et al. (1982), Guffon et al. (1993), Whittle et al. (2023)
Pheochromocytoma/paraganglioma syndrome-7	Perturbed overall activity and tumor suppressor function due to heterozygous missense <i>DLST</i> mutations	Remacha et al. (2019)
Epilepsy	Decreased overall activity	McDonald et al. (2017)
PDHc for review, see: Ganetzky et al. (1993), Patel et al. (2012), Pavlu-Pereira et al. (2020)		
Alzheimer's disease	Decreased overall activity	Sorbi et al. (1983), Jia et al. (2023)

Disease	α -Keto acid dehydrogenase complex malfunction	References
Friedrich's ataxia	Decreased overall activity and downregulation of E1p subunit	Blass et al. (1976), Melancon et al. (1979), Costantini et al. (2016), Pathak et al. (2019)
Huntington's disease	Decreased overall activity	Naseri et al. (2015); Sorbi et al. (1983), Butterworth et al. (1985)
Epilepsy	Decreased overall activity due to mutations in <i>PDHA1</i> , <i>PDHB</i> , and <i>PDK1</i>	Bhandary and Aguan (2015), Prasad et al. (2011), Vaz et al. (2022), McDonald et al. (2017)
Wernicke-Korsakoff syndrome	Decreased overall activity	Butterworth et al. (1993)
Lactic acidosis in childhood	Decreased overall activity due to genetic defect in the E1p subunit, reduced PDP activity	Brown (1992), Robinson and Sherwood (1975)
Cancer	Low overall activity due to upregulation of PDK, PDP1 activity maintains HIF activity	Stacpoole (2017), Schoenmann et al. (2023), Anwar et al. (2021), Sradhanjali and Reddy (2018), Woolbright et al. (2019), Guertin and Wellen (2023), Karagiota et al. (2023)
PDHc-E2 deficiency	Decreased overall activity due to <i>DLAT</i> gene mutation	McWilliam et al. (2010), Friedman et al. (2017)
Dihydrolipoamide dehydrogenase (E3) deficiency	Decreased overall activity due to <i>DLD</i> gene mutation	See above
PDHc-E1 α deficiency	Decreased overall activity due to <i>PDHA1</i> gene mutation	Brown (1992), Dahl (1995), Laxmi et al. (2023), Tanner et al. (2023), Yoshida et al. (2017), Kara et al. (2017), Drakulic et al. (2018), Pavlu-Pereira et al. (2021)
PDHc-E1 β deficiency	Decreased overall activity due to <i>PDHB</i> gene mutation	Prasad et al. (2011), Okajima et al. (2008)
Leigh syndrome	Decreased overall activity due to <i>PDHB</i> , <i>PDHA1</i> , <i>DLD</i> , and <i>PDHX</i> gene mutations	Chabrol et al. (1994), Arai et al. (1991), Quintana et al. (2009), Schiff et al. (2006), Grafakou et al. (2003), Craigen (1996), Briones et al. (1996)
PDHc-E3BP deficiency	Decreased overall activity due to <i>PDHX</i> gene mutation	Brown et al. (2002, 2006); Hargreaves et al. (2003), Fazeli et al. (2016)
PDP deficiency	Decreased overall activity due to <i>PDP1</i> gene mutation	Bedoyan et al. (2019), Maj et al. (2005, 2006), Cameron et al. (2009)
Charcot-Marie-Tooth disease	Decreased overall activity due to <i>PDK3</i> gene mutation	Narayanan et al. (2021), Perez-Siles et al. (2016), Kennerson et al. (2013, 2016)

Disease	α -Keto acid dehydrogenase complex malfunction	References
BCKDHc		
Maple syrup urine disease, types 1 and 2 (MSUD 1 and 2)	Decreased overall activity due to mutations in the E1b or E2b subunit <i>PPMIK</i> genetic mutation (phosphatase PP2Cm)	Burrage et al. (2014), Dancis et al. (1959), Strauss et al. (2020), Oyarzabal et al. (2013)
Obesity, diabetes mellitus type 2	Suppressed hE1b activity in obesity models in adipose tissue and liver	Nadler et al. (2000), She et al. (2007)
Diabetic nephropathy associated with chronic hyperinsulinemia	Increased overall activity	Shin et al. (2014)
Breast cancer	Increased mRNA level of hE1b α and hE1b β	Zhang and Han (2017)
Hepatocellular carcinoma	Low protein level from <i>BCKDHA</i> , low overall BCKDHc activity	Ericksen et al. (2019)
Chronic kidney disease	Increased overall activity	Hara et al. (1987)
Heart failure	Decreased overall activity	Sun et al. (2016), Wang et al. (2016)
Dihydrolipoamide dehydrogenase (E3) deficiency, MSUD 3	Decreased overall activity due to <i>DLD</i> gene mutation	See above
KADHC		
α -Aminoacidic and α -ketoacidic aciduria	Decreased overall activity due to <i>DHTKD1</i> gene mutation	Danhauser et al. (2012), Stiles et al. (2016), Hagen et al. (2015), Leandro and Houten (2020)
Charcot-Marie-Tooth disease type 2	Decreased overall activity due to <i>DHTKD1</i> gene mutation	Xu et al. (2018, 2012), Luan et al. (2020)
Amyotrophic lateral sclerosis	Decreased overall activity due to <i>DHTKD1</i> gene mutation	Osmanovic et al. (2021)
Spinal muscular atrophy-infantile onset with cognitive delay	Decreased overall activity due to <i>DHTKD1</i> gene mutation	Karakaya et al. (2018)
Eosinophilic esophagitis	Decreased overall activity due to <i>DHTKD1</i> gene mutation	Sherrill et al. (2018)
Oral squamous cell carcinoma	Circular RNA of <i>DHTKD1</i> promotes tumor formation	Wu et al. (2021)
Cancer	Over- or underexpression of <i>DHTKD1</i>	Papatheodorou et al. (2018), http://gepia.cancer-pku.cn/detail.php?gene=DHTKD1###

**Appendix 2: Most Important Structural Coordinates
Deposited in the Protein Data Bank (PDB) Related
to the α -Keto Acid Dehydrogenase Complexes^{a,b}**

Protein	Species	PDB ID	Small molecules
E1k (homodimer)	<i>Homo sapiens</i>	7WGR*	ThDP
	<i>Mycolicibacterium smegmatis</i>	2XT6	ThDP
		2XTA	ThDP acetyl-CoA
		2Y0P	ThDP intermediate acetyl-CoA
		2YIC	ThDP
		2YID	ThDP intermediate
		3ZHS	ThDP intermediate
		3ZHT	ThDP intermediate
		3ZHU	ThDP intermediate
		3ZHV	ThDP intermediate
	<i>Escherichia coli</i>	2JGD	AMP
		6VEF*	AMP oxaloacetate
	<i>Vibrio vulnificus</i>	6KM9	ThDP
E1a (homodimer)	<i>Homo sapiens</i>	6U3J	ThDP
		6SY1	ThDP
		5RW1	ThDP
E2k			
→ LD	<i>Escherichia coli</i>	1PMR [#]	—
	<i>Azotobacter vinelandii</i>	1GHJ [#]	—
		1GHK [#]	—
→ PSBD	<i>Escherichia coli</i>	1BAL [#]	—
		1BBL [#]	—
→ CD	<i>Homo sapiens</i>	6H05*	—
	<i>Bos taurus</i>	7UOL*	—
	<i>Thermochaetoides thermophila</i>	7Q5Q*	—
		8OIU*	—
	<i>Escherichia coli</i>	6PBR	—

Protein	Species	PDB ID	Small molecules
E1k-E2k ^c (homohexamer)	<i>Mycolicibacterium smegmatis</i>	8P5R (hetero-16-mer with glycogen accumulation regulator)	ThDP
	<i>Corynebacterium glutamicum</i>	8P5S	ThDP acetyl-CoA CoA
		8P5T*	ThDP acetyl-CoA
		8P5U*	ThDP acetyl-CoA CoA
		8P5V*	ThDP acetyl-CoA succinyl-CoA
8P5W*	ThDP derivative acetyl-CoA		
E1b (heterotetramer)	<i>Homo sapiens</i>	1DTW	ThDP
		1OLS	ThDP
		1U5B	ThDP
		1V1R	–
		1WCI	ThDP intermediate
		2BEU	ThDP intermediate
		2BEV	ThDP intermediate
		2BFB	ThDP
		2BFE	ThDP
	<i>Pseudomonas putida</i>	2BP7	–
	<i>Thermus thermophilus</i>	1UM9	–
		1UMB	ThDP
		1UMD	ThDP α-ketoiso-caproate
E1b (homodimer)	<i>Lactococcus lactis</i>	2VBF	ThDP
		2VBG	ThDP intermediate
E2b			
→ LD	<i>Homo sapiens</i>	1K8M [#]	–
		1K8O [#]	–
→ PSBD	<i>Homo sapiens</i>	2COO [#]	
		1ZWV [#]	

Protein	Species	PDB ID	Small molecules
→ CD (24-mer)	<i>Bos taurus</i>	2IHW	–
		2II4	CoA
		2II5	isobutyryl-CoA
	<i>Mycobacterium tuberculosis</i>	6ZZN	–
→ CD (42-mer)	<i>Thermoplasma acidophilum</i>	4OFS (truncated to form functional homotrimer)	–
E1p (heterotetramer)	<i>Homo sapiens</i>	3EXE	ThDP
	<i>Geobacillus stearothermophilus</i>	1W85(with E2-PSBD)	ThDP
	<i>Pyrobaculum aerophilum</i>	1IK6 (β subunit homotetramer)	–
E1p (homodimer)	<i>Escherichia coli</i>	2G67	–
		1L8A	ThDP
		2IEA	ThDP
		4QOY (with E2-LD-PSBD)	–
E2p			
→LD2/2 (inner)	<i>Homo sapiens</i>	1FYC [#]	–
		3CRK (with PDK2 <i>Rattus norvegicus</i>)	L-Lys-DHLA
		2PNR (with PDK3)	DHLA
		1Y8N (with PDK3)	DHLA
		1Y8O (with PDK3)	DHLA ADP
		1Y8P (with PDK3)	DHLA ATP
→ LD (only)	<i>Geobacillus stearothermophilus</i>	1LAB [#]	–
		1LAC [#]	–
	<i>Thermoplasma acidophilum</i>	2L5T [#]	–
→ LD1/3 (outermost)	<i>Azotobacter vinelandii</i>	1IYU [#]	–
		1IYV [#]	–
→ LD3/3 (innermost)	<i>Escherichia coli</i>	1QJO [#]	–
		7B9K* (with CD)	L-Lys-DHLA
→ PSBD	<i>Geobacillus stearothermophilus</i>	2PDD [#]	–
		2PDE [#]	–
		1W3D [#]	–
		1EBD (with E3)	–

Protein	Species	PDB ID	Small molecules
→ CD (60-mer)	<i>Homo sapiens</i>	3B8K*	–
		6CT0*	–
		6H55*	–
	<i>Bos taurus</i>	7UOM*	–
	<i>Neurospora crassa</i>	6ZLO*	–
	<i>Geobacillus stearothermophilus</i>	1B5S	–
	<i>Thermochaetoides thermophila</i>	7OTT*	–
		7Q5R*	–
		7BGJ*	–
→ CD (24-mer)	<i>Escherichia coli</i>	7B9K* (with LD)	L-Lys-DHLA
		4N72 (truncated to form functional homotrimers)	–
	<i>Azotobacter vinelandii</i>	1EAA	–
		1EAB	CoA lipoamide
		1EAE	lipoamide
→ CD (trimer)	<i>Corynebacterium glutamicum</i>	1EAF	–
		6ZZI	–
		6ZZK	CoA dihydro-lipoamide
	<i>Corynebacterium mustelae</i>	6ZZL	–
		6ZZM	CoA
→ CD (60-mer + 12-mer)	<i>Neurospora crassa</i>	6ZLM* (with E3BP)	–
		7R5M* (trimer with E3BP-core-binding domain (CBD))	–
		8OHS* (trimer with E3BP-CBD)	–
E3BP	<i>Homo sapiens</i>	6H60* (monomer)	–
	<i>Neurospora crassa</i>	6ZLM* (+E2p)	–

Protein	Species	PDB ID	Small molecules
E3	<i>Homo sapiens</i>	1ZMC	FAD NAD ⁺
		1ZMD	FAD NADH
		5NHG	FAD
		6I4Q	FAD
		2F5Z	FAD
		1ZY8 (with E3BP-E3BD)	FAD
		3RNM (with mutated E2b-PSBD)	FAD
	<i>Sus scrofa</i>	1IVI	–
	<i>Saccharomyces cerevisiae</i>	1JEH	FAD
		1V59	FAD NAD ⁺
	<i>Geobacillus stearothermophilus</i>	1EBD (with E2p-PSBD)	FAD
	<i>Escherichia coli</i>	4JDR	FAD
	<i>Azotobacter vinelandii</i>	3LAD	FAD
	<i>Pseudomonas fluorescens</i>	1LPF	FAD
	<i>Pseudomonas aeruginosa</i>	5U8W	FAD NADH
	<i>Neisseria gonorrhoeae</i>	5U25	FAD

^aAll structures were determined by X-ray crystallography unless noted otherwise

^bArtificial and pathogenic mutants as well as complexes with non-physiological substrates/ligands and inhibitors are not included in this table

^cThe E1 and E2 chains are joint protein products and are both present in the structures

^{*}Structures determined by EM

[#]Structures determined by NMR spectroscopy

Appendix 3: Chain Stoichiometry and Structural Topology in the α -Keto Acid Dehydrogenase Complexes

Type of complex	Species	Chain stoichiometry (E1:E2:E3)	Location of the peripheral subunits ^a	References
PDHc	<i>Homo sapiens</i>	120:60:12:12 (E1[$\alpha_2\beta_2$]:E2:E3BP:E3)	E1: outer sphere E3: outer sphere	Smolle et al. (2006), Ambrus et al. (2015b)
	<i>Homo sapiens</i>	160:40:20:40 (E1[$\alpha_2\beta_2$]:E2:E3BP:E3)	E1: outer sphere E3: outer sphere	Brautigam et al. (2009)
	<i>Homo sapiens</i>	48:12 (E2:E3BP)	–	Jiang et al. (2018)
	<i>Homo sapiens</i>	40:20 (E2:E3BP)	–	Prajapati et al. (2019)
	<i>Bos taurus</i>	120:60:10-12 (E1[$\alpha_2\beta_2$]:E2:E3)	E1: edges E3: faces	Barrera et al. (1972)
	<i>Bos taurus</i>	60:12:12 (E2:E3BP:E3)	E3: faces	Sanderson et al. (1996)
	<i>Bos taurus</i>	88:60:12:12 (E1[$\alpha_2\beta_2$]:E2:E3BP:E3)	E1: outer sphere E3: outer sphere	Zhou et al. (2001)
	<i>Bos taurus</i>	48:12 (E2:E3BP)	–	Hiromasa et al. (2004)
	<i>Bos taurus</i>	–	E1: edges E3: faces	Liu et al. (2022)
	<i>Saccharomyces cerevisiae</i>	60:12:12 (E2:E3BP:E3)	E3: outer sphere, faces	Stoops et al. (1997)
	<i>Geobacillus stearothermophilus</i>	48:60:12	E1: outer sphere E3: outer sphere	Domingo et al. (1999), Milne et al. (2002, 2006), Frank et al. (2005)
	<i>Escherichia coli</i>	24:24:12	E1: edges E3: faces	Ambrus et al. (2015b), Yang et al. (1985, 1986)
	<i>Azotobacter vinelandii</i>	24:24:12	E1: edges E3: faces	Mattevi et al. (1992a)

Type of complex	Species	Chain stoichiometry (E1:E2:E3)	Location of the peripheral subunits ^a	References
KGDHc	<i>Homo sapiens</i>	24:24:12	–	Ambrus et al. (2015b)
	<i>Mammalian</i>	24:24:12	E1: edges E3: faces	Bunik and Sievers (2002), Gibson et al. (2000)
	<i>Bos taurus</i>	24	E1: edges E3: faces	Liu et al. (2022)
	<i>Bos taurus</i>	24 (E2)	E1: edges	Hackert et al. (1989)
	<i>Bos taurus</i>	(2:2:3) _x	–	Hirashima et al. (1967)
	<i>Bos taurus</i>	(6:1:6) _x	–	Tanaka et al. (1972)
	<i>Rattus norvegicus</i>	10:12:14	–	Bunik et al. (2008)
	<i>Escherichia coli</i>	12:24:12	E1: between edges and faces E3: between edges and faces	Ambrus et al. (2015b), Wagenknecht et al. (1983, 1986, 1987)
	<i>Acetobacter xylinum</i>	12:12:12	–	de Kok et al. (1980)
BCKDHc	<i>Homo sapiens</i>	48:24:12 E1: [$\alpha_2\beta_2$]	E1: edges E3: faces	AEvarsson et al. (2000)
	<i>Bos taurus</i>	24 (E2)	E1: edges	Hackert et al. (1989)

^aThe positions are described with respect to the (generally) 24-meric (cubic) and 60-meric (pentagonal dodecahedron-like) E2 core geometries

Appendix 4: Moonlighting Functions of the α -Keto Acid Dehydrogenase Complexes and Their Subcomplexes/Components/Subunits

Enzyme	Moonlighting function	Details	References
KGDHc	Lysine acetyltransferase 2A (KAT2A) binding	It takes place in the nucleus. KGDHc is a source of succinyl-CoA for KAT2A, a histone succinyltransferase.	Wang et al. (2017)
	Generation of amino acids necessary for the assembly as well as growth of myofibrils	KGDHc is localized in the Z-disc of the indirect flight muscle. The E1 subunit of the complex is recruited to the Z-disc by Zasp52 (<i>Drosophila melanogaster</i>).	González Morales et al. (2023)
	Regulation of HIF1 α prolyl hydroxylases (PHDs)	KGDHc mutations indirectly inhibit the PHDs via promoting L-2-hydroxyglutarate production.	Burr et al. (2016)
PDHc	Histon acetylation	Mammalian PDHc can translocate to the nucleus and generate acetyl-CoA for the acetylation of core histons.	Sutendra et al. (2014)
	Mitochondrial autophagy	PDHc directly regulates the associated kinases and phosphatases, which eventually determines the phosphorylation states and mitophagic fates of selected mitochondrial matrix proteins (<i>Saccharomyces cerevisiae</i>).	Kolitsida et al. (2023)
PDHc and its subunits	Contribution to the virulence of <i>Mycoplasma</i> species	They are present at the surfaces of the bacteria.	Grundel et al. (2015)
PDHc subunits (mainly E2p, but also E1p α , E1p β , and likely the overall complex)	Interaction with the transcription factor STAT5A	It takes place in various cell types and modulates gene expression.	Chueh et al. (2011), Richard et al. (2017)
E2p	RNA binding	Binding to the mRNA that encodes the D1 protein of photosystem II links the acetyl-CoA-consuming lipid syntheses and biogenesis of the photosynthetic membranes (<i>Chlamydomonas reinhardtii</i>). In vitro, E2p could (unspecifically) bind RNA molecules originating from different prokaryotic and eukaryotic species (cyanobacteria, yeast, <i>Homo sapiens</i> , etc.).	Bohne et al. (2013)
E2p and E3BP	Autoantigen formation	It is involved in primary biliary cirrhosis.	Gershwin et al. (2000), Hirschfield et al. (2010), Pacini et al. (2015), Selmi et al. (2011), Wang et al. (2014b), Yamagiwa et al. (2014)

Enzyme	Moonlighting function	Details	References
E2p (other PDHc enzymes are also assumed)	Intramolecular nitroxyl (HNO) formation by E2p-lipoate and nitric oxide	HNO interacts with thiols, leading to irreversible modifications including sulfonamide formation. E3 subunit proved to be an HNO target.	Palmieri et al. (2023)
E1k and E1p	Defense against glutamate anaplerosis and nitroxidative stress	They can provide electrons for peroxidase and peroxynitrite reductase activities (<i>Mycobacterium tuberculosis</i>).	Maksymiuk et al. (2015)
E1k (non-human) and E1p	Generation of ROS	Isolated E1 subunits generate ROS (considered to be at a physiologically insignificant rate, <i>Bacillus stearothermophilus</i> / <i>E. coli</i> /human E1p, <i>E. coli</i> E1k)	Frank et al. (2008), Ambrus et al. (2015b)
KGDHc, PDHc, BCKDHc (all can be either prokaryotic or eukaryotic), and hE1k-hE2k	Generation of ROS	In contrast to the other α -keto acid dehydrogenase complexes, the hE1k-hE2k subcomplex also produces physiologically and pathologically significant amounts of ROS	Ambrus et al. (2009, 2015b), Tretter and Adam-Vizi (2004), Starkov et al. (2004b), Quinlan et al. (2014)

Enzyme	Moonlighting function	Details	References
E3	Diaphorase activity	Broadly, it is a reaction in the reverse catalytic direction (using NADH as substrate) with O ₂ (oxidase activity, producing superoxide and H ₂ O ₂), ubiquinone, organic dyes (e.g., 2,6-dichlorophenolindophenol (DCPIP), strictly the diaphorase activity), nitric oxide, inorganic complexes, labile ferric iron (electron-transferase activity), etc. This activity is enhanced by selected pathogenic, often dimer interface, mutations, and acidosis.	Vaubel et al. (2011), Klyachko et al. (2005), Corran et al. (1939), Gazaryan et al. (2002), Kalse and Veeger (1968), Massey et al. (1960), Straub (1939), Massey (1960b), Ambrus et al. (2011), Igamberdiev et al. (2004)
	Protease activity	This activity is enhanced by selected dimer interface mutations. The protease function may be implicated in the pathogenesis of E3 deficiency.	Vaubel et al. (2011), Babady et al. (2007)
	Adherence to metal-oxide surfaces	It is present in bacteria, like <i>Rhodococcus ruber</i> . However, the ability is conserved and also identified in human E3. A modified enzyme is able to interact with integrin-rich cancer cells followed by “integrin-assisted-endocytosis”. This allows harnessing the enzyme for cancer therapy.	Fleminger and Dayan (2021)
	DNA binding	E3 interacts with double-stranded DNA. Multiple molecules can bind to one DNA molecule. E3 was proposed to play a role in apoptosis.	Fleminger and Dayan (2021)
	G4-DNA oligonucleotide (G-quadruplex) binding	G4-DNA binding was reported in <i>Tetrahymena thermophila</i> . Further experiments showed that mammalian E3 also interacts with G4-DNA.	Kee et al. (1998)
	Binding to complement regulator proteins (Vitronectin, Factor H, FHL-1, and Clusterin) and plasminogen	To block the complement attack, <i>Pseudomonas aeruginosa</i> E3 interacts with human regulator proteins.	Hallström et al. (2012, 2015)
	Removal of ethoxylate chains of non-ionic ethoxylated surfactants via formation of oxygen radicals	<i>Pseudomonas nitroreducens</i> strain TX1 can use these surfactants as sole carbon source.	Hung et al. (2019)
	Menadione-mediated cytochrome <i>c</i> reduction	It is observed in <i>Streptomyces seoulensis</i> via superoxide anion radical production.	Youn and Kang (2000)
	Poly(ADP-ribose) polymerase (PARP) activity	It is prompted by nitrosative stress (investigated only in KGDHc).	Pankotai et al. (2009)

Enzyme	Moonlighting function	Details	References
E3 containing a lipoyl domain	Interaction with the regulatory domains of several sugar (galactose, raffinose, stachyose) transport proteins	It happens to membrane-bound E3 in <i>Streptococcus pneumoniae</i> , <i>E. coli</i> , and other pathogens (<i>Trypanosoma brucei</i> is also suggested).	Tyx et al. (2011), Fleminger and Dayan (2021)
	Protein acetyltransferase (transacetylase, TAase) activity	The N-terminal region of this TAase is LADH. It is involved in the transfer of acetyl groups from polyphenolic acetates to glutathione-S-transferase (GST) (<i>Starkeyomyces koorchalomoides</i> and <i>Streptococcus pneumoniae</i>).	Tyagi et al. (2009)

References

- Adam-Vizi V, Starkov AA (2010) Calcium and mitochondrial reactive oxygen species generation: how to read the facts. *J Alzheimers Dis* 20(Suppl 2):S413–S426. <https://doi.org/10.3233/JAD-2010-100465>
- AEvarsson A, Chuang JL, Wynn RM, Turley S, Chuang DT, Hol WG (2000) Crystal structure of human branched-chain alpha-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease. *Structure* 8(3):277–291. [https://doi.org/10.1016/s0969-2126\(00\)00105-2](https://doi.org/10.1016/s0969-2126(00)00105-2)
- Albers DS, Augood SJ, Park LC, Browne SE, Martin DM, Adamson J, Hutton M, Standaert DG, Vonsattel JP, Gibson GE, Beal MF (2000) Frontal lobe dysfunction in progressive supranuclear palsy: evidence for oxidative stress and mitochondrial impairment. *J Neurochem* 74(2):878–881. <https://doi.org/10.1046/j.1471-4159.2000.740878.x>
- Alfarsi A, Alfadhel M, Alameer S, Alhashem A, Tabarki B, Ababneh F, Al Mutairi F (2021) The phenotypic spectrum of dihydrolipoamide dehydrogenase deficiency in Saudi Arabia. *Mol Genet Metab Rep* 29:100817. <https://doi.org/10.1016/j.ymgmr.2021.100817>
- Almo SC, Bonanno JB, Sauder JM, Emtage S, Dilorenzo TP, Malashkevich V, Wasserman SR, Swaminathan S, Eswaramoorthy S, Agarwal R, Kumaran D, Madegowda M, Ragumani S, Patskovsky Y, Alvarado J, Ramagopal UA, Faber-Barata J, Chance MR, Sali A, Fiser A, Zhang ZY, Lawrence DS, Burley SK (2007) Structural genomics of protein phosphatases. *J Struct Funct Genomics* 8(2-3):121–140. <https://doi.org/10.1007/s10969-007-9036-1>
- Ambrus A, Adam-Vizi V (2013) Molecular dynamics study of the structural basis of dysfunction and the modulation of reactive oxygen species generation by pathogenic mutants of human dihydrolipoamide dehydrogenase. *Archiv Biochem Biophys* 538(2):145–155. <https://doi.org/10.1016/j.abb.2013.08.015>
- Ambrus A, Adam-Vizi V (2018) Human dihydrolipoamide dehydrogenase (E3) deficiency: Novel insights into the structural basis and molecular pathomechanism. *Neurochem Int* 117:5–14. <https://doi.org/10.1016/j.neuint.2017.05.018>
- Ambrus A, Tretter L, Adam-Vizi V (2009) Inhibition of the alpha-ketoglutarate dehydrogenase-mediated reactive oxygen species generation by lipoic acid. *J Neurochem* 109(Suppl 1):222–229. <https://doi.org/10.1111/j.1471-4159.2009.05942.x>
- Ambrus A, Torocsik B, Tretter L, Ozohanics O, Adam-Vizi V (2011) Stimulation of reactive oxygen species generation by disease-causing mutations of lipoamide dehydrogenase. *Hum Mol Genet* 20(15):2984–2995. <https://doi.org/10.1093/hmg/ddr202>
- Ambrus A, Mizsei R, Adam-Vizi V (2015a) Structural alterations by five disease-causing mutations in the low-pH conformation of human dihydrolipoamide dehydrogenase (hLADH) analyzed by molecular dynamics – implications in functional loss and modulation of reac-

- tive oxygen species generation by pathogenic hLADH forms. *Biochem Biophys Rep* 2:50–56. <https://doi.org/10.1016/j.bbrep.2015.04.006>
- Ambrus A, Nemeria NS, Torocsik B, Tretter L, Nilsson M, Jordan F, Adam-Vizi V (2015b) Formation of reactive oxygen species by human and bacterial pyruvate and 2-oxoglutarate dehydrogenase multienzyme complexes reconstituted from recombinant components. *Free Radic Biol Med* 89:642–650. <https://doi.org/10.1016/j.freeradbiomed.2015.10.001>
- Ambrus A, Wang JJ, Mizsei R, Zambo Z, Torocsik B, Jordan F, Adam-Vizi V (2016) Structural alterations induced by ten disease-causing mutations of human dihydrolipoamide dehydrogenase analyzed by hydrogen/deuterium-exchange mass spectrometry: Implications for the structural basis of E3 deficiency. *Biochim Biophys Acta Mol Basis Dis* 1862 11:2098–2109. <https://doi.org/10.1016/j.bbadis.2016.08.013>
- Anderson NM, Li D, Peng HL, Laroche FJ, Mansour MR, Gjini E, Aioub M, Helman DJ, Roderick JE, Cheng T, Harrold I, Samaha Y, Meng L, Amsterdam A, Neuberg DS, Denton TT, Sanda T, Kelliher MA, Singh A, Look AT, Feng H (2016) The TCA cycle transferase DLST is important for MYC-mediated leukemogenesis. *Leukemia* 30(6):1365–1374. <https://doi.org/10.1038/leu.2016.26>
- Anderson NM, Qin X, Finan JM, Lam A, Athoe J, Missiaen R, Skuli N, Kennedy A, Saini AS, Tao T, Zhu S, Nissim I, Look AT, Qing G, Simon MC, Feng H (2021) Metabolic enzyme DLST promotes tumor aggression and reveals a vulnerability to OXPHOS inhibition in high-risk neuroblastoma. *Cancer Res* 81(17):4417–4430. <https://doi.org/10.1158/0008-5472.CAN-20-2153>
- Andreyev AY, Kushnareva YE, Starkov AA (2005) Mitochondrial metabolism of reactive oxygen species. *Biochemistry (Mosc)* 70(2):200–214. <https://doi.org/10.1007/s10541-005-0102-7>
- Andreyev AY, Kushnareva YE, Murphy AN, Starkov AA (2015) Mitochondrial ROS metabolism: 10 years later. *Biochemistry (Mosc)* 80(5):517–531. <https://doi.org/10.1134/S0006297915050028>
- Andreyev AY, Yang H, Doulias P-T, Dolatabadi N, Zhang X, Luevanos M, Blanco M, Baal C, Putra I, Nakamura T, Ischiropoulos H, Tannenbaum SR, Lipton SA (2023) Metabolic bypass rescues aberrant S-nitrosylation-induced TCA cycle inhibition and synapse loss in Alzheimer's disease human neurons. *bioRxiv*:2023.2010.2012.562118. doi:<https://doi.org/10.1101/2023.10.12.562118>
- Anwar S, Shamsi A, Mohammad T, Islam A, Hassan MI (2021) Targeting pyruvate dehydrogenase kinase signaling in the development of effective cancer therapy. *Biochim Biophys Acta Rev Cancer* 1876 1:188568. <https://doi.org/10.1016/j.bbcan.2021.188568>
- Arai Y, Miyasato Y, Koide H (1991) Characteristic changes on brain CT in a case of Leigh encephalopathy with deficiency of pyruvate dehydrogenase. *Brain Dev* 13(6):457–458. [https://doi.org/10.1016/s0387-7604\(12\)80050-7](https://doi.org/10.1016/s0387-7604(12)80050-7)
- Arjunan P, Nemeria N, Brunskill A, Chandrasekhar K, Sax M, Yan Y, Jordan F, Guest JR, Furey W (2002) Structure of the pyruvate dehydrogenase multienzyme complex E1 component from *Escherichia coli* at 1.85 Å resolution. *Biochemistry* 41(16):5213–5221. <https://doi.org/10.1021/bi0118557>
- Artukhovich AV, Grabarska A, Gumbarewicz E, Aleshin VA, Kahne T, Obata T, Kazantsev AV, Lukashev NV, Stepulak A, Fernie AR, Bunik VI (2020) Synthetic analogues of 2-oxo acids discriminate metabolic contribution of the 2-oxoglutarate and 2-oxoadipate dehydrogenases in mammalian cells and tissues. *Sci Rep* 10(1):22. <https://doi.org/10.1038/s41598-020-58701-4>
- Babady NE, Pang YP, Elpeleg O, Isaya G (2007) Cryptic proteolytic activity of dihydrolipoamide dehydrogenase. *Proc Natl Acad Sci U S A* 104(15):6158–6163. <https://doi.org/10.1073/pnas.061018104>
- Balakrishnan A, Nemeria NS, Chakraborty S, Kakalis L, Jordan F (2012) Determination of pre-steady-state constants on the *Escherichia coli* pyruvate dehydrogenase complex reveals that loop movement controls the rate-limiting step. *J Am Chem Soc* 134:18644–18655

- Bao HY, Kasten SA, Yan XH, Roche TE (2004) Pyruvate dehydrogenase kinase isoform 2 activity limited and further inhibited by slowing down the rate of dissociation of ADP. *Biochemistry* 43(42):13432–13441. <https://doi.org/10.1021/bi049488x>
- Barak N, Hummer D, Segal T, Ben Ari Z, Halevy J, Kasper RT (1998) Lipoamide dehydrogenase deficiency: a newly discovered cause of acute hepatitis in adults. *J Hepatol* 29(3):482–484. [https://doi.org/10.1016/S0168-8278\(98\)80069-x](https://doi.org/10.1016/S0168-8278(98)80069-x)
- Barrera CR, Namihira G, Hamilton L, Munk P, Eley MH, Linn TC, Reed LJ (1972) α -Keto acid dehydrogenase complexes. XVI. Studies on the subunit structure of the pyruvate dehydrogenase complexes from bovine kidney and heart. *Arch Biochem Biophys* 148(2):343–358
- Bedoyan JK, Hecht L, Zhang S, Tarrant S, Bergin A, Demirbas D, Yang E, Shin HK, Grahame GJ, DeBrosse SD, Hoppel CL, Kerr DS, Berry GT (2019) A novel null mutation in the pyruvate dehydrogenase phosphatase catalytic subunit gene (PDP1) causing pyruvate dehydrogenase complex deficiency. *JIMD Rep* 48(1):26–35. <https://doi.org/10.1002/jmd2.12054>
- Bell D, Berchuck A, Birrer M, Chien J, Cramer DW, Dao F, Dhir R, DiSaia P, Gabra H, Glenn P, Godwin AK, Gross J, Hartmann L, Huang M, Huntsman DG, Iacocca M, Imielinski M, Kalloger S, Karlan BY, Levine DA, Mills GB, Morrison C, Mutch D, Olvera N, Orsulic S, Park K, Petrelli N, Rabeno B, Rader JS, Sikic BI, Smith-McCune K, Sood AK, Bowtell D, Penny R, Testa JR, Chang K, Creighton CJ, Dinh HH, Drummond JA, Fowler G, Gunaratne P, Hawes AC, Kovar CL, Lewis LR, Morgan MB, Newsham IF, Santibanez J, Reid JG, Trevino LR, Wu YQ, Wang M, Muzny DM, Wheeler DA, Gibbs RA, Getz G, Lawrence MS, Cibulskis K, Sivachenko AY, Sougnez C, Voet D, Wilkinson J, Bloom T, Ardlie K, Fennell T, Baldwin J, Nichol R, Fisher S, Gabriel S, Lander ES, Ding L, Fulton RS, Koboldt DC, McLellan MD, Wylie T, Walker J, O’Laughlin M, Dooling DJ, Fulton L, Abbott R, Dees ND, Zhang Q, Kandoth C, Wendl M, Schierding W, Shen D, Harris CC, Schmidt H, Kalicki J, Delehaunty KD, Fronick CC, Demeter R, Cook L, Wallis JW, Lin L, Magrini VJ, Hodges JS, Eldred JM, Smith SM, Pohl CS, Vandin F, Upfal E, Raphael BJ, Weinstock GM, Mardis ER, Wilson RK, Meyerson M, Winckler W, Getz G, RGW V, Carter SL, Mermel CH, Saksena G, Nguyen H, Onofrio RC, Lawrence MS, Hubbard D, Gupta S, Crenshaw A, Ramos AH, Ardlie K, Chin L, Protopopov A, Zhang J, Kim TM, Perna I, Xiao Y, Zhang H, Ren G, Sathiamoorthy N, Park RW, Lee E, Park PJ, Kucherlapati R, Absher DM, Waite L, Sherlock G, Brooks JD, Li JZ, Xu J, Myers RM, Laird PW, Cope L, Herman JG, Shen H, Weisenberger DJ, Noushmehr H, Pan F, Triche T Jr, Berman BP, Van Den Berg DJ, Buckley J, Baylin SB, Spellman PT, Purdom E, Neuvial P, Bengtsson H, Jakkula LR, Durinck S, Han J, Dorton S, Marr H, Choi YG, Wang V, Wang NJ, Ngai J, Conboy JG, Parvin B, Feiler HS, Speed TP, Gray JW, Levine DA, Succi ND, Liang Y, Taylor BS, Schultz N, Borsu L, Lash AE, Brennan C, Viale A, Sander C, Ladanyi M, Hoadley KA, Meng S, Du Y, Shi Y, Li L, Turman YJ, Zang D, Helms EB, Balu S, Zhou X, Wu J, Topal MD, Hayes DN, Perou CM, Getz G, Voet D, Saksena G, Zhang J, Zhang H, Wu CJ, Shukla S, Cibulskis K, Lawrence MS, Sivachenko A, Jing R, Park RW, Liu Y, Park PJ, Noble M, Chin L, Carter H, Kim D, Samayoa J, Karchin R, Spellman PT, Purdom E, Neuvial P, Bengtsson H, Durinck S, Han J, Korkola JE, Heiser LM, Cho RJ, Hu Z, Parvin B, Speed TP, Gray JW, Schultz N, Cerami E, Taylor BS, Olshen A, Reva B, Antipin Y, Shen R, Mankoo P, Sheridan R, Ciriello G, Chang WK, Bernanke JA, Borsu L, Levine DA, Ladanyi M, Sander C, Haussler D, Benz CC, Stuart JM, Benz SC, Sanborn JZ, Vaske CJ, Zhu J, Szeto C, Scott GK, Yau C, Hoadley KA, Du Y, Balu S, Hayes DN, Perou CM, Wilkerson MD, Zhang N, Akbani R, Baggerly KA, Yung WK, Mills GB, Weinstein JN, Penny R, Shelton T, Grimm D, Hatfield M, Morris S, Yena P, Rhodes P, Sherman M, Paulauskis J, Millis S, Kahn A, Greene JM, The Cancer Genome Atlas Research N, Disease working g, tissue source s, Genome sequencing centres: Baylor College of M, Broad I, Washington University in St L, Cancer genome characterization centres: Broad Institute/Dana-Farber Cancer I, Harvard Medical S, HudsonAlpha Institute/Stanford U, University of Southern California/Johns Hopkins U, Lawrence Berkeley National L, Memorial Sloan-Kettering Cancer C, University of North Carolina at Chapel H, Genome data analysis centres: Broad I, Johns Hopkins U, University of California Santa Cruz/Buck I, The University of Texas MDACC, Biospecimen core r, Data coordination c (2011)

- Integrated genomic analyses of ovarian carcinoma. *Nature* 474(7353):609–615. <https://doi.org/10.1038/nature10166>
- Benen J, van Berkel W, Dieteren N, Arscott D, Williams C Jr, Veeger C, de Kok A (1992) Lipoamide dehydrogenase from *Azotobacter vinelandii*: site-directed mutagenesis of the His450-Glu455 diad. Kinetics of wild-type and mutated enzymes. *Eur J Biochem* 207(2):487–497
- Ben-Yoseph O, Badar-Goffer RS, Morris PG, Bachelard HS (1993) Glycerol 3-phosphate and lactate as indicators of the cerebral cytoplasmic redox state in severe and mild hypoxia respectively: a ¹³C- and ³¹P-n.m.r. study. *Biochem J* 291(Pt 3):915–919. <https://doi.org/10.1042/bj2910915>
- Berg A, Vervoort J, de Kok A (1996) Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. *J Mol Biol* 261(3):432–442. <https://doi.org/10.1006/jmbi.1996.0474>
- Berg A, Westphal AH, Bosma HJ, de Kok A (1998) Kinetics and specificity of reductive acylation of wild-type and mutated lipoyl domains of 2-oxo-acid dehydrogenase complexes from *Azotobacter vinelandii*. *Eur J Biochem* 252(1):45–50. <https://doi.org/10.1046/j.1432-1327.1998.2520045.x>
- Berthold CL, Gocke D, Wood MD, Leeper FJ, Pohl M, Schneider G (2007) Structure of the branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* provides insights into the structural basis for the chemoselective and enantioselective carbonylation reaction. *Acta Crystallogr D Biol Crystallogr* 63(Pt 12):1217–1224. <https://doi.org/10.1107/s0907444907050433>
- Bezerra GA, Foster WR, Bailey HJ, Hicks KG, Sauer SW, Dimitrov B, McCorvie TJ, Okun JG, Rutter J, Kolker S, Yue WW (2020) Crystal structure and interaction studies of human DHTKD1 provide insight into a mitochondrial megacomplex in lysine catabolism. *IUCrJ* 7(Pt 4):693–706. <https://doi.org/10.1107/S205225252000696X>
- Bhandary S, Aguan K (2015) Pyruvate dehydrogenase complex deficiency and its relationship with epilepsy frequency—an overview. *Epilepsy Res* 116:40–52. <https://doi.org/10.1016/j.epilepsyres.2015.07.002>
- Bixel M, Shimomura Y, Hutson S, Hamprecht B (2001) Distribution of key enzymes of branched-chain amino acid metabolism in glial and neuronal cells in culture. *J Histochem Cytochem* 49(3):407–418. <https://doi.org/10.1177/002215540104900314>
- Blackburn PR, Gass JM, Vairo FPE, Farnham KM, Atwal HK, Macklin S, Klee EW, Atwal PS (2017) Maple syrup urine disease: mechanisms and management. *Appl Clin Genet* 10:57–66. <https://doi.org/10.2147/TACG.S125962>
- Blass JP, Gibson GE (1991) The role of oxidative abnormalities in the pathophysiology of Alzheimer's disease. *Rev Neurol (Paris)* 147(6-7):513–525
- Blass JP, Kark RA, Menon NK (1976) Low activities of the pyruvate and oxoglutarate dehydrogenase complexes in five patients with Friedreich's ataxia. *N Engl J Med* 295(2):62–67. <https://doi.org/10.1056/NEJM197607082950202>
- Bohne AV, Schwarz C, Schottkowski M, Lidschreiber M, Piotrowski M, Zerges W, Nickelsen J (2013) Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. *PLoS Biol* 11(2):e1001482. <https://doi.org/10.1371/journal.pbio.1001482>
- Bollinger E, Peloquin M, Libera J, Albuquerque B, Pashos E, Shipstone A, Hadjipanayis A, Sun Z, Xing G, Clasquin M, Stansfield JC, Tierney B, Gernhardt S, Siddall CP, Greizer T, Geoly FJ, Vargas SR, Gao LC, Williams G, Marshall M, Rosado A, Steppan C, Filipski KJ, Zhang BB, Miller RA, Roth Flach RJ (2022) BDK inhibition acts as a catabolic switch to mimic fasting and improve metabolism in mice. *Mol Metab* 66:101611. <https://doi.org/10.1016/j.molmet.2022.101611>
- Bonnefont JP, Chretien D, Rustin P, Robinson B, Vassault A, Aupetit J, Charpentier C, Rabier D, Saudubray JM, Munnich A (1992) Alpha-ketoglutarate dehydrogenase deficiency presenting as congenital lactic acidosis. *J Pediatr* 121(2):255–258. [https://doi.org/10.1016/s0022-3476\(05\)81199-0](https://doi.org/10.1016/s0022-3476(05)81199-0)

- Bouzier-Sore AK, Voisin P, Bouchaud V, Bezancon E, Franconi JM, Pellerin L (2006) Competition between glucose and lactate as oxidative energy substrates in both neurons and astrocytes: a comparative NMR study. *Eur J Neurosci* 24(6):1687–1694. <https://doi.org/10.1111/j.1460-9568.2006.05056.x>
- Boyer B, Odessey R (1991) Kinetic characterization of branched chain ketoacid dehydrogenase. *Arch Biochem Biophys* 285(1):1–7. [https://doi.org/10.1016/0003-9861\(91\)90321-9](https://doi.org/10.1016/0003-9861(91)90321-9)
- Boyko A, Tsepkova P, Aleshin V, Artiukhov A, Mkrtchyan G, Ksenofontov A, Baratova L, Ryabov S, Graf A, Bunik V (2021) Severe spinal cord injury in rats induces chronic changes in the spinal cord and cerebral cortex metabolism, adjusted by thiamine that improves locomotor performance. *Front Mol Neurosci* 14:620593. <https://doi.org/10.3389/fnmol.2021.620593>
- Boyko AI, Karlina IS, Zavileyskiy LG, Aleshin VA, Artiukhov AV, Kaehne T, Ksenofontov AL, Ryabov SI, Graf AV, Tramonti A, Bunik VI (2022) Delayed impact of 2-oxoadipate dehydrogenase inhibition on the rat brain metabolism is linked to protein glutarylation. *Front Med* 9:18. <https://doi.org/10.3389/fmed.2022.896263>
- Brand MD (2016) Mitochondrial generation of superoxide and hydrogen peroxide as the source of mitochondrial redox signaling. *Free Radic Biol Med* 100:14–31. <https://doi.org/10.1016/j.freeradbiomed.2016.04.001>
- Brassier A, Ottolenghi C, Boutron A, Bertrand AM, Valmary-Degano S, Cervoni JP, Chretien D, Arnoux JB, Hubert L, Rabier D, Lacaille F, de Keyser Y, Di Martino V, de Lonlay P (2013) Dihydrolipoamide dehydrogenase deficiency: A still overlooked cause of recurrent acute liver failure and Reye-like syndrome. *Molecular Genetics and Metabolism* 109(1):28–32. <https://doi.org/10.1016/j.ymgme.2013.01.017>
- Brautigam CA, Chuang JL, Tomchick DR, Machius M, Chuang DT (2005) Crystal structure of human dihydrolipoamide dehydrogenase: NAD(+)/NADH binding and the structural basis of disease-causing mutations. *J Mol Biol* 350(3):543–552. <https://doi.org/10.1016/j.jmb.2005.05.014>
- Brautigam CA, Wynn RM, Chuang JL, Machius M, Tomchick DR, Chuang DT (2006) Structural insight into interactions between dihydrolipoamide dehydrogenase (E3) and E3 binding protein of human pyruvate dehydrogenase complex. *Structure* 14(3):611–621
- Brautigam CA, Wynn RM, Chuang JL, Chuang DT (2009) Subunit and catalytic component stoichiometries of an in vitro reconstituted human pyruvate dehydrogenase complex. *J Biol Chem* 284(19):13086–13098. <https://doi.org/10.1074/jbc.M806563200>
- Brautigam CA, Wynn RM, Chuang JL, Naik MT, Young BB, Huang TH, Chuang DT (2011) Structural and thermodynamic basis for weak interactions between dihydrolipoamide dehydrogenase and subunit-binding domain of the branched-chain alpha-ketoacid dehydrogenase complex. *J Biol Chem* 286(26):23476–23488. <https://doi.org/10.1074/jbc.M110.202960>
- Briones P, Lopez MJ, De Meirleir L, Ribes A, Rodes M, Martinez-Costa C, Peris M, Lissens W (1996) Leigh syndrome due to pyruvate dehydrogenase E1 alpha deficiency (point mutation R263G) in a Spanish boy. *J Inher Metab Dis* 19(6):795–796. <https://doi.org/10.1007/BF01799177>
- Brosnan JT, Brosnan ME (2006) Branched-chain amino acids: enzyme and substrate regulation. *J Nutr* 136(1 Suppl):207S–211S. <https://doi.org/10.1093/jn/136.1.207S>
- Brown GK (1992) Pyruvate dehydrogenase E1 alpha deficiency. *J Inher Metab Dis* 15(4):625–633. <https://doi.org/10.1007/BF01799619>
- Brown GK, Otero LJ, Legris M, Brown RM (1994) Pyruvate-dehydrogenase deficiency. *J Med Genet* 31(11):875–879. <https://doi.org/10.1136/jmg.31.11.875>
- Brown RM, Head RA, Brown GK (2002) Pyruvate dehydrogenase E3 binding protein deficiency. *Hum Genet* 110(2):187–191. <https://doi.org/10.1007/s00439-001-0665-3>
- Brown RM, Head RA, Morris AA, Raiman JA, Walter JH, Whitehouse WP, Brown GK (2006) Pyruvate dehydrogenase E3 binding protein (protein X) deficiency. *Dev Med Child Neurol* 48(9):756–760. <https://doi.org/10.1017/S0012162206001617>
- Bryk R, Arango N, Maksymiuk C, Balakrishnan A, Wu YT, Wong CH, Masquelin T, Hipskind P, Lima CD, Nathan C (2013) Lipoamide channel-binding sulfonamides selectively inhibit myco-

- bacterial lipoamide dehydrogenase. *Biochemistry* 52(51):9375–9384. <https://doi.org/10.1021/bi401077f>
- Bubber P, Haroutunian V, Fisch G, Blass JP, Gibson GE (2005) Mitochondrial abnormalities in Alzheimer brain: mechanistic implications. *Ann Neurol* 57(5):695–703. <https://doi.org/10.1002/ana.20474>
- Bunik VI, Degtyarev D (2008) Structure-function relationships in the 2-oxo acid dehydrogenase family: substrate-specific signatures and functional predictions for the 2-oxoglutarate dehydrogenase-like proteins. *Proteins* 71(2):874–890. <https://doi.org/10.1002/prot.21766>
- Bunik VI, Sievers C (2002) Inactivation of the 2-oxo acid dehydrogenase complexes upon generation of intrinsic radical species. *Eur J Biochem* 269(20):5004–5015
- Bunik V, Kaehne T, Degtyarev D, Shcherbakova T, Reiser G (2008) Novel isoenzyme of 2-oxoglutarate dehydrogenase is identified in brain, but not in heart. *FEBS J* 275(20):4990–5006. <https://doi.org/10.1111/j.1742-4658.2008.06632.x>
- Burr SP, Costa AS, Grice GL, Timms RT, Lobb IT, Freisinger P, Dodd RB, Dougan G, Lehner PJ, Frezza C, Nathan JA (2016) Mitochondrial protein lipoylation and the 2-oxoglutarate dehydrogenase complex controls HIF1alpha stability in aerobic conditions. *Cell Metab* 24(5):740–752. <https://doi.org/10.1016/j.cmet.2016.09.015>
- Burrage LC, Nagamani SC, Campeau PM, Lee BH (2014) Branched-chain amino acid metabolism: from rare Mendelian diseases to more common disorders. *Hum Mol Genet* 23(R1):R1–R8. <https://doi.org/10.1093/hmg/ddu123>
- Butterworth J, Yates CM, Reynolds GP (1985) Distribution of phosphate-activated glutaminase, succinic dehydrogenase, pyruvate dehydrogenase and gamma-glutamyl transpeptidase in post-mortem brain from Huntington's disease and agonal cases. *J Neurol Sci* 67(2):161–171. [https://doi.org/10.1016/0022-510x\(85\)90112-1](https://doi.org/10.1016/0022-510x(85)90112-1)
- Butterworth RF, Kril JJ, Harper CG (1993) Thiamine-dependent enzyme changes in the brains of alcoholics: relationship to the Wernicke-Korsakoff syndrome. *Alcohol Clin Exp Res* 17(5):1084–1088. <https://doi.org/10.1111/j.1530-0277.1993.tb05668.x>
- Byron O, Lindsay JG (2017) The pyruvate dehydrogenase complex and related assemblies in health and disease. *Subcell Biochem* 83:523–550. https://doi.org/10.1007/978-3-319-46503-6_19
- Cai Y, Lv L, Lu T, Ding M, Yu Z, Chen X, Zhou X, Wang X (2023) alpha-KG inhibits tumor growth of diffuse large B-cell lymphoma by inducing ROS and TP53-mediated ferroptosis. *Cell Death Discov* 9(1):182. <https://doi.org/10.1038/s41420-023-01475-1>
- Cameron JM, Levandovskiy V, MacKay N, Raiman J, Renaud DL, Clarke JTR, Feigenbaum A, Elpeleg O, Robinson BH (2006) Novel mutations in dihydrolipoamide dehydrogenase deficiency in two cousins with borderline-normal PDH complex activity. *Am J Med Genet A* 140A(14):1542–1552. <https://doi.org/10.1002/ajmg.a.31313>
- Cameron JM, Maj M, Levandovskiy V, Barnett CP, Blaser S, Mackay N, Raiman J, Feigenbaum A, Schulze A, Robinson BH (2009) Pyruvate dehydrogenase phosphatase 1 (PDP1) null mutation produces a lethal infantile phenotype. *Hum Genet* 125(3):319–326. <https://doi.org/10.1007/s00439-009-0629-6>
- Cardenas C, Lovy A, Silva-Pavez E, Urrea F, Mizzoni C, Ahumada-Castro U, Bustos G, Jana F, Cruz P, Farias P, Mendoza E, Huerta H, Murgas P, Hunter M, Rios M, Cerda O, Georgakoudi I, Zakarian A, Molgo J, Foscett JK (2020) Cancer cells with defective oxidative phosphorylation require endoplasmic reticulum-to-mitochondria Ca(2+) transfer for survival. *Sci Signal* 13(640). <https://doi.org/10.1126/scisignal.aay1212>
- Carrozzo R, Torraco A, Fiermonte G, Martinelli D, Di Nottia M, Rizza T, Voza A, Verrigni D, Diodato D, Parisi G, Maiorana A, Rizzo C, Pierri CL, Zucano S, Piemonte F, Bertini E, Dionisi-Vici C (2014) Riboflavin responsive mitochondrial myopathy is a new phenotype of dihydrolipoamide dehydrogenase deficiency. The chaperon-like effect of vitamin B2. *Mitochondrion* 18:49–57. <https://doi.org/10.1016/j.mito.2014.09.006>
- Cate RL, Roche TE, Davis LC (1980) Rapid intersite transfer of acetyl groups and movement of pyruvate dehydrogenase component in the kidney pyruvate dehydrogenase complex. *J Biol Chem* 255(16):7556–7562

- Cerna L, Wenchich L, Hansiková H, Kmoch S, Peskova K, Chrastina P, Brynda J, Zeman J (2001) Novel mutations in a boy with dihydrolipoamide dehydrogenase deficiency. *Med Sci Monit* 7(6):1319–1325
- Chabrol B, Mancini J, Benelli C, Gire C, Munnich A (1994) Leigh syndrome: pyruvate dehydrogenase defect. A case with peripheral neuropathy. *J Child Neurol* 9(1):52–55. <https://doi.org/10.1177/088307389400900113>
- Chakraborty J, Nemeria NS, Farinas E, Jordan F (2018) Catalysis of transthioacylation in the active centers of dihydrolipoamide acyltransferase components of 2-oxo acid dehydrogenase complexes. *FEBS Open Bio* 8(6):880–896. <https://doi.org/10.1002/2211-5463.12431>
- Chandrasekhar K, Wang JJ, Arjunan P, Sax M, Park YH, Nemeria NS, Kumaran S, Song JY, Jordan F, Furey W (2013) Insight to the interaction of the dihydrolipoamide acetyltransferase (E2) core with the peripheral components in the Escherichia coli pyruvate dehydrogenase complex via multifaceted structural approaches. *J Biol Chem* 288(21):15402–15417. <https://doi.org/10.1074/jbc.M113.466789>
- Chang CF, Chou HT, Chuang JL, Chuang DT, Huang TH (2002) Solution structure and dynamics of the lipoic acid-bearing domain of human mitochondrial branched-chain alpha-keto acid dehydrogenase complex. *J Biol Chem* 277(18):15865–15873. <https://doi.org/10.1074/jbc.M110952200>
- Chang LC, Chiang SK, Chen SE, Hung MC (2022) Targeting 2-oxoglutarate dehydrogenase for cancer treatment. *Am J Cancer Res* 12(4):1436–1455
- Chinopoulos C (2011a) The “B space” of mitochondrial phosphorylation. *J Neurosci Res* 89(12):1897–1904. <https://doi.org/10.1002/jnr.22659>
- Chinopoulos C (2011b) Mitochondrial consumption of cytosolic ATP: not so fast. *FEBS letters* 585(9):1255–1259. <https://doi.org/10.1016/j.febslet.2011.04.004>
- Chinopoulos C, Adam-Vizi V (2006) Calcium, mitochondria and oxidative stress in neuronal pathology. Novel aspects of an enduring theme. *FEBS J* 273(3):433–450. <https://doi.org/10.1111/j.1742-4658.2005.05103.x>
- Chinopoulos C, Seyfried TN (2018) Mitochondrial substrate-level phosphorylation as energy source for glioblastoma: review and hypothesis. *ASN Neuro* 10:1759091418818261. <https://doi.org/10.1177/1759091418818261>
- Chiu CC, Chung A, Barletta G, Jordan F (1996) Intramolecular model for the reductive acyl transfer catalyzed by α -keto acid dehydrogenases. *J Am Chem Soc* 118(45):11026–11029. <https://doi.org/10.1021/ja954322s>
- Chouchani ET, Pell VR, Gaude E, Aksentijevic D, Sundier SY, Robb EL, Logan A, Nadtochiy SM, Ord ENJ, Smith AC, Eyassu F, Shirley R, Hu CH, Dare AJ, James AM, Rogatti S, Hartley RC, Eaton S, Costa ASH, Brookes PS, Davidson SM, Duchon MR, Saeb-Parsy K, Shattock MJ, Robinson AJ, Work LM, Frezza C, Krieg T, Murphy MP (2014) Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. *Nature* 515(7527):431–435. <https://doi.org/10.1038/nature13909>
- Chuang JL, Wynn RM, Chuang DT (2002) The C-terminal hinge region of lipoic acid-bearing domain of E2b is essential for domain interaction with branched-chain alpha-keto acid dehydrogenase kinase. *J Biol Chem* 277(40):36905–36908. <https://doi.org/10.1074/jbc.C200430200>
- Chuang JL, Wynn RM, Moss CC, Song JL, Li J, Awad N, Mandel H, Chuang DT (2004) Structural and biochemical basis for novel mutations in homozygous Israeli maple syrup urine disease patients: a proposed mechanism for the thiamin-responsive phenotype. *J Biol Chem* 279(17):17792–17800. <https://doi.org/10.1074/jbc.M313879200>
- Chuang DT, Shih VE, Max Wynn R (2019) Maple syrup urine disease (Branched-chain ketoaciduria). In: Valle DL, Antonarakis S, Ballabio A, Beaudet AL, Mitchell GA (eds) *The online metabolic and molecular bases of inherited disease*. McGraw-Hill Education, New York, NY
- Chueh FY, Leong KF, Cronk RJ, Venkitachalam S, Pabich S, Yu CL (2011) Nuclear localization of pyruvate dehydrogenase complex-E2 (PDC-E2), a mitochondrial enzyme, and its role in signal transducer and activator of transcription 5 (STAT5)-dependent gene transcription. *Cell Signal* 23(7):1170–1178. <https://doi.org/10.1016/j.cellsig.2011.03.004>

- Ciszak EM, Korotchkina LG, Dominiak PM, Sidhu S, Patel MS (2003) Structural basis for flip-flop action of thiamin pyrophosphate-dependent enzymes revealed by human pyruvate dehydrogenase. *J Biol Chem* 278(23):21240–21246. <https://doi.org/10.1074/jbc.M300339200>
- Ciszak EM, Makal A, Hong YS, Vettaikkorumakankauv AK, Korotchkina LG, Patel MS (2006) How dihydrolipoamide dehydrogenase-binding protein binds dihydrolipoamide dehydrogenase in the human pyruvate dehydrogenase complex. *J Biol Chem* 281(1):648–655
- Cole JT, Sweatt AJ, Hutson SM (2012) Expression of mitochondrial branched-chain aminotransferase and alpha-keto-acid dehydrogenase in rat brain: implications for neurotransmitter metabolism. *Front Neuroanat* 6:18. <https://doi.org/10.3389/fnana.2012.00018>
- Cooney GJ, Denyer GS, Jenkins AB, Storlien LH, Kraegen EW, Caterson ID (1993) In vivo insulin sensitivity of the pyruvate dehydrogenase complex in tissues of the rat. *Am J Physiol* 265(1 Pt 1):E102–E107. <https://doi.org/10.1152/ajpendo.1993.265.1.E102>
- Corran HS, Green DE, Straub FB (1939) On the catalytic function of heart flavoprotein. *Biochem J* 33(5):793–801
- Costantini A, Laureti T, Pala MI, Colangeli M, Cavalieri S, Pozzi E, Brusco A, Salvarani S, Serrati C, Fancellu R (2016) Long-term treatment with thiamine as possible medical therapy for Friedreich ataxia. *J Neurol* 263(11):2170–2178. <https://doi.org/10.1007/s00415-016-8244-7>
- Craigden WJ (1996) Leigh disease with deficiency of lipoamide dehydrogenase: treatment failure with dichloroacetate. *Pediatr Neurol* 14(1):69–71. [https://doi.org/10.1016/0887-8994\(96\)00005-7](https://doi.org/10.1016/0887-8994(96)00005-7)
- Dahl HH (1995) Pyruvate dehydrogenase E1 alpha deficiency: males and females differ yet again. *Am J Hum Genet* 56(3):553–557
- Dahl HH, Maragos C, Brown RM, Hansen LL, Brown GK (1990) Pyruvate dehydrogenase deficiency caused by deletion of a 7-bp repeat sequence in the E1 alpha gene. *Am J Hum Genet* 47(2):286–293
- Dahl HH, Brown GK, Brown RM, Hansen LL, Kerr DS, Wexler ID, Patel MS, De Meirleir L, Lissens W, Chun K et al (1992) Mutations and polymorphisms in the pyruvate dehydrogenase E1 alpha gene. *Hum Mutat* 1(2):97–102. <https://doi.org/10.1002/humu.1380010203>
- Damuni Z, Reed LJ (1987) Purification and characterization of a divalent cation-independent, spermine-stimulated protein phosphatase from bovine kidney mitochondria. *J Biol Chem* 262(11):5133–5138
- Dancis J, Levitz M, Miller S, Westall RG (1959) Maple syrup urine disease. *Br Med J* 1(5114):91–93. <https://doi.org/10.1136/bmj.1.5114.91>
- Danhauser K, Sauer SW, Haack TB, Wieland T, Staufner C, Graf E, Zschocke J, Strom TM, Traub T, Okun JG, Meitinger T, Hoffmann GF, Prokisch H, Kolker S (2012) DHTKD1 mutations cause 2-aminoadipic and 2-oxoadipic aciduria. *Am J Hum Genet* 91(6):1082–1087. <https://doi.org/10.1016/j.ajhg.2012.10.006>
- Dardel F, Davis AL, Laue ED, Perham RN (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J Mol Biol* 229(4):1037–1048. <https://doi.org/10.1006/jmbi.1993.1103>
- Davie JR, Wynn RM, Meng M, Huang YS, Aalund G, Chuang DT, Lau KS (1995) Expression and characterization of branched-chain alpha-ketoacid dehydrogenase kinase from the rat. Is it a histidine-protein kinase? *J Biol Chem* 270(34):19861–19867. <https://doi.org/10.1074/jbc.270.34.19861>
- de Kok A, van Berkel WJH (1996) Lipoamide dehydrogenase. In: Patel MS, Roche TE, Harris RA (eds) Alpha-keto acid dehydrogenase complexes. Birkhäuser, pp 52–69. <https://doi.org/10.1007/978-3-0348-8981-0>
- de Kok A, Kornfeld S, Benziman M, Milner Y (1980) Subunit composition and partial reactions of the 2-oxoglutarate dehydrogenase complex of *Acetobacter xylinum*. *Eur J Biochem* 106(1):49–58
- de Kok A, Berg A, van Berkel W, Fabisz-Kijowska A, Westphal A, van den Akker F, Mattevi A, Hol WGJ (1994) The pyruvate dehydrogenase complex from *Azotobacter vinelandii*. In: Yagi K (ed) Flavins and flavoproteins 1993, Nagoya, Japan, 1993. Walter de Gruyter, pp 587–596

- Denton RM (2009) Regulation of mitochondrial dehydrogenases by calcium ions. *Biochim Biophys Acta* 1787(11):1309–1316. <https://doi.org/10.1016/j.bbambio.2009.01.005>
- Denton RM, Pullen TJ, Armstrong CT, Heesom KJ, Rutter GA (2016) Calcium-insensitive splice variants of mammalian E1 subunit of 2-oxoglutarate dehydrogenase complex with tissue-specific patterns of expression. *Biochem J*. <https://doi.org/10.1042/bcj20160135>
- Devedjiev Y, Steussy CN, Vassilyev DG (2007) Crystal structure of an asymmetric complex of pyruvate dehydrogenase kinase 3 with lipoyl domain 2 and its biological implications. *J Mol Biol* 370(3):407–416. <https://doi.org/10.1016/j.jmb.2007.04.083>
- Dimou A, Tsimihodimos V, Bairaktari E (2022) The critical role of the branched chain amino acids (BCAAs) catabolism-regulating enzymes, branched-chain aminotransferase (BCAT) and branched-chain alpha-keto acid dehydrogenase (BCKD), in human pathophysiology. *Int J Mol Sci* 23(7). <https://doi.org/10.3390/ijms23074022>
- Dobolyi A, Bago A, Palkovits M, Nemeria NS, Jordan F, Doczi J, Ambrus A, Adam-Vizi V, Chinopoulos C (2020) Exclusive neuronal detection of KGDHC-specific subunits in the adult human brain cortex despite pancellular protein lysine succinylation. *Brain Struct Funct* 225(2):639–667. <https://doi.org/10.1007/s00429-020-02026-5>
- Dolatabad MR, Guo LL, Xiao P, Zhu Z, He QT, Yang DX, Qu CX, Guo SC, Fu XL, Li RR, Ge L, Hu KJ, Liu HD, Shen YM, Yu X, Sun JP, Zhang PJ (2019) Crystal structure and catalytic activity of the PPM1K N94K mutant. *J Neurochem* 148(4):550–560. <https://doi.org/10.1111/jnc.14631>
- Domingo GJ, Chauhan HJ, Lessard IA, Fuller C, Perham RN (1999) Self-assembly and catalytic activity of the pyruvate dehydrogenase multienzyme complex from *Bacillus stearothermophilus*. *Eur J Biochem* 266(3):1136–1146. <https://doi.org/10.1046/j.1432-1327.1999.00966.x>
- Drakulic S, Rai J, Petersen SV, Golas MM, Sander B (2018) Folding and assembly defects of pyruvate dehydrogenase deficiency-related variants in the E1 α subunit of the pyruvate dehydrogenase complex. *Cell Mol Life Sci* 75(16):3009–3026. <https://doi.org/10.1007/s00018-018-2775-2>
- Dudkina NV, Kouřil R, Peters K, Braun H-P, Boekema EJ (2010) Structure and function of mitochondrial supercomplexes. *Biochim Biophys Acta Bioenerg* 1797(6):664–670. <https://doi.org/10.1016/j.bbambio.2009.12.013>
- Dutta R, Inouye M (2000) GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci* 25(1):24–28. [https://doi.org/10.1016/s0968-0004\(99\)01503-0](https://doi.org/10.1016/s0968-0004(99)01503-0)
- Endo H, Hasegawa K, Narisawa K, Tada K, Kagawa Y, Ohta S (1989) Defective gene in lactic acidosis: abnormal pyruvate dehydrogenase E1 α -subunit caused by a frame shift. *Am J Hum Genet* 44(3):358–364
- Ericksen RE, Lim SL, McDonnell E, Shuen WH, Vadiveloo M, White PJ, Ding Z, Kwok R, Lee P, Radda GK, Toh HC, Hirschey MD, Han W (2019) Loss of BCAA catabolism during carcinogenesis enhances mTORC1 activity and promotes tumor development and progression. *Cell Metab* 29(5):1151–1165. e1156. <https://doi.org/10.1016/j.cmet.2018.12.020>
- Fan J, Shan C, Kang HB, Elf S, Xie J, Tucker M, Gu TL, Aguiar M, Lonning S, Chen H, Mohammadi M, Britton LM, Garcia BA, Aleckovic M, Kang Y, Kaluz S, Devi N, Van Meir EG, Hitosugi T, Seo JH, Lonial S, Gaddh M, Arellano M, Khoury HJ, Khuri FR, Boggon TJ, Kang S, Chen J (2014) Tyr phosphorylation of PDP1 toggles recruitment between ACAT1 and SIRT3 to regulate the pyruvate dehydrogenase complex. *Mol Cell* 53(4):534–548. <https://doi.org/10.1016/j.molcel.2013.12.026>
- Faure M, Bourguignon J, Neuburger M, MacHerel D, Sieker L, Ober R, Kahn R, Cohen-Addad C, Douce R (2000) Interaction between the lipamide-containing H-protein and the lipamide dehydrogenase (L-protein) of the glycine decarboxylase multienzyme system 2. Crystal structures of H- and L-proteins. *Eur J Biochem* 267(10):2890–2898. <https://doi.org/10.1046/j.1432-1033.2000.01330.x>
- Fazeli W, Karakaya M, Herkenrath P, Vierzig A, Dotsch J, von Kleist-Retzow JC, Cirak S (2016) Mendeliome sequencing enables differential diagnosis and treatment of neonatal lactic acidosis. *Mol Cell Pediatr* 3(1):22. <https://doi.org/10.1186/s40348-016-0050-x>

- Felig P, Marliss E, Cahill GF Jr (1969) Plasma amino acid levels and insulin secretion in obesity. *N Engl J Med* 281(15):811–816. <https://doi.org/10.1056/NEJM196910092811503>
- Filla A, Butterworth RF, Geoffroy G, Lemieux B, Barbeau A (1978) Serum and platelet lipoamide dehydrogenase in Friedreich's ataxia. *Can J Neurol Sci* 5(1):111–114
- Fisher-Wellman KH, Gilliam LAA, Lin CT, Cathey BL, Lark DS, Darrell Neuffer P (2013) Mitochondrial glutathione depletion reveals a novel role for the pyruvate dehydrogenase complex as a key H₂O₂-emitting source under conditions of nutrient overload. *Free Radic Biol Med* 65:1201–1208. <https://doi.org/10.1016/j.freeradbiomed.2013.09.008>
- Fleminger G, Dayan A (2021) The moonlighting activities of dihydrolipoamide dehydrogenase: biotechnological and biomedical applications. *J Mol Recognit* 34(11):e2924. <https://doi.org/10.1002/jmr.2924>
- Forsberg BO, Aibara S, Howard RJ, Mortezaei N, Lindahl E (2020) Arrangement and symmetry of the fungal E3BP-containing core of the pyruvate dehydrogenase complex. *Nat Commun* 11(1):4667. <https://doi.org/10.1038/s41467-020-18401-z>
- Frank RA, Titman CM, Pratap JV, Luisi BF, Perham RN (2004) A molecular switch and proton wire synchronize the active sites in thiamine enzymes. *Science (New York, NY)* 306(5697):872–876. <https://doi.org/10.1126/science.1101030>
- Frank RA, Pratap JV, Pei XY, Perham RN, Luisi BF (2005) The molecular origins of specificity in the assembly of a multienzyme complex. *Structure* 13(8):1119–1130. <https://doi.org/10.1016/j.str.2005.04.021>
- Frank RA, Price AJ, Northrop FD, Perham RN, Luisi BF (2007) Crystal structure of the E1 component of the *Escherichia coli* 2-oxoglutarate dehydrogenase multienzyme complex. *J Mol Biol* 368(3):639–651. <https://doi.org/10.1016/j.jmb.2007.01.080>
- Frank RAW, Kay CWM, Hirs J, Luisi BF (2008) Off-pathway, oxygen-dependent thiamine radical in the Krebs cycle. *J Am Chem Soc* 130(5):1662–1668. <https://doi.org/10.1021/ja076468k>
- Friedman J, Feigenbaum A, Chuang N, Silhavy J, Gleeson JG (2017) Pyruvate dehydrogenase complex-E2 deficiency causes paroxysmal exercise-induced dyskinesia. *Neurology* 89(22):2297–2298. <https://doi.org/10.1212/WNL.0000000000004689>
- Ganetzky R, McCormick EM, Falk MJ (1993) Primary pyruvate dehydrogenase complex deficiency overview. In: Adam MP, Feldman J, Mirzaa GM et al (eds) *GeneReviews*(R), Seattle, WA
- Gazaryan IG, Krasnikov BF, Ashby GA, Thorneley RNF, Kristal BS, Brown AM (2002) Zinc is a potent inhibitor of thiol oxidoreductase activity and stimulates reactive oxygen species production by lipoamide dehydrogenase. *J Biol Chem* 277(12):10064–10072
- Gershwin ME, Ansari AA, Mackay IR, Nakanuma Y, Nishio A, Rowley MJ, Coppel RL (2000) Primary biliary cirrhosis: an orchestrated immune response against epithelial cells. *Immunol Rev* 174:210–225. <https://doi.org/10.1034/j.1600-0528.2002.017402.x>
- Gherardi G, Monticelli H, Rizzuto R, Mammucari C (2020) The mitochondrial Ca(2+) uptake and the fine-tuning of aerobic metabolism. *Front Physiol* 11:554904. <https://doi.org/10.3389/fphys.2020.554904>
- Gibson GE, Sheu KF, Blass JP, Baker A, Carlson KC, Harding B, Perrino P (1988) Reduced activities of thiamine-dependent enzymes in the brains and peripheral tissues of patients with Alzheimer's disease. *Arch Neurol* 45(8):836–840. <https://doi.org/10.1001/archneur.1988.00520320022009>
- Gibson GE, Zhang H, Sheu KF, Bogdanovich N, Lindsay JG, Lannfelt L, Vestling M, Cowburn RF (1998) Alpha-ketoglutarate dehydrogenase in Alzheimer brains bearing the APP670/671 mutation. *Ann Neurol* 44(4):676–681. <https://doi.org/10.1002/ana.410440414>
- Gibson GE, Park LC, Sheu KF, Blass JP, Calingasan NY (2000) The alpha-ketoglutarate dehydrogenase complex in neurodegeneration. *Neurochem Int* 36(2):97–112
- Gibson GE, Chen HL, Xu H, Qiu L, Xu Z, Denton TT, Shi Q (2012) Deficits in the mitochondrial enzyme alpha-ketoglutarate dehydrogenase lead to Alzheimer's disease-like calcium dysregulation. *Neurobiol Aging* 33(6):1121 e1113–1121 e1124. <https://doi.org/10.1016/j.neurobiolaging.2011.11.003>
- Glasser NR, Wang BX, Hoy JA, Newman DK (2017) The pyruvate and alpha-ketoglutarate dehydrogenase complexes of *Pseudomonas aeruginosa* catalyze pyocyanin and phenazine-

- l-carboxylic acid reduction via the subunit dihydrolipoamide dehydrogenase. *J Biol Chem* 292(13):5593–5607. <https://doi.org/10.1074/jbc.M116.772848>
- Goncalves RL, Bunik VI, Brand MD (2016) Production of superoxide/hydrogen peroxide by the mitochondrial 2-oxoadipate dehydrogenase complex. *Free Radic Biol Med* 91:247–255. <https://doi.org/10.1016/j.freeradbiomed.2015.12.020>
- González Morales N, Marescal O, Szikora S, Katzemich A, Correia-Mesquita T, Bíró P, Erdelyi M, Mihály J, Schöck F (2023) The oxoglutarate dehydrogenase complex is involved in myofibril growth and Z-disc assembly in *Drosophila*. *J Cell Sci* 136(13). <https://doi.org/10.1242/jcs.260717>
- Gorlach A, Dimova EY, Petry A, Martinez-Ruiz A, Hernansanz-Agustin P, Rolo AP, Palmeira CM, Kietzmann T (2015) Reactive oxygen species, nutrition, hypoxia and diseases: problems solved? *Redox Biol* 6:372–385. <https://doi.org/10.1016/j.redox.2015.08.016>
- Grafakou O, Oexle K, van den Heuvel L, Smeets R, Trijbels F, Goebel HH, Bosshard N, Superti-Furga A, Steinmann B, Smeitink J (2003) Leigh syndrome due to compound heterozygosity of dihydrolipoamide dehydrogenase gene mutations. Description of the first E3 splice site mutation. *Eur J Pediatr* 162(10):714–718. <https://doi.org/10.1007/s00431-003-1282-z>
- Graham LD, Packman LC, Perham RN (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* 28(4):1574–1581. <https://doi.org/10.1021/bi00430a023>
- Green JD, Laue ED, Perham RN, Ali ST, Guest JR (1995) Three-dimensional structure of a lipoyl domain from the dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J Mol Biol* 248(2):328–343. [https://doi.org/10.1016/s0022-2836\(95\)80054-9](https://doi.org/10.1016/s0022-2836(95)80054-9)
- Green T, Grigorian A, Klyuyeva A, Tuganova A, Luo M, Popov KM (2008) Structural and functional insights into the molecular mechanisms responsible for the regulation of pyruvate dehydrogenase kinase 2. *J Biol Chem* 283(23):15789–15798. <https://doi.org/10.1074/jbc.M800311200>
- Grundel A, Friedrich K, Pfeiffer M, Jacobs E, Dumke R (2015) Subunits of the pyruvate dehydrogenase cluster of *Mycoplasma pneumoniae* are surface-displayed proteins that bind and activate human plasminogen. *PLoS One* 10(5):e0126600. <https://doi.org/10.1371/journal.pone.0126600>
- Guertin DA, Wellen KE (2023) Acetyl-CoA metabolism in cancer. *Nat Rev Cancer* 23(3):156–172. <https://doi.org/10.1038/s41568-022-00543-5>
- Guest JR, Lewis HM, Graham LD, Packman LC, Perham RN (1985) Genetic reconstruction and functional analysis of the repeating lipoyl domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J Mol Biol* 185(4):743–754. [https://doi.org/10.1016/0022-2836\(85\)90059-2](https://doi.org/10.1016/0022-2836(85)90059-2)
- Guffon N, Lopez-Mediavilla C, Dumoulin R, Mousson B, Godinot C, Carrier H, Collombet JM, Divry P, Mathieu M, Guibaud P (1993) 2-Ketoglutarate dehydrogenase deficiency, a rare cause of primary hyperlactataemia: report of a new case. *J Inherit Metab Dis* 16(5):821–830. <https://doi.org/10.1007/BF00714273>
- Guo Y, Qiu W, Roche TE, Hackert ML (2020) Crystal structure of the catalytic subunit of bovine pyruvate dehydrogenase phosphatase. *Acta Crystallogr F Struct Biol Commun* 76(Pt 7):292–301. <https://doi.org/10.1107/S2053230X20007943>
- Hackert ML, Xu WX, Oliver RM, Wall JS, Hainfeld JF, Mullinax TR, Reed LJ (1989) Branched-chain alpha-keto acid dehydrogenase complex from bovine kidney: radial distribution of mass determined from dark-field electron micrographs. *Biochemistry* 28(17):6816–6821. <https://doi.org/10.1021/bi00443a006>
- Hagen J, te Brinke H, Wanders RJ, Knecht AC, Oussoren E, Hoogeboom AJ, Ruijter GJ, Becker D, Schwab KO, Franke I, Duran M, Waterham HR, Sass JO, Houten SM (2015) Genetic basis of alpha-aminoaciduria and alpha-ketoadipic aciduria. *J Inherit Metab Dis* 38(5):873–879. <https://doi.org/10.1007/s10545-015-9841-9>

- Halim ND, McFate T, Mohyeldin A, Okagaki P, Korotchikina LG, Patel MS, Jeoung NH, Harris RA, Schell MJ, Verma A (2010) Phosphorylation status of pyruvate dehydrogenase distinguishes metabolic phenotypes of cultured rat brain astrocytes and neurons. *Glia* 58(10):1168–1176. <https://doi.org/10.1002/glia.20996>
- Hallström T, Mörgelin M, Barthel D, Raguse M, Kunert A, Hoffmann R, Skerka C, Zipfel PF (2012) Dihydrolipoamide dehydrogenase of *Pseudomonas aeruginosa* is a surface-exposed immune evasion protein that binds three members of the factor H family and plasminogen. *J Immunol* 189(10):4939–4950. <https://doi.org/10.4049/jimmunol.1200386>
- Hallstrom T, Uhde M, Singh B, Skerka C, Riesbeck K, Zipfel PF (2015) *Pseudomonas aeruginosa* uses dihydrolipoamide dehydrogenase (Lpd) to bind to the human terminal pathway regulators vitronectin and clusterin to inhibit terminal pathway complement attack. *Plos One* 10(9):e0137630. <https://doi.org/10.1371/journal.pone.0137630>
- Hansen GE, Gibson GE (2022) The alpha-ketoglutarate dehydrogenase complex as a hub of plasticity in neurodegeneration and regeneration. *Int J Mol Sci* 23(20). <https://doi.org/10.3390/ijms232012403>
- Hara Y, May RC, Kelly RA, Mitch WE (1987) Acidosis, not azotemia, stimulates branched-chain, amino acid catabolism in uremic rats. *Kidney Int* 32(6):808–814. <https://doi.org/10.1038/ki.1987.280>
- Hargreaves IP, Heales SJ, Briddon A, Lee PJ, Hanna MG, Land JM (2003) Primary pyruvate dehydrogenase E3 binding protein deficiency with mild hyperlactataemia and hyperalaninaemia. *J Inherit Metab Dis* 26(5):505–506. <https://doi.org/10.1023/a:1025181512847>
- Harper AE, Miller RH, Block KP (1984) Branched-chain amino acid metabolism. *Annu Rev Nutr* 4:409–454. <https://doi.org/10.1146/annurev.nu.04.070184.002205>
- Harris RA, Hawes JW, Popov KM, Zhao Y, Shimomura Y, Sato J, Jaskiewicz J, Hurley TD (1997) Studies on the regulation of the mitochondrial alpha-ketoacid dehydrogenase complexes and their kinases. *Adv Enzyme Regul* 37:271–293. [https://doi.org/10.1016/s0065-2571\(96\)00009-x](https://doi.org/10.1016/s0065-2571(96)00009-x)
- Hassan SA, Gupta V (2023) Maple syrup urine disease. In: StatPearls. Treasure Island (FL) ineligible companies. Disclosure: Vikas Gupta declares no relevant financial relationships with ineligible companies.
- Hatazawa Y, Tadaishi M, Nagaike Y, Morita A, Ogawa Y, Ezaki O, Takai-Igarashi T, Kitaura Y, Shimomura Y, Kamei Y, Miura S (2014) PGC-1alpha-mediated branched-chain amino acid metabolism in the skeletal muscle. *PLoS One* 9(3):e91006. <https://doi.org/10.1371/journal.pone.0091006>
- Hauschildt S, Neuman B, Brand K (1982) Calcium plays an important role in the regulation of hepatic branched-chain 2-oxo acid dehydrogenase activity. *FEBS Lett* 143(1):77–80. [https://doi.org/10.1016/0014-5793\(82\)80277-9](https://doi.org/10.1016/0014-5793(82)80277-9)
- Heldt HW, Schwalbach K (1967) The participation of GTP-AMP-P transferase in substrate level phosphate transfer of rat liver mitochondria. *Eur J Biochem* 1(2):199–206. https://doi.org/10.1007/978-3-662-25813-2_31
- Hevler JF, Albanese P, Cabrera-Orefice A, Potter A, Jankevics A, Misic J, Scheltema RA, Brandt U, Arnold S, Heck AJR (2023) MRPS36 provides a structural link in the eukaryotic 2-oxoglutarate dehydrogenase complex. *Open Biol* 13(3):220363. <https://doi.org/10.1098/rsob.220363>
- Hirashima M, Hayakawa T, Koike M (1967) Mammalian alpha-keto acid dehydrogenase complexes. II. An improved procedure for the preparation of 2-oxoglutarate dehydrogenase complex from pig heart muscle. *J Biol Chem* 242(5):902–907
- Hiromasa Y, Fujisawa T, Aso Y, Roche TE (2004) Organization of the cores of the mammalian pyruvate dehydrogenase complex formed by E2 and E2 plus the E3-binding protein and their capacities to bind the E1 and E3 components. *J Biol Chem* 279(8):6921–6933. <https://doi.org/10.1074/jbc.M308172200>
- Hiromasa Y, Hu L, Roche TE (2006) Ligand-induced effects on pyruvate dehydrogenase kinase isoform 2. *J Biol Chem* 281(18):12568–12579. <https://doi.org/10.1074/jbc.M513514200>

- Hirschfield GM, Heathcote EJ, Gershwin ME (2010) Pathogenesis of cholestatic liver disease and therapeutic approaches. *Gastroenterology* 139(5):1481–1496. <https://doi.org/10.1053/j.gastro.2010.09.004>
- Hitosugi T, Fan J, Chung TW, Lythgoe K, Wang X, Xie J, Ge Q, Gu TL, Polakiewicz RD, Roessel JL, Chen GZ, Boggon TJ, Lonial S, Fu H, Khuri FR, Kang S, Chen J (2011) Tyrosine phosphorylation of mitochondrial pyruvate dehydrogenase kinase 1 is important for cancer metabolism. *Mol Cell* 44(6):864–877. <https://doi.org/10.1016/j.molcel.2011.10.015>
- Hoffelder M, Raasch K, van Ooyen J, Eggeling L (2010) The E2 domain of OdhA of *Corynebacterium glutamicum* has succinyltransferase activity dependent on lipoyl residues of the acetyltransferase AceF. *J Bacteriol* 192(19):5203–5211. <https://doi.org/10.1128/jb.00597-10>
- Holecek M (2023) Role of impaired glycolysis in perturbations of amino acid metabolism in diabetes mellitus. *Int J Mol Sci* 24(2). <https://doi.org/10.3390/ijms24021724>
- Hong YS, Kerr DS, Craigen WJ, Tan J, Pan YZ, Lusk M, Patel MS (1996) Identification of two mutations in a compound heterozygous child with dihydrolipoamide dehydrogenase deficiency. *Hum Mol Genet* 5(12):1925–1930. <https://doi.org/10.1093/hmg/5.12.1925>
- Hong YS, Kerr DS, Liu TC, Lusk M, Powell BR, Patel MS (1997) Deficiency of dihydrolipoamide dehydrogenase due to two mutant alleles (E340K and G101del) - analysis of a family and prenatal testing. *Biochim Biophys Acta Mol Basis Dis* 1362(2-3):160–168. [https://doi.org/10.1016/S0925-4439\(97\)00073-2](https://doi.org/10.1016/S0925-4439(97)00073-2)
- Hong YS, Korman SH, Lee J, Ghoshal P, Qu Q, Barash V, Kang S, Oh S, Kwon M, Gutman A, Rachmel A, Patel MS (2003) Identification of a common mutation (Gly194Cys) in both Arab Moslem and Ashkenazi Jewish patients with dihydrolipoamide dehydrogenase (E3) deficiency: Possible beneficial effect of vitamin therapy. *J Inher Metab Dis* 26(8):816–818
- Horga A, Woodward CE, Mills A, Parees I, Hargreaves IP, Brown RM, Bugiardini E, Brooks T, Manole A, Remzova E, Rahman S, Reilly MM, Houlden H, Sweeney MG, Brown GK, Polke JM, Gago F, Parton MJ, Pitceathly RDS, Hanna MG (2019) Differential phenotypic expression of a novel PDHA1 mutation in a female monozygotic twin pair. *Hum Genet* 138(11-12):1313–1322. <https://doi.org/10.1007/s00439-019-02075-9>
- Horvath G, Svab G, Komlodi T, Ravasz D, Kacso G, Doczi J, Chinopoulos C, Ambrus A, Tretter L (2022) Reverse and Forward Electron Flow-Induced H₂O₂ Formation Is Decreased in alpha-Ketoglutarate Dehydrogenase (alpha-KGDH) Subunit (E2 or E3) Heterozygote Knock Out Animals. *Antioxidants (Basel)* 11(8). <https://doi.org/10.3390/antiox11081487>
- Huang B, Wu P, Bowker-Kinley MM, Harris RA (2002) Regulation of pyruvate dehydrogenase kinase expression by peroxisome proliferator-activated receptor-alpha ligands, glucocorticoids, and insulin. *Diabetes* 51(2):276–283. <https://doi.org/10.2337/diabetes.51.2.276>
- Humphries KM, Szweda LI (1998) Selective inactivation of alpha-ketoglutarate dehydrogenase and pyruvate dehydrogenase: reaction of lipoic acid with 4-hydroxy-2-nonenal. *Biochemistry* 37(45):15835–15841. <https://doi.org/10.1021/bi981512h>
- Hung KC, Nguyen NT, Sun YL, Huang SL (2019) Bio-Fenton reaction involved in the cleavage of the ethoxylate chain of nonionic surfactants by dihydrolipoamide dehydrogenase from *Pseudomonas nitroreducens* TX1. *Sci Rep* 9(1):6827. <https://doi.org/10.1038/s41598-019-43266-8>
- Hurd LR, Collins Y, Abakumova I, Chouchani ET, Baranowski B, Fearnley IM, Prime TA, Murphy MP, James AM (2012) Inactivation of pyruvate dehydrogenase kinase 2 by mitochondrial reactive oxygen species. *J Biol Chem* 287(42):35153–35160. <https://doi.org/10.1074/jbc.M112.400002>
- Hutson SM, Sweatt AJ, Lanoue KF (2005) Branched-chain [corrected] amino acid metabolism: implications for establishing safe intakes. *J Nutr* 135(6 Suppl):1557S–1564S. <https://doi.org/10.1093/jn/135.6.1557S>
- Ibrahim SL, Abed MN, Mohamed G, Price JC, Abdullah MI, Richardson A (2023) Inhibition of branched-chain alpha-keto acid dehydrogenase kinase augments the sensitivity of ovarian

- and breast cancer cells to paclitaxel. *Br J Cancer* 128(5):896–906. <https://doi.org/10.1038/s41416-022-02095-9>
- Igamberdiev AU, Bykova NV, Ens W, Hill RD (2004) Dihydrolipoamide dehydrogenase from porcine heart catalyzes NADH-dependent scavenging of nitric oxide. *FEBS Lett* 568(1-3):146–150
- Ilic N, Birsoy K, Aguirre AJ, Kory N, Pacold ME, Singh S, Moody SE, DeAngelo JD, Spardy NA, Freinkman E, Weir BA, Tsherniak A, Cowley GS, Root DE, Asara JM, Vazquez F, Widlund HR, Sabatini DM, Hahn WC (2017) PIK3CA mutant tumors depend on oxoglutarate dehydrogenase. *Proc Natl Acad Sci U S A* 114(17):E3434–E3443. <https://doi.org/10.1073/pnas.1617922114>
- Islam MM, Wallin R, Wynn RM, Conway M, Fujii H, Mobley JA, Chuang DT, Hutson SM (2007) A novel branched-chain amino acid metabolon. Protein-protein interactions in a supramolecular complex. *J Biol Chem* 282(16):11893–11903. <https://doi.org/10.1074/jbc.M700198200>
- Javadov S, Jang S, Chapa-Dubocq XR, Khuchua Z, Camara AKS (2021) Mitochondrial respiratory supercomplexes in mammalian cells: structural versus functional role. *J Mol Med* 99(1):57–73. <https://doi.org/10.1007/s00109-020-02004-8>
- Jia D, Wang F, Yu H (2023) Systemic alterations of tricarboxylic acid cycle enzymes in Alzheimer's disease. *Front Neurosci* 17:1206688. <https://doi.org/10.3389/fnins.2023.1206688>
- Jiang J, Baiesc FL, Hiromasa Y, Yu X, Hui WH, Dai X, Roche TE, Zhou ZH (2018) Atomic structure of the E2 inner core of human pyruvate dehydrogenase complex. *Biochemistry* 57(16):2325–2334. <https://doi.org/10.1021/acs.biochem.8b00357>
- Johnson MT, Yang HS, Magnuson T, Patel MS (1997) Targeted disruption of the murine dihydrolipoamide dehydrogenase gene (Dld) results in perigastrulation lethality. *Proc Natl Acad Sci U S A* 94(26):14512–14517. <https://doi.org/10.1073/pnas.94.26.14512>
- Jones DD, Perham RN (2008) The role of loop and beta-turn residues as structural and functional determinants for the lipoyl domain from the *Escherichia coli* 2-oxoglutarate dehydrogenase complex. *Biochem J* 409(2):357–366. <https://doi.org/10.1042/bj20071119>
- Jones SM, Yeaman SJ (1986) Oxidative decarboxylation of 4-methylthio-2-oxobutryate by branched-chain 2-oxo acid dehydrogenase complex. *Biochem J* 237(2):621–623. <https://doi.org/10.1042/bj2370621>
- Jones DD, Horne HJ, Reche PA, Perham RN (2000) Structural determinants of post-translational modification and catalytic specificity for the lipoyl domains of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J Mol Biol* 295(2):289–306. <https://doi.org/10.1006/jmbi.1999.3335>
- Jung HI, Cooper A, Perham RN (2003) Interactions of the peripheral subunit-binding domain of the dihydrolipoyl acetyltransferase component in the assembly of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *Eur J Biochem* 270(22):4488–4496. <https://doi.org/10.1046/j.1432-1033.2003.03842.x>
- Kafkia E, Andres-Pons A, Ganter K, Seiler M, Smith TS, Andrejeva A, Jouhten P, Pereira F, Franco C, Kuroshchenkova A, Leone S, Sawarkar R, Boston R, Thaventhiran J, Zaugg JB, Lilley KS, Lancrin C, Beck M, Patil KR (2022) Operation of a TCA cycle subnetwork in the mammalian nucleus. *Sci Adv* 8(35):eabq5206. <https://doi.org/10.1126/sciadv.abq5206>
- Kalse JF, Veeger C (1968) Relation between conformations and activities of lipoamide dehydrogenase. I. Relation between diaphorase and lipoamide dehydrogenase activities upon binding of FAD by the apoenzyme. *Biochim Biophys Acta* 159(2):244–256
- Kaplon J, Zheng L, Meissl K, Chaneton B, Selivanov VA, Mackay G, van der Burg SH, Verdegaa EM, Cascante M, Shlomi T, Gottlieb E, Peeper DS (2013) A key role for mitochondrial gatekeeper pyruvate dehydrogenase in oncogene-induced senescence. *Nature* 498(7452):109–112. <https://doi.org/10.1038/nature12154>
- Kara B, Genc HM, Uyur-Yalcin E, Sakarya-Gunes A, Topcu U, Mulayim S, Ceylaner S (2017) Pyruvate dehydrogenase-E1alpha deficiency presenting as recurrent acute proximal muscle weakness of upper and lower extremities in an 8-year-old boy. *Neuromuscul Disord* 27(1):94–97. <https://doi.org/10.1016/j.nmd.2016.11.001>

- Karagiota A, Kanoura A, Paraskeva E, Simos G, Chachami G (2023) Pyruvate dehydrogenase phosphatase 1 (PDP1) stimulates HIF activity by supporting histone acetylation under hypoxia. *FEBS J* 290(8):2165–2179. <https://doi.org/10.1111/febs.16694>
- Karakaya M, Storbeck M, Strathmann EA, Delle Vedove A, Holker I, Altmueller J, Naghiyeva L, Schmitz-Steinkruger L, Vezyroglou K, Motameny S, Alawbathani S, Thiele H, Polat AI, Okur D, Boostani R, Karimiani EG, Wunderlich G, Ardicli D, Topaloglu H, Kirschner J, Schrank B, Maroofian R, Magnusson O, Yis U, Nurnberg P, Heller R, Wirth B (2018) Targeted sequencing with expanded gene profile enables high diagnostic yield in non-5q-spinal muscular atrophies. *Hum Mutat* 39(9):1284–1298. <https://doi.org/10.1002/humu.23560>
- Kark RA, Rodriguez-Budelli MM (1979) Clinical correlations of partial deficiency of lipoamide dehydrogenase. *Neurology* 29(7):1006–1013. <https://doi.org/10.1212/wnl.29.7.1006>
- Karpova T, Danchuk S, Kolobova E, Popov KM (2003) Characterization of the isozymes of pyruvate dehydrogenase phosphatase: implications for the regulation of pyruvate dehydrogenase activity. *Biochim Biophys Acta* 1652(2):126–135. <https://doi.org/10.1016/j.bbapap.2003.08.010>
- Kato M, Chuang JL, Tso SC, Wynn RM, Chuang DT (2005) Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex. *EMBO J* 24(10):1763–1774. <https://doi.org/10.1038/sj.emboj.7600663>
- Kato M, Wynn RM, Chuang JL, Brautigam CA, Custorio M, Chuang DT (2006) A synchronized substrate-gating mechanism revealed by cubic-core structure of the bovine branched-chain alpha-ketoacid dehydrogenase complex. *The EMBO journal* 25(24):5983–5994. <https://doi.org/10.1038/sj.emboj.7601444>
- Kato M, Li J, Chuang JL, Chuang DT (2007) Distinct structural mechanisms for inhibition of pyruvate dehydrogenase kinase isoforms by AZD7545, dichloroacetate, and radicicol. *Structure* 15(8):992–1004. <https://doi.org/10.1016/j.str.2007.07.001>
- Kato M, Wynn RM, Chuang JL, Tso SC, Machius M, Li J, Chuang DT (2008) Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops. *Structure* 16(12):1849–1859. <https://doi.org/10.1016/j.str.2008.10.010>
- Kee K, Niu LM, Henderson E (1998) A *Tetrahymena thermophila* G4-DNA binding protein with dihydrolipoamide dehydrogenase activity. *Biochemistry* 37(12):4224–4234. <https://doi.org/10.1021/bi9716377>
- Kennerson ML, Yiu EM, Chuang DT, Kidambi A, Tso SC, Ly C, Chaudhry R, Drew AP, Rance G, Delatycki MB, Zuchner S, Ryan MM, Nicholson GA (2013) A new locus for X-linked dominant Charcot-Marie-Tooth disease (CMTX6) is caused by mutations in the pyruvate dehydrogenase kinase isoenzyme 3 (PDK3) gene. *Hum Mol Genet* 22(7):1404–1416. <https://doi.org/10.1093/hmg/dd557>
- Kennerson ML, Kim EJ, Siddell A, Kidambi A, Kim SM, Hong YB, Hwang SH, Chung KW, Choi BO (2016) X-linked Charcot-Marie-Tooth disease type 6 (CMTX6) patients with a p.R158H mutation in the pyruvate dehydrogenase kinase isoenzyme 3 gene. *J Peripher Nerv Syst* 21(1):45–51. <https://doi.org/10.1111/jns.12160>
- Khailova LS, Korochkina LG, Severin SE (1991) Intersite cooperativity in enzyme action of pyruvate-dehydrogenase. *Biochemistry and physiology of thiamin diphosphate enzymes*, pp 251–265
- Kikuchi G (1973) The glycine cleavage system: composition, reaction mechanism, and physiological significance. *Mol Cell Biochem* 1(2):169–187. <https://doi.org/10.1007/bf01659328>
- Kim H (2012) Characterization of two naturally occurring mutations (Gly-101 deletion and Glu-340 to lys substitution) in human dihydrolipoamide dehydrogenase of a patient with metabolic acidosis. *Bull Korean Chem Soc* 33(8):2477–2478. <https://doi.org/10.5012/bkcs.2012.33.8.2477>
- Kim H, Patel MS (1992) Characterization of 2 site specifically mutated human dihydrolipoamide dehydrogenases (His-452-Gln and Glu-457-Gln). *J Biol Chem* 267(8):5128–5132. <https://doi.org/10.0000/jbc.267/8/5128>

- Kim JW, Tchernyshyov I, Semenza GL, Dang CV (2006) HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. *Cell Metab* 3(3):177–185. <https://doi.org/10.1016/j.cmet.2006.02.002>
- Kiss G, Konrad C, Doczi J, Starkov AA, Kawamata H, Manfredi G, Zhang SF, Gibson GE, Beal MF, Adam-Vizi V, Chinopoulos C (2013) The negative impact of alpha-ketoglutarate dehydrogenase complex deficiency on matrix substrate-level phosphorylation. *FASEB J* 27(6):2392–2406. <https://doi.org/10.1096/fj.12-220202>
- Klivenyi P, Starkov AA, Calingasan NY, Gardian G, Browne SE, Yang L, Bubber P, Gibson GE, Patel MS, Beal MF (2004) Mice deficient in dihydrolipoamide dehydrogenase show increased vulnerability to MPTP, malonate and 3-nitropropionic acid neurotoxicity. *J Neurochem* 88(6):1352–1360. <https://doi.org/10.1046/j.1471-4159.2003.02263.x>
- Klyachko NL, Shchedrina VA, Efimov AV, Kazakov SV, Gazaryan IG, Kristal BS, Brown AM (2005) pH-dependent substrate preference of pig heart lipoamide dehydrogenase varies with oligomeric state - Response to mitochondrial matrix acidification. *J Biol Chem* 280(16):16106–16114
- Klyuyeva A, Tuganova A, Kedishvili N, Popov KM (2019) Tissue-specific kinase expression and activity regulate flux through the pyruvate dehydrogenase complex. *J Biol Chem* 294(3):838–851. <https://doi.org/10.1074/jbc.RA118.006433>
- Knapp JE, Mitchell DT, Yazdi MA, Ernst SR, Reed LJ, Hackert ML (1998) Crystal structure of the truncated cubic core component of the Escherichia coli 2-oxoglutarate dehydrogenase multi-enzyme complex. *J Mol Biol* 280(4):655–668
- Ko LW, Sheu KF, Thaler HT, Markesbery WR, Blass JP (2001) Selective loss of KGDHC-enriched neurons in Alzheimer temporal cortex: does mitochondrial variation contribute to selective vulnerability? *J Mol Neurosci* 17(3):361–369. <https://doi.org/10.1385/JMN:17:3:361>
- Kobayashi T, Matsumine H, Matuda S, Mizuno Y (1998) Association between the gene encoding the E2 subunit of the alpha-ketoglutarate dehydrogenase complex and Parkinson's disease. *Ann Neurol* 43(1):120–123. <https://doi.org/10.1002/ana.410430121>
- Kohlschutter A, Behbehani A, Langenbeck U, Albani M, Heidemann P, Hoffmann G, Kleineke J, Lehnert W, Wendel U (1982) A familial progressive neurodegenerative disease with 2-oxoglutaric aciduria. *Eur J Pediatr* 138(1):32–37
- Kolitsida P, Nolic V, Zhou J, Stumpe M, Niemi NM, Dengjel J, Abeliovich H (2023) The pyruvate dehydrogenase complex regulates mitophagic trafficking and protein phosphorylation. *Life Sci Alliance* 6(9). <https://doi.org/10.26508/lsa.202302149>
- Kolobova E, Tuganova A, Boulatnikov I, Popov KM (2001) Regulation of pyruvate dehydrogenase activity through phosphorylation at multiple sites. *Biochem J* 358:69–77. <https://doi.org/10.1042/0264-6021:3580069>
- Korotchikina LG, Patel MS (1995) Mutagenesis studies of the phosphorylation sites of recombinant human pyruvate dehydrogenase. Site-specific regulation. *J Biol Chem* 270(24):14297–14304. <https://doi.org/10.1074/jbc.270.24.14297>
- Kovina MV, Kochetov GA (1998) Cooperativity and flexibility of active sites in homodimeric transketolase. *FEBS Lett* 440(1–2):81–84. [https://doi.org/10.1016/s0014-5793\(98\)01423-9](https://doi.org/10.1016/s0014-5793(98)01423-9)
- Kwon HS, Huang B, Unterman TG, Harris RA (2004) Protein kinase B- α inhibits human pyruvate dehydrogenase kinase-4 gene induction by dexamethasone through inactivation of FOXO transcription factors. *Diabetes* 53(4):899–910. <https://doi.org/10.2337/diabetes.53.4.899>
- Lai J, Cooper A (1986) Brain alpha-ketoglutarate dehydrogenase complex: kinetic properties, regional distribution, and effects of inhibitors. *J Neurochem* 47(5):1376–1386
- Laxmi V, Gunasekaran PK, Kumar A, Manjunathan S, Tiwari S, Saini L (2023) Acute flaccid paralysis due to pyruvate dehydrogenase E1- α deficiency. *Indian J Pediatr*. <https://doi.org/10.1007/s12098-023-04853-w>
- Le BH, Burrage LC (2021) Maple syrup urine disease. <https://www.medlink.com/articles/maple-syrup-urine-disease>.

- Leandro J, Houten SM (2020) The lysine degradation pathway: Subcellular compartmentalization and enzyme deficiencies. *Mol Genet Metab* 131(1-2):14–22. <https://doi.org/10.1016/j.ymgme.2020.07.010>
- Leandro J, Dodatko T, Aten J, Nemeria NS, Zhang X, Jordan F, Hendrickson RC, Sanchez R, Yu C, DeVita RJ, Houten SM (2020a) DHTKD1 and OGDH display substrate overlap in cultured cells and form a hybrid 2-oxo acid dehydrogenase complex in vivo. *Hum Mol Genet* 29(7):1168–1179. <https://doi.org/10.1093/hmg/ddaa037>
- Leandro J, Khamrui S, Wang H, Suebsuwong C, Nemeria NS, Huynh K, Moustakim M, Secor C, Wang M, Dodatko T, Stauffer B, Wilson CG, Yu CL, Arkin MR, Jordan F, Sanchez R, DeVita RJ, Lazarus MB, Houten SM (2020b) Inhibition and crystal structure of the human DHTKD1-thiamin diphosphate complex. *ACS Chem Biol* 15(8):2041–2047. <https://doi.org/10.1021/acscchembio.0c00114>
- Li de la Sierra I, Pernot L, Prange T, Saludjian P, Schiltz M, Fourme R, Padron G (1997) Molecular structure of the lipoamide dehydrogenase domain of a surface antigen from *Neisseria meningitidis*. *J Mol Biol* 269(1):129–141. <https://doi.org/10.1006/jmbi.1997.1009>
- Li J, Wynn RM, Machius M, Chuang JL, Karthikeyan S, Tomchick DR, Chuang DT (2004) Cross-talk between thiamin diphosphate binding and phosphorylation loop conformation in human branched-chain alpha-keto acid decarboxylase/dehydrogenase. *J Biol Chem* 279(31):32968–32978. <https://doi.org/10.1074/jbc.M403611200>
- Li J, Machius M, Chuang JL, Wynn RM, Chuang DT (2007) The two active sites in human branched-chain alpha-keto acid dehydrogenase operate independently without an obligatory alternating-site mechanism. *J Biol Chem* 282(16):11904–11913. <https://doi.org/10.1074/jbc.M610843200>
- Lim J, Liu Z, Apontes P, Feng D, Pessin JE, Sauve AA, Angeletti RH, Chi Y (2014) Dual mode action of mangiferin in mouse liver under high fat diet. *PLoS One* 9(3):e90137. <https://doi.org/10.1371/journal.pone.0090137>
- Lindqvist Y, Schneider G, Ermler U, Sundström M (1992) Three-dimensional structure of transketolase, a thiamine diphosphate dependent enzyme, at 2.5 Å resolution. *The EMBO journal* 11(7):2373–2379. <https://doi.org/10.1002/j.1460-2075.1992.tb05301.x>
- Liu TC, Kim H, Arizmendi C, Kitano A, Patel MS (1993) Identification of two missense mutations in a dihydrolipoamide dehydrogenase-deficient patient. *Proc Natl Acad Sci U S A* 90(11):5186–5190
- Liu TC, Korotchikina LG, Hyatt SL, Vettakkorumakankav NN, Patel MS (1995) Spectroscopic studies of the characterization of recombinant human dihydrolipoamide dehydrogenase and its side-directed mutants. *J Biol Chem* 270(26):15545–15550
- Liu Y, Fiskum G, Schubert D (2002) Generation of reactive oxygen species by the mitochondrial electron transport chain. *J Neurochem* 80(5):780–787. <https://doi.org/10.1046/j.0022-3042.2002.00744.x>
- Liu X, Si W, He L, Yang J, Peng Y, Ren J, Liu X, Jin T, Yu H, Zhang Z, Cheng X, Zhang W, Xia L, Huang Y, Wang Y, Liu S, Shan L, Zhang Y, Yang X, Li H, Liang J, Sun L, Shang Y (2021) The existence of a nonclassical TCA cycle in the nucleus that wires the metabolic-epigenetic circuitry. *Signal Transduct Target Ther* 6(1):375. <https://doi.org/10.1038/s41392-021-00774-2>
- Liu S, Xia X, Zhen J, Li Z, Zhou ZH (2022) Structures and comparison of endogenous 2-oxoglutarate and pyruvate dehydrogenase complexes from bovine kidney. *Cell Discov* 8(1):126. <https://doi.org/10.1038/s41421-022-00487-y>
- Liu S, Kormos BL, Knafels JD, Sahasrabudhe PV, Rosado A, Sommese RF, Reyes AR, Ward J, Roth Flach RJ, Wang X, Buzon LM, Reese MR, Bhattacharya SK, Omoto K, Filipinski KJ (2023) Structural studies identify angiotensin II receptor blocker-like compounds as branched-chain ketoacid dehydrogenase kinase inhibitors. *J Biol Chem* 299(3):102959. <https://doi.org/10.1016/j.jbc.2023.102959>
- Losman JA, Koivunen P, Kaelin WG Jr (2020) 2-Oxoglutarate-dependent dioxygenases in cancer. *Nat Rev Cancer* 20(12):710–726. <https://doi.org/10.1038/s41568-020-00303-3>

- Lu CW, Lin SC, Chen KF, Lai YY, Tsai SJ (2008) Induction of pyruvate dehydrogenase kinase-3 by hypoxia-inducible factor-1 promotes metabolic switch and drug resistance. *J Biol Chem* 283(42):28106–28114. <https://doi.org/10.1074/jbc.M803508200>
- Luan CJ, Guo W, Chen L, Wei XW, He Y, Chen Y, Dang SY, Prior R, Li X, Kuang Y, Wang ZG, Van Den Bosch L, Gu MM (2020) CMT2Q-causing mutation in the Dhtkd1 gene lead to sensory defects, mitochondrial accumulation and altered metabolism in a knock-in mouse model. *Acta Neuropathol Commun* 8(1):32. <https://doi.org/10.1186/s40478-020-00901-0>
- Lyubarev AE, Kurganov BI (1989) Supramolecular organization of tricarboxylic-acid cycle enzymes. *Biosystems* 22(2):91–102
- Machius M, Chuang JL, Wynn RM, Tomchick DR, Chuang DT (2001) Structure of rat BCKD kinase: nucleotide-induced domain communication in a mitochondrial protein kinase. *Proc Natl Acad Sci U S A* 98(20):11218–11223. <https://doi.org/10.1073/pnas.201220098>
- Machius M, Wynn RM, Chuang JL, Li J, Kluger R, Yu D, Tomchick DR, Brautigam CA, Chuang DT (2006) A versatile conformational switch regulates reactivity in human branched-chain alpha-ketoacid dehydrogenase. *Structure* 14(2):287–298. <https://doi.org/10.1016/j.str.2005.10.009>
- Maguolo A, Rodella G, Giorgetti A, Nicolodi M, Ribeiro R, Dianin A, Cantalupo G, Monge I, Carcereri S, De Bernardi ML, Delledonne M, Pasini A, Campostrini N, Ion Popa F, Piacentini G, Teofoli F, Vincenzi M, Camilot M, Bordugo A (2022) A gain-of-function mutation on BCKDK gene and its possible pathogenic role in branched-chain amino acid metabolism. *Genes (Basel)* 13(2). doi:<https://doi.org/10.3390/genes13020233>
- Maj MC, MacKay N, Levandovskiy V, Addis J, Baumgartner ER, Baumgartner MR, Robinson BH, Cameron JM (2005) Pyruvate dehydrogenase phosphatase deficiency: identification of the first mutation in two brothers and restoration of activity by protein complementation. *J Clin Endocrinol Metab* 90(7):4101–4107. <https://doi.org/10.1210/jc.2005-0123>
- Maj MC, Cameron JM, Robinson BH (2006) Pyruvate dehydrogenase phosphatase deficiency: orphan disease or an under-diagnosed condition? *Mol Cell Endocrinol* 249(1-2):1–9. <https://doi.org/10.1016/j.mce.2006.02.003>
- Maksymiuk C, Balakrishnan A, Bryk R, Rhee KY, Nathan CF (2015) E1 of alpha-ketoglutarate dehydrogenase defends Mycobacterium tuberculosis against glutamate anaplerosis and nitrooxidative stress. *Proc Natl Acad Sci U S A* 112(43):E5834–E5843. <https://doi.org/10.1073/pnas.1510932112>
- Mande SS, Sarfaty S, Allen MD, Perham RN, Hol WG (1996) Protein-protein interactions in the pyruvate dehydrogenase multienzyme complex: dihydrolipoamide dehydrogenase complexed with the binding domain of dihydrolipoamide acetyltransferase. *Structure* 4(3):277–286. [https://doi.org/10.1016/S0969-2126\(96\)00032-9](https://doi.org/10.1016/S0969-2126(96)00032-9)
- Marshall L, DiGeorge A (1981) Maple syrup urine disease in the old order Mennonites. *Am J Hum Genet* 6, A139–A139. University of Chicago Press, Chicago, IL
- Massey V (1960a) The composition of the ketoglutarate dehydrogenase complex. *Biochim Biophys Acta* 38:447–460
- Massey V (1960b) The identity of diaphorase and lipoyl dehydrogenase. *Biochim Biophys Acta* 37:314–322
- Massey V, Veeger C (1961) Studies on the reaction mechanism of lipoyl dehydrogenase. *Biochim Biophys Acta* 48:33–47. [https://doi.org/10.1016/0006-3002\(61\)90512-1](https://doi.org/10.1016/0006-3002(61)90512-1)
- Massey V, Gibson QH, Veeger C (1960) Intermediates in the catalytic action of lipoyl dehydrogenase (diaphorase). *Biochem J* 77:341–351
- Mastrogriacomo F, Bergeron C, Kish SJ (1993) Brain alpha-ketoglutarate dehydrogenase complex activity in Alzheimer's disease. *J Neurochem* 61(6):2007–2014. <https://doi.org/10.1111/j.1471-4159.1993.tb07436.x>
- Mastrogriacomo F, LaMarche J, Dozic S, Lindsay G, Bettendorff L, Robitaille Y, Schut L, Kish SJ (1996) Immunoreactive levels of alpha-ketoglutarate dehydrogenase subunits in Friedreich's ataxia and spinocerebellar ataxia type 1. *Neurodegeneration* 5(1):27–33. <https://doi.org/10.1006/neur.1996.0004>

- Mathias RA, Greco TM, Oberstein A, Budayeva HG, Chakrabarti R, Rowland EA, Kang YB, Shenk T, Cristea IM (2014) Sirtuin 4 is a lipoamidase regulating pyruvate dehydrogenase complex activity. *Cell* 159(7):1615–1625. <https://doi.org/10.1016/j.cell.2014.11.046>
- Mattevi A, Schierbeek AJ, Hol WG (1991) Refined crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 Å resolution. A comparison with the structure of glutathione reductase. *J Mol Biol* 220(4):975–994. [https://doi.org/10.1016/0022-2836\(91\)90367-F](https://doi.org/10.1016/0022-2836(91)90367-F)
- Mattevi A, Obmolova G, Schulze E, Kalk KH, Westphal AH, de Kok A, Hol WG (1992a) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* (New York, NY) 255(5051):1544–1550
- Mattevi A, Obmolova G, Sokatch JR, Betzel C, Hol WGJ (1992b) The refined crystal structure of *Pseudomonas putida* lipoamide dehydrogenase complexed with NAD⁺ at 2.45 Å resolution. *Proteins Struct Funct Genet* 13(4):336–351. <https://doi.org/10.1002/prot.340130406>
- Mattevi A, Obmolova G, Kalk KH, Teplyakov A, Hol WG (1993a) Crystallographic analysis of substrate binding and catalysis in dihydrolipoyl transacetylase (E2p). *Biochemistry* 32(15):3887–3901. <https://doi.org/10.1021/bi00066a007>
- Mattevi A, Obmolova G, Kalk KH, van Berkel WJ, Hol WG (1993b) Three-dimensional structure of lipoamide dehydrogenase from *Pseudomonas fluorescens* at 2.8 Å resolution. Analysis of redox and thermostability properties. *J Mol Biol* 230(4):1200–1215. <https://doi.org/10.1006/jmbi.1993.1236>
- Matthews PM, Marchington DR, Squier M, Land J, Brown RM, Brown GK (1993) Molecular genetic characterization of an X-linked form of Leigh's syndrome. *Ann Neurol* 33(6):652–655. <https://doi.org/10.1002/ana.410330616>
- Matthews PM, Brown RM, Otero LJ, Marchington DR, LeGris M, Howes R, Meadows LS, Shevell M, Scriver CR, Brown GK (1994) Pyruvate dehydrogenase deficiency. Clinical presentation and molecular genetic characterization of five new patients. *Brain* 117(Pt 3):435–443. <https://doi.org/10.1093/brain/117.3.435>
- McCartney RG, Rice JE, Sanderson SJ, Bunik V, Lindsay H, Lindsay JG (1998) Subunit interactions in the mammalian alpha-ketoglutarate dehydrogenase complex - Evidence for direct association of the alpha-ketoglutarate dehydrogenase and dihydrolipoamide dehydrogenase components. *J Biol Chem* 273(37):24158–24164. <https://doi.org/10.1074/jbc.273.37.24158>
- McCormack SE, Shaham O, McCarthy MA, Deik AA, Wang TJ, Gerszten RE, Clish CB, Mootha VK, Grinspoon SK, Fleischman A (2013) Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. *Pediatr Obes* 8(1):52–61. <https://doi.org/10.1111/j.2047-6310.2012.00087.x>
- McDonald TS, Carrasco-Pozo C, Hodson MP, Borges K (2017) Alterations in cytosolic and mitochondrial [U-(13C)]glucose metabolism in a chronic epilepsy mouse model. *eNeuro* 4(1). <https://doi.org/10.1523/ENEURO.0341-16.2017>
- McLain AL, Szweda PA, Szweda LI (2011) alpha-Ketoglutarate dehydrogenase: a mitochondrial redox sensor. *Free Radic Res* 45(1):29–36. <https://doi.org/10.3109/10715762.2010.534163>
- McWilliam CA, Ridout CK, Brown RM, McWilliam RC, Tolmie J, Brown GK (2010) Pyruvate dehydrogenase E2 deficiency: a potentially treatable cause of episodic dystonia. *Eur J Paediatr Neurol* 14(4):349–353. <https://doi.org/10.1016/j.ejpn.2009.11.001>
- Melancon SB, Potier M, Dallaire L, Rollin P, Fontaine G, Grenier B (1979) Pyruvate dehydrogenase, lipoamide dehydrogenase and citrate synthase activity in fibroblasts from patients with Friedreich's and Charlevoix-Saguenay ataxia. *Can J Neurol Sci* 6(2):241–242. <https://doi.org/10.1017/s0317167100119717>
- Meng M, Chuang DT (1994) Site-directed mutagenesis and functional analysis of the active-site residues of the E2 component of bovine branched-chain alpha-keto acid dehydrogenase complex. *Biochemistry* 33(43):12879–12885. <https://doi.org/10.1021/bi00209a020>
- Menkes JH, Hurst PL, Craig JM (1954) A new syndrome: progressive familial infantile cerebral dysfunction associated with an unusual urinary substance. *Pediatr* 14(5):462–467
- Michelakis ED, Sutendra G, Dromparis P, Webster L, Haromy A, Niven E, Maguire C, Gammer TL, Mackey JR, Fulton D, Abdulkarim B, McMurtry MS, Petruk KC (2010) Metabolic

- modulation of glioblastoma with dichloroacetate. *Sci Transl Med* 2(31):31ra34. <https://doi.org/10.1126/scitranslmed.3000677>
- Miki Y, Tanji K, Mori F, Kakita A, Takahashi H, Wakabayashi K (2017) Alteration of mitochondrial protein PDHA1 in Lewy body disease and PARK14. *Biochem Biophys Res Commun* 489(4):439–444. <https://doi.org/10.1016/j.bbrc.2017.05.162>
- Milne JL, Shi D, Rosenthal PB, Sunshine JS, Domingo GJ, Wu X, Brooks BR, Perham RN, Henderson R, Subramaniam S (2002) Molecular architecture and mechanism of an icosahedral pyruvate dehydrogenase complex: a multifunctional catalytic machine. *EMBO J* 21(21):5587–5598. <https://doi.org/10.1093/emboj/cdf574>
- Milne JL, Wu X, Borgnia MJ, Lengyel JS, Brooks BR, Shi D, Perham RN, Subramaniam S (2006) Molecular structure of a 9-MDa icosahedral pyruvate dehydrogenase subcomplex containing the E2 and E3 enzymes using cryoelectron microscopy. *J Biol Chem* 281(7):4364–4670. <https://doi.org/10.1074/jbc.M504363200>
- Mizuno Y, Matuda S, Yoshino H, Mori H, Hattori N, Ikebe S (1994) An immunohistochemical study on alpha-ketoglutarate dehydrogenase complex in Parkinson's disease. *Ann Neurol* 35(2):204–210. <https://doi.org/10.1002/ana.410350212>
- Mkrtchyan GV, Ucal M, Mullebner A, Dumitrescu S, Kames M, Moldzio R, Molcanyi M, Schaefer S, Weidinger A, Schaefer U, Hescheler J, Duvigneau JC, Redl H, Bunik VI, Kozlov AV (2018) Thiamine preserves mitochondrial function in a rat model of traumatic brain injury, preventing inactivation of the 2-oxoglutarate dehydrogenase complex. *Biochim Biophys Acta Bioenerg* 1859 9:925–931. <https://doi.org/10.1016/j.bbabi.2018.05.005>
- Moro T, Ebert SM, Adams CM, Rasmussen BB (2016) Amino acid sensing in skeletal muscle. *Trends Endocrinol Metab* 27(11):796–806. <https://doi.org/10.1016/j.tem.2016.06.010>
- Morrison AJ (2022) Cancer cell metabolism connects epigenetic modifications to transcriptional regulation. *FEBS J* 289(5):1302–1314. <https://doi.org/10.1111/febs.16032>
- Moxley MA, Beard DA, Bazil JN (2014) A pH-dependent kinetic model of dihydrolipoamide dehydrogenase from multiple organisms. *Biophys J* 107(12):2984–2998. <https://doi.org/10.1016/j.bpj.2014.09.025>
- Murin R, Mohammadi G, Leibfritz D, Hamprecht B (2009a) Glial metabolism of isoleucine. *Neurochem Res* 34(1):194–204. <https://doi.org/10.1007/s11064-008-9840-4>
- Murin R, Mohammadi G, Leibfritz D, Hamprecht B (2009b) Glial metabolism of valine. *Neurochem Res* 34(7):1195–1203. <https://doi.org/10.1007/s11064-008-9895-2>
- Nadler ST, Stoeckl JP, Schueler KL, Tanimoto G, Yandell BS, Attie AD (2000) The expression of adipogenic genes is decreased in obesity and diabetes mellitus. *Proc Natl Acad Sci U S A* 97(21):11371–11376. <https://doi.org/10.1073/pnas.97.21.11371>
- Nagaraj R, Sharpley MS, Chi F, Braas D, Zhou Y, Kim R, Clark AT, Banerjee U (2017) Nuclear localization of mitochondrial TCA cycle enzymes as a critical step in mammalian zygotic genome activation. *Cell* 168(1–2):210–223 e211. <https://doi.org/10.1016/j.cell.2016.12.026>
- Nagy B, Polak M, Ozohanics O, Zambo Z, Szabo E, Hubert A, Jordan F, Novacek J, Adam-Vizi V, Ambrus A (2021) Structure of the dihydrolipoamide succinyltransferase (E2) component of the human alpha-ketoglutarate dehydrogenase complex (hKGDHc) revealed by cryo-EM and cross-linking mass spectrometry: Implications for the overall hKGDHc structure. *Biochim Biophys Acta Gen Subj* 1865 6:129889. <https://doi.org/10.1016/j.bbagen.2021.129889>
- Naia L, Cunha-Oliveira T, Rodrigues J, Rosenstock TR, Oliveira A, Ribeiro M, Carmo C, Oliveira-Sousa SI, Duarte AI, Hayden MR, Rego AC (2017) Histone deacetylase inhibitors protect against pyruvate dehydrogenase dysfunction in Huntington's disease. *J Neurosci* 37(10):2776–2794. <https://doi.org/10.1523/JNEUROSCI.2006-14.2016>
- Naito E, Ito M, Yokota I, Saijo T, Matsuda J, Osaka H, Kimura S, Kuroda Y (1997) Biochemical and molecular analysis of an X-linked case of Leigh syndrome associated with thiamin-responsive pyruvate dehydrogenase deficiency. *J Inher Metab Dis* 20(4):539–548. <https://doi.org/10.1023/a:1005305614374>
- Nakai T, Nakagawa N, Maoka N, Masui R, Kuramitsu S, Kamiya N (2004) Ligand-induced conformational changes and a reaction intermediate in branched-chain 2-oxo acid dehydroge-

- nase (E1) from *Thermus thermophilus* HB8, as revealed by X-ray crystallography. *J Mol Biol* 337(4):1011–1033. <https://doi.org/10.1016/j.jmb.2004.02.011>
- Nakai T, Kuramitsu S, Kamiya N (2008) Structural bases for the specific interactions between the E2 and E3 components of the *Thermus thermophilus* 2-oxo acid dehydrogenase complexes. *J Biochem* 143(6):747–758. <https://doi.org/10.1093/jb/mvn033>
- Nakano K, Matuda S, Yamanaka T, Tsubouchi H, Nakagawa S, Titani K, Ohta S, Miyata T (1991) Purification and molecular cloning of succinyltransferase of the rat α -ketoglutarate dehydrogenase complex. Absence of a sequence motif of the putative E3 and/or E1 binding site. *J Biol Chem* 266(28):19013–19017
- Nakano K, Takase C, Sakamoto T, Nakagawa S, Inazawa J, Ohta S, Matuda S (1994) Isolation, characterization and structural organization of the gene and pseudogene for the dihydroli-poamide succinyltransferase component of the human 2-oxoglutarate dehydrogenase complex. *Eur J Biochem* 224(1):179–189
- Narayanan RK, Brewer MH, Perez-Siles G, Ellis M, Ly C, Burgess A, Neumann B, Nicholson GA, Vucic S, Kennerson ML (2021) Charcot-Marie-tooth disease causing mutation (p.R158H) in pyruvate dehydrogenase kinase 3 (PDK3) affects synaptic transmission, ATP production and causes neurodegeneration in a CMTX6 *C. elegans* model. *Hum Mol Genet* 31(1):133–145. <https://doi.org/10.1093/hmg/ddab228>
- Naseri NN, Xu H, Bonica J, Vonsattel JP, Cortes EP, Park LC, Arjomand J, Gibson GE (2015) Abnormalities in the tricarboxylic Acid cycle in Huntington disease and in a Huntington disease mouse model. *J Neuropathol Exp Neurol* 74(6):527–537. <https://doi.org/10.1097/NEN.0000000000000197>
- Neinast M, Murashige D, Arany Z (2019) Branched chain amino acids. *Annu Rev Physiol* 81:139–164. <https://doi.org/10.1146/annurev-physiol-020518-114455>
- Nemeria NS, Arjunan P, Chandrasekhar K, Mossad M, Tittmann K, Furey W, Jordan F (2010) Communication between thiamin cofactors in the *Escherichia coli* pyruvate dehydrogenase complex E1 component active centers: evidence for a “direct pathway” between the 4'-aminopyrimidine N1' atoms. *J Biol Chem* 285(15):11197–11209. <https://doi.org/10.1074/jbc.M109.069179>
- Nemeria NS, Ambrus A, Patel H, Gerfen G, Adam-Vizi V, Tretter L, Zhou J, Wang J, Jordan F (2014) Human 2-oxoglutarate dehydrogenase complex E1 component forms a thiamin-derived radical by aerobic oxidation of the enamine intermediate. *J Biol Chem* 289(43):29859–29873. <https://doi.org/10.1074/jbc.M114.591073>
- Nemeria NS, Gerfen G, Nareddy PR, Yang LY, Zhang X, Szostak M, Jordan F (2018) The mitochondrial 2-oxoadipate and 2-oxoglutarate dehydrogenase complexes share their E2 and E3 components for their function and both generate reactive oxygen species. *Free Radic Biol Med* 115:136–145. <https://doi.org/10.1016/j.freeradbiomed.2017.11.018>
- Nemeria NS, Zhang X, Leandro J, Zhou JY, Yang LY, Houten SM, Jordan F (2021) Toward an understanding of the structural and mechanistic aspects of protein-protein interactions in 2-oxo-acid dehydrogenase complexes. *Life-Basel* 11(5):24. <https://doi.org/10.3390/life11050407>
- Nemeria NS, Nagy B, Sanchez R, Zhang X, Leandro J, Ambrus A, Houten SM, Jordan F (2022) Functional versatility of the human 2-oxoadipate dehydrogenase in the L-lysine degradation pathway toward its non-cognate substrate 2-oxopimelic acid. *Int J Mol Sci* 23(15):20. <https://doi.org/10.3390/ijms23158213>
- Ng F, Tang BL (2014) Pyruvate dehydrogenase complex (PDC) export from the mitochondrial matrix. *Mol Membr Biol* 31(7-8):207–210. <https://doi.org/10.3109/09687688.2014.987183>
- Nguyen NH, Gonzalez SV, Hassel B (2007) Formation of glycerol from glucose in rat brain and cultured brain cells. Augmentation with kainate or ischemia. *J Neurochem* 101(6):1694–1700. <https://doi.org/10.1111/j.1471-4159.2006.04433.x>
- Noguchi S, Kondo Y, Ito R, Katayama T, Kazama S, Kadota Y, Kitaura Y, Harris RA, Shimomura Y (2018) Ca(2+)-dependent inhibition of branched-chain α -ketoacid dehydrogenase kinase by thiamine pyrophosphate. *Biochem Biophys Res Commun* 504(4):916–920. <https://doi.org/10.1016/j.bbrc.2018.09.038>

- Nulton-Persson AC, Starke DW, Mieyal JJ, Szweda LI (2003) Reversible inactivation of alpha-ketoglutarate dehydrogenase in response to alterations in the mitochondrial glutathione status. *Biochemistry* 42(14):4235–4242. <https://doi.org/10.1021/bi027370f>
- Odievre MH, Chretien D, Munnich A, Robinson BH, Dumoulin R, Masmoudi S, Kadhom N, Rotig A, Rustin P, Bonnefont JP (2005) A novel mutation in the dihydrolipoamide dehydrogenase E3 subunit gene (DLD) resulting in an atypical form of alpha-ketoglutarate dehydrogenase deficiency. *Hum Mutat* 25(3):323–324. <https://doi.org/10.1002/humu.9319>
- Okajima K, Korotchkina LG, Prasad C, Rupar T, Phillips JA 3rd, Ficicioglu C, Hertecant J, Patel MS, Kerr DS (2008) Mutations of the E1beta subunit gene (PDHB) in four families with pyruvate dehydrogenase deficiency. *Mol Genet Metab* 93(4):371–380. <https://doi.org/10.1016/j.ymgme.2007.10.135>
- Osmanovic A, Gogol I, Martens H, Widjaja M, Muller K, Schreiber-Katz O, Feuerhake F, Langhans CD, Schmidt G, Andersen PM, Ludolph AC, Weishaupt JH, Brand F, Petri S, Weber RG (2021) Heterozygous DHTKD1 variants in two European cohorts of amyotrophic lateral sclerosis patients. *Genes (Basel)* 13(1). <https://doi.org/10.3390/genes13010084>
- Oyarzabal A, Martinez-Pardo M, Merinero B, Navarrete R, Desviat LR, Ugarte M, Rodriguez-Pombo P (2013) A novel regulatory defect in the branched-chain alpha-keto acid dehydrogenase complex due to a mutation in the PPM1K gene causes a mild variant phenotype of maple syrup urine disease. *Hum Mutat* 34(2):355–362. <https://doi.org/10.1002/humu.22242>
- Ozcelik F, Arslan S, Ozguc Caliskan B, Kardas F, Ozkul Y, Dundar M (2023) PPM1K defects cause mild maple syrup urine disease: the second case in the literature. *Am J Med Genet A* 191(5):1360–1365. <https://doi.org/10.1002/ajmg.a.63129>
- Ozohanics O, Zhang X, Nemeria NS, Ambrus A, Jordan F (2023) Probing the E1 α -E2 α and E1 α -E2 α interactions in binary subcomplexes of the human 2-oxoglutarate dehydrogenase and 2-oxoadipate dehydrogenase complexes by chemical cross-linking mass spectrometry and molecular dynamics simulation. *Int J Mol Sci* 24(5). <https://doi.org/10.3390/ijms24054555>
- Pacini G, Carotenuto A, Rentier C, Nuti F, Real-Fernandez F, Brancaccio D, Sabatino G, Larregola M, Peroni E, Migliorini P, Novellino E, Battezzati PM, Selmi C, Papini AM, Rovero P (2015) Role of lipoylation of the immunodominant epitope of pyruvate dehydrogenase complex: toward a peptide-based diagnostic assay for primary biliary cirrhosis. *J Med Chem* 58(16):6619–6629. <https://doi.org/10.1021/acs.jmedchem.5b00783>
- Palmieri EM, Holeywinski R, McGinity CL, Pierri CL, Maio N, Weiss JM, Tragni V, Miranda KM, Rouault TA, Andresson T, Wink DA, McVicar DW (2023) Pyruvate dehydrogenase operates as an intramolecular nitroxyl generator during macrophage metabolic reprogramming. *Nat Commun* 14(1):5114. <https://doi.org/10.1038/s41467-023-40738-4>
- Pandya JD, Musyaju S, Modi HR, Cao Y, Flerlage WJ, Huynh L, Kociuba B, Visavadiya NP, Kobeissy F, Wang K, Gilsdorf JS, Scultetus AH, Shear DA (2023) Comprehensive evaluation of mitochondrial redox profile, calcium dynamics, membrane integrity and apoptosis markers in a preclinical model of severe penetrating traumatic brain injury. *Free Radic Biol Med* 198:44–58. <https://doi.org/10.1016/j.freeradbiomed.2023.02.001>
- Pankotai E, Lacza Z, Muranyi M, Szabo C (2009) Intra-mitochondrial poly(ADP-ribosyl)ation: potential role for alpha-ketoglutarate dehydrogenase. *Mitochondrion* 9(2):159–164. <https://doi.org/10.1016/j.mito.2009.01.013>
- Papandreou I, Cairns RA, Fontana L, Lim AL, Denko NC (2006) HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption. *Cell Metab* 3(3):187–197. <https://doi.org/10.1016/j.cmet.2006.01.012>
- Papatheodorou I, Fonseca NA, Keays M, Tang YA, Barrera E, Bazant W, Burke M, Fullgrabe A, Fuentes AM, George N, Huerta L, Koskinen S, Mohammed S, Geniza M, Preece J, Jaiswal P, Jarnuczak AF, Huber W, Stegle O, Vizcaino JA, Brazma A, Petryszak R (2018) Expression Atlas: gene and protein expression across multiple studies and organisms. *Nucleic Acids Res* 46(D1):D246–D251. <https://doi.org/10.1093/nar/gkx1158>

- Park LC, Calingasan NY, Uchida K, Zhang H, Gibson GE (2000) Metabolic impairment elicits brain cell type-selective changes in oxidative stress and cell death in culture. *J Neurochem* 74(1):114–124. <https://doi.org/10.1046/j.1471-4159.2000.0740114.x>
- Park LC, Albers DS, Xu H, Lindsay JG, Beal MF, Gibson GE (2001) Mitochondrial impairment in the cerebellum of the patients with progressive supranuclear palsy. *J Neurosci Res* 66(5):1028–1034. <https://doi.org/10.1002/jnr.10062>
- Park YH, Wei W, Zhou L, Nemeria N, Jordan F (2004) Amino-terminal residues 1–45 of the *Escherichia coli* pyruvate dehydrogenase complex E1 subunit interact with the E2 subunit and are required for activity of the complex but not for reductive acetylation of the E2 subunit. *Biochemistry* 43(44):14037–14046. <https://doi.org/10.1021/bi049027b>
- Parker PJ, Randle PJ (1978) Partial purification and properties of branched-chain 2-oxo acid dehydrogenase of ox liver. *Biochem J* 171(3):751–757. <https://doi.org/10.1042/bj1710751>
- Patel MS, Korotchikina LG (2001) Regulation of mammalian pyruvate dehydrogenase complex by phosphorylation: complexity of multiple phosphorylation sites and kinases. *Exp Mol Med* 33(4):191–197. <https://doi.org/10.1038/emm.2001.32>
- Patel MS, Korotchikina LG (2006) Regulation of the pyruvate dehydrogenase complex. *Biochem Soc Trans* 34(Pt 2):217–222. <https://doi.org/10.1042/BST20060217>
- Patel TB, Olson MS (1982) Evidence for the regulation of the branched chain alpha-keto acid dehydrogenase multienzyme complex by a phosphorylation/dephosphorylation mechanism. *Biochemistry* 21(18):4259–4265. <https://doi.org/10.1021/bi00261a012>
- Patel MS, Korotchikina LG, Sidhu S (2009) Interaction of E1 and E3 components with the core proteins of the human pyruvate dehydrogenase complex. *J Mol Catal B-Enzym* 61(1–2):2–6. <https://doi.org/10.1016/j.molcatb.2009.05.001>
- Patel KP, O'Brien TW, Subramony SH, Shuster J, Stacpoole PW (2012) The spectrum of pyruvate dehydrogenase complex deficiency: clinical, biochemical and genetic features in 371 patients. *Mol Genet Metab* 106(3):385–394. <https://doi.org/10.1016/j.ymgme.2012.03.017>
- Patel MS, Nemeria NS, Furey W, Jordan F (2014) The pyruvate dehydrogenase complexes: structure-based function and regulation. *J Biol Chem* 289(24):16615–16623. <https://doi.org/10.1074/jbc.R114.563148>
- Pathak D, Srivastava AK, Padma MV, Gulati S, Rajeswari MR (2019) Quantitative proteomic and network analysis of differentially expressed proteins in PBMC of Friedreich's Ataxia (FRDA) patients. *Front Neurosci* 13:1054. <https://doi.org/10.3389/fnins.2019.01054>
- Patrick M, Gu Z, Zhang G, Wynn RM, Kaphle P, Cao H, Vu H, Cai F, Gao X, Zhang Y, Chen M, Ni M, Chuang DT, DeBerardinis RJ, Xu J (2022) Metabolon formation regulates branched-chain amino acid oxidation and homeostasis. *Nat Metab* 4(12):1775–1791. <https://doi.org/10.1038/s42255-022-00689-4>
- Pavlu-Pereira H, Silva MJ, Florindo C, Sequeira S, Ferreira AC, Duarte S, Rodrigues AL, Janeiro P, Oliveira A, Gomes D, Bandeira A, Martins E, Gomes R, Soares S, Tavares de Almeida I, Vicente JB, Rivera I (2020) Pyruvate dehydrogenase complex deficiency: updating the clinical, metabolic and mutational landscapes in a cohort of Portuguese patients. *Orphanet J Rare Dis* 15(1):298. <https://doi.org/10.1186/s13023-020-01586-3>
- Pavlu-Pereira H, Lousa D, Tome CS, Florindo C, Silva MJ, de Almeida IT, Leandro P, Rivera I, Vicente JB (2021) Structural and functional impact of clinically relevant E1alpha variants causing pyruvate dehydrogenase complex deficiency. *Biochimie* 183:78–88. <https://doi.org/10.1016/j.biochi.2021.02.007>
- Perez-Siles G, Ly C, Grant A, Drew AP, Yiu EM, Ryan MM, Chuang DT, Tso SC, Nicholson GA, Kennerson ML (2016) Pathogenic mechanisms underlying X-linked Charcot-Marie-Tooth neuropathy (CMTX6) in patients with a pyruvate dehydrogenase kinase 3 mutation. *Neurobiol Dis* 94:237–244. <https://doi.org/10.1016/j.nbd.2016.07.001>
- Perham RN (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes - a paradigm in the design of a multifunctional protein. *Biochemistry* 30(35):8501–8512. <https://doi.org/10.1021/bi00099a001>

- Pettit FH, Yeaman SJ, Reed LJ (1978) Purification and characterization of branched chain alpha-keto acid dehydrogenase complex of bovine kidney. *Proc Natl Acad Sci U S A* 75(10):4881–4885
- Piroli GG, Manuel AM, McCain RS, Smith HH, Ozohanics O, Mellid S, Cox JH, Cotham WE, Walla MD, Cascon A, Ambrus A, Frizzell N (2023) Defective function of alpha-ketoglutarate dehydrogenase exacerbates mitochondrial ATP deficits during complex I deficiency. *Redox Biol* 67:102932. <https://doi.org/10.1016/j.redox.2023.102932>
- Plokhikh KS, Nesterov SV, Chesnokov YM, Rogov AG, Kamyshinsky RA, Vasiliev AL, Yaguzhinsky LS, Vasilov RG (2023) Association of 2-oxoacid dehydrogenase complexes with respirasomes in mitochondria. *FEBS J*. <https://doi.org/10.1111/febs.16965>
- Plubell DL, Fenton AM, Wilmarth PA, Bergstrom P, Zhao Y, Minnier J, Heinecke JW, Yang X, Pamiir N (2018) GM-CSF driven myeloid cells in adipose tissue link weight gain and insulin resistance via formation of 2-aminoadipate. *Sci Rep* 8(1):11485. <https://doi.org/10.1038/s41598-018-29250-8>
- Prajapati S, Haselbach D, Wittig S, Patel MS, Chari A, Schmidt C, Stark H, Tittmann K (2019) Structural and functional analyses of the human PDH complex suggest a “division-of-labor” mechanism by local E1 and E3 clusters. *Structure* 27(7):1124–1136. e1124. <https://doi.org/10.1016/j.str.2019.04.009>
- Prasad C, Rupar T, Prasad AN (2011) Pyruvate dehydrogenase deficiency and epilepsy. *Brain Dev* 33(10):856–865. <https://doi.org/10.1016/j.braindev.2011.08.003>
- Quinlan CL, Goncalves RL, Hey-Mogensen M, Yadava N, Bunik VI, Brand MD (2014) The 2-oxo-acid dehydrogenase complexes in mitochondria can produce superoxide/hydrogen peroxide at much higher rates than complex I. *J Biol Chem* 289(12):8312–8325. <https://doi.org/10.1074/jbc.M113.545301>
- Quinonez SC, Thoene JG (1993) Dihydrolipoamide dehydrogenase deficiency. In: Adam MP, Feldman J, Mirzaa GM et al (eds) *GeneReviews*(®), Seattle, WA
- Quinonez SC, Thoene JG (2014) Dihydrolipoamide dehydrogenase deficiency. *GeneReviews*® [Internet] 1993-2016. University of Washington, Seattle, WA
- Quinonez SC, Leber SM, Martin DM, Thoene JG, Bedoyan JK (2013) Leigh syndrome in a girl with a novel DLD mutation causing E3 deficiency. *Pediatr Neurol* 48(1):67–72. <https://doi.org/10.1016/j.pediatrneurol.2012.09.013>
- Quintana E, Mayr JA, Garcia Silva MT, Font A, Tortoledo MA, Moliner S, Ozaez L, Lluch M, Cabello A, Ricoy JR, Koch J, Ribes A, Sperl W, Briones P (2009) PDH E1beta deficiency with novel mutations in two patients with Leigh syndrome. *J Inherit Metab Dis* 32(Suppl 1):S339–S343. <https://doi.org/10.1007/s10545-009-1343-1>
- Quintana E, Pineda M, Font A, Vilaseca MA, Tort F, Ribes A, Briones P (2010) Dihydrolipoamide dehydrogenase (DLD) deficiency in a Spanish patient with myopathic presentation due to a new mutation in the interface domain. *J Inherit Metab Dis* 33:S315–S319. <https://doi.org/10.1007/s10545-010-9169-4>
- Rahmatullah M, Gopalakrishnan S, Andrews PC, Chang CL, Radke GA, Roche TE (1989) Subunit associations in the mammalian pyruvate dehydrogenase complex. Structure and role of protein X and the pyruvate dehydrogenase component binding domain of the dihydrolipoyl transacetylase component. *J Biol Chem* 264(4):2221–2227
- Rajashankar KR, Bryk R, Kniewel R, Buglino JA, Nathan CF, Lima CD (2005) Crystal structure and functional analysis of lipoamide dehydrogenase from *Mycobacterium tuberculosis*. *J Biol Chem* 280(40):33977–33983
- Reed LJ (1974) Multienzyme complexes. *Acc Chem Res* 7:40–46. <https://doi.org/10.1021/ar50074a002>
- Reed LJ, Hackert ML (1990) Structure-function relationships in dihydrolipoamide acyltransferases. *J Biol Chem* 265(16):8971–8974
- Reed LJ, Damuni Z, Merryfield ML (1985) Regulation of mammalian pyruvate and branched-chain alpha-keto acid dehydrogenase complexes by phosphorylation-dephosphorylation. *Curr Top Cell Regul* 27:41–49

- Remacha L, Pirman D, Mahoney CE, Coloma J, Calsina B, Curras-Freixes M, Leton R, Torres-Perez R, Richter S, Pita G, Herraiz B, Cianchetta G, Honrado E, Maestre L, Urioste M, Aller J, Garcia-Uriarte O, Galvez MA, Luque RM, Lahera M, Moreno-Rengel C, Eisenhofer G, Montero-Conde C, Rodriguez-Antona C, Llorca O, Smolen GA, Robledo M, Cascon A (2019) Recurrent germline DLST mutations in individuals with multiple pheochromocytomas and paragangliomas. *Am J Hum Genet* 104(5):1008–1010. <https://doi.org/10.1016/j.ajhg.2019.04.010>
- Ricaud PM, Howard MJ, Roberts EL, Broadhurst RW, Perham RN (1996) Three-dimensional structure of the lipoyl domain from the dihydrolipoyl succinyltransferase component of the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. *J Mol Biol* 264(1):179–190
- Rice JE, Dunbar B, Lindsay JG (1992) Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. *EMBO J* 11(9):3229–3235
- Richard AJ, Hang H, Stephens JM (2017) Pyruvate dehydrogenase complex (PDC) subunits moonlight as interaction partners of phosphorylated STAT5 in adipocytes and adipose tissue. *J Biol Chem* 292(48):19733–19742. <https://doi.org/10.1074/jbc.M117.811794>
- Robinson BH, Sherwood WG (1975) Pyruvate dehydrogenase phosphatase deficiency: a cause of congenital chronic lactic acidosis in infancy. *Pediatr Res* 9(12):935–939. <https://doi.org/10.1203/00006450-197512000-00015>
- Robinson JN, Norwitz ER, Mulkern R, Brown SA, Rybicki F, Tempany CMC (2001) Prenatal diagnosis of pyruvate dehydrogenase deficiency using magnetic resonance imaging. *Prenatal Diag* 21(12):1053–1056. <https://doi.org/10.1002/pd.187>
- Roche TE, Baker JC, Yan X, Hiromasa Y, Gong X, Peng T, Dong J, Turkan A, Kasten SA (2001) Distinct regulatory properties of pyruvate dehydrogenase kinase and phosphatase isoforms. *Prog Nucl Acid Res Mol Biol* 70:33–75. [https://doi.org/10.1016/s0079-6603\(01\)70013-x](https://doi.org/10.1016/s0079-6603(01)70013-x)
- Roche TE, Hiromasa Y, Turkan A, Gong X, Peng T, Yan X, Kasten SA, Bao H, Dong J (2003) Essential roles of lipoyl domains in the activated function and control of pyruvate dehydrogenase kinases and phosphatase isoform 1. *Eur J Biochem* 270(6):1050–1056. <https://doi.org/10.1046/j.1432-1033.2003.03468.x>
- Roth Flach RJ, Bollinger E, Reyes AR, Laforest B, Kormos BL, Liu S, Reese MR, Martinez Alsina LA, Buzon L, Zhang Y, Bechle B, Rosado A, Sahasrabudhe PV, Knafels J, Bhattacharya SK, Omoto K, Stansfield JC, Hurley LD, Song L, Luo L, Breitkopf SB, Monetti M, Cunio T, Tierney B, Geoly FJ, Delmore J, Siddall CP, Xue L, Yip KN, Kalgutkar AS, Miller RA, Zhang BB, Filipinski KJ (2023) Small molecule branched-chain ketoacid dehydrogenase kinase (BDK) inhibitors with opposing effects on BDK protein levels. *Nat Commun* 14(1):4812. <https://doi.org/10.1038/s41467-023-40536-y>
- Russell GC, Guest JR (1990) Overexpression of restructured pyruvate dehydrogenase complexes and site-directed mutagenesis of a potential active-site histidine residue. *Biochem J* 269(2):443–450. <https://doi.org/10.1042/bj2690443>
- Sakaguchi Y, Yoshino M, Aramaki S, Yoshida I, Yamashita F, Kuhara T, Matsumoto I, Hayashi T (1986) Dihydrolipoyl dehydrogenase deficiency - a therapeutic trial with branched-chain amino acid restriction. *Eur J Pediatr* 145(4):271–274
- Sanderson SJ, Khan SS, McCartney RG, Miller C, Lindsay JG (1996) Reconstitution of mammalian pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase complexes: Analysis of protein X involvement and interaction of homologous and heterologous dihydrolipoamide dehydrogenases. *Biochem J* 319:109–116
- Schierbeek AJ, Swarte MB, Dijkstra BW, Vriend G, Read RJ, Hol WG, Drenth J, Betzel C (1989) X-ray structure of lipoamide dehydrogenase from *Azotobacter vinelandii* determined by a combination of molecular and isomorphous replacement techniques. *J Mol Biol* 206(2):365–379. [https://doi.org/10.1016/0022-2836\(89\)90486-5](https://doi.org/10.1016/0022-2836(89)90486-5)

- Schiff M, Mine M, Brivet M, Marsac C, Elmaleh-Berges M, Evrard P, Ogier de Baulny H (2006) Leigh's disease due to a new mutation in the PDHX gene. *Ann Neurol* 59(4):709–714. <https://doi.org/10.1002/ana.20818>
- Schoenmann N, Tannenbaum N, Hodgeman RM, Raju RP (2023) Regulating mitochondrial metabolism by targeting pyruvate dehydrogenase with dichloroacetate, a metabolic messenger. *Biochim Biophys Acta Mol Basis Dis* 1869 7:166769. <https://doi.org/10.1016/j.bbadis.2023.166769>
- Schröder-Tittmann K, Meyer D, Arens J, Wechsler C, Tietzel M, Golbik R, Tittmann K (2013) Alternating sites reactivity is a common feature of thiamin diphosphate-dependent enzymes as evidenced by isothermal titration calorimetry studies of substrate binding. *Biochemistry* 52(15):2505–2507. <https://doi.org/10.1021/bi301591e>
- Selmi C, Mackay IR, Gershwin ME (2011) The autoimmunity of primary biliary cirrhosis and the clonal selection theory. *Immunol Cell Biol* 89(1):70–80. <https://doi.org/10.1038/icb.2010.126>
- Sergienko EA, Jordan F (2002) New model for activation of yeast pyruvate decarboxylase by substrate consistent with the alternating sites mechanism: demonstration of the existence of two active forms of the enzyme. *Biochemistry* 41(12):3952–3967. <https://doi.org/10.1021/bi011860a>
- Shaag A, Saada A, Berger I, Mandel H, Joseph A, Feigenbaum A, Elpeleg ON (1999) Molecular basis of lipamide dehydrogenase deficiency in Ashkenazi Jews. *Am J Med Genet* 82(2):177–182
- Shany E, Saada A, Landau D, Shaag A, HersHKovitz E, Elpeleg ON (1999) Lipamide dehydrogenase deficiency due to a novel mutation in the interface domain. *Biochem Biophys Res Commun* 262(1):163–166
- She P, Van Horn C, Reid T, Hutson SM, Cooney RN, Lynch CJ (2007) Obesity-related elevations in plasma leucine are associated with alterations in enzymes involved in branched-chain amino acid metabolism. *Am J Physiol Endocrinol Metab* 293(6):E1552–E1563. <https://doi.org/10.1152/ajpendo.00134.2007>
- Shen N, Korm S, Karantanos T, Li D, Zhang X, Ritou E, Xu H, Lam A, English J, Zong WX, Liu CT, Shirihi O, Feng H (2021) DLST-dependence dictates metabolic heterogeneity in TCA-cycle usage among triple-negative breast cancer. *Commun Biol* 4(1):1289. <https://doi.org/10.1038/s42003-021-02805-8>
- Sherrill JD, Kc K, Wang X, Wen T, Chamberlin A, Stucke EM, Collins MH, Abonia JP, Peng Y, Wu Q, Putnam PE, Dexheimer PJ, Aronow BJ, Kottyan LC, Kaufman KM, Harley JB, Huang T, Rothenberg ME (2018) Whole-exome sequencing uncovers oxidoreductases DHTKD1 and OGDHL as linkers between mitochondrial dysfunction and eosinophilic esophagitis. *JCI Insight* 3(8). <https://doi.org/10.1172/jci.insight.99922>
- Sheu KFR, Blass JP (1999) The alpha-ketoglutarate dehydrogenase complex. In: *Oxidative/Energy metabolism in neurodegenerative disorders*, vol 893. *Annals of the New York Academy of Sciences*. New York Acad Sciences, New York, pp 61–78
- Shim DJ, Nemeria NS, Balakrishnan A, Patel H, Song J, Wang JJ, Jordan F, Farinas ET (2011) Assignment of function to histidines 260 and 298 by engineering the E1 component of the *Escherichia coli* 2-oxoglutarate dehydrogenase complex; Substitutions that lead to acceptance of substrates lacking the 5-carboxyl group. *Biochemistry* 50(35):7705–7709. <https://doi.org/10.1021/bi200936n>
- Shimomura Y, Kuntz MJ, Suzuki M, Ozawa T, Harris RA (1988) Monovalent cations and inorganic phosphate alter branched-chain alpha-ketoacid dehydrogenase-kinase activity and inhibitor sensitivity. *Arch Biochem Biophys* 266(1):210–218. [https://doi.org/10.1016/0003-9861\(88\)90252-4](https://doi.org/10.1016/0003-9861(88)90252-4)
- Shimomura Y, Fujii H, Suzuki M, Murakami T, Fujitsuka N, Nakai N (1995) Branched-chain alpha-keto acid dehydrogenase complex in rat skeletal muscle: regulation of the activity and gene expression by nutrition and physical exercise. *J Nutr* 125(6 Suppl):1762S–1765S. https://doi.org/10.1093/jn/125.suppl_6.1762S

- Shin AC, Fasshauer M, Filatova N, Grundell LA, Zielinski E, Zhou JY, Scherer T, Lindtner C, White PJ, Lapworth AL, Ilkayeva O, Knippschild U, Wolf AM, Scheja L, Grove KL, Smith RD, Qian WJ, Lynch CJ, Newgard CB, Buettner C (2014) Brain insulin lowers circulating BCAA levels by inducing hepatic BCAA catabolism. *Cell Metab* 20(5):898–909. <https://doi.org/10.1016/j.cmet.2014.09.003>
- Shin D, Lee J, You JH, Kim D, Roh JL (2020) Dihydrolipoamide dehydrogenase regulates cystine deprivation-induced ferroptosis in head and neck cancer. *Redox Biol* 30:101418. <https://doi.org/10.1016/j.redox.2019.101418>
- Sies H, Jones DP (2020) Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. *Nat Rev Mol Cell Biol* 21(7):363–383. <https://doi.org/10.1038/s41580-020-0230-3>
- Skerlova J, Berndtsson J, Nolte H, Ott M, Stenmark P (2021) Structure of the native pyruvate dehydrogenase complex reveals the mechanism of substrate insertion. *Nat Commun* 12(1):5277. <https://doi.org/10.1038/s41467-021-25570-y>
- Smolle M, Prior AE, Brown AE, Cooper A, Byron O, Lindsay JG (2006) A new level of architectural complexity in the human pyruvate dehydrogenase complex. *J Biol Chem* 281(28):19772–19780. <https://doi.org/10.1074/jbc.M601140200>
- Sokatch JR, McCully V, Gebrosky J, Sokatch DJ (1981) Isolation of a specific lipoamide dehydrogenase for a branched-chain keto acid dehydrogenase from *Pseudomonas putida*. *J Bacteriol* 148(2):639–646. <https://doi.org/10.1128/jb.148.2.639-646.1981>
- Song JY, Jordan F (2012) Interchain acetyl transfer in the E2 component of bacterial pyruvate dehydrogenase suggests a model with different roles for each chain in a trimer of the homooligomeric component. *Biochemistry* 51(13):2795–2803. <https://doi.org/10.1021/bi201614n>
- Song J, Park YH, Nemeria N, Kale S, Kakalis L, Jordan F (2010) Nuclear magnetic resonance evidence for the role of the flexible regions of the E1 component of the pyruvate dehydrogenase complex from gram-negative bacteria. *J Biol Chem* 285:4680–4694
- Sorbi S, Bird ED, Blass JP (1983) Decreased pyruvate dehydrogenase complex activity in Huntington and Alzheimer brain. *Ann Neurol* 13(1):72–78. <https://doi.org/10.1002/ana.410130116>
- Sperl W, Fleuren L, Freisinger P, Haack TB, Ribes A, Feichtinger RG, Rodenburg RJ, Zimmermann FA, Koch J, Rivera I, Prokisch H, Smeitink JA, Mayr JA (2015) The spectrum of pyruvate oxidation defects in the diagnosis of mitochondrial disorders. *J Inher Metab Dis* 38(3):391–403. <https://doi.org/10.1007/s10545-014-9787-3>
- Spriet LL, Heigenhauser GJ (2002) Regulation of pyruvate dehydrogenase (PDH) activity in human skeletal muscle during exercise. *Exerc Sport Sci Rev* 30(2):91–95. <https://doi.org/10.1097/00003677-200204000-00009>
- Sradhanjali S, Reddy MM (2018) Inhibition of pyruvate dehydrogenase kinase as a therapeutic strategy against cancer. *Curr Top Med Chem* 18(6):444–453. <https://doi.org/10.2174/1568026618666180523105756>
- Srivastava S, Gajwani P, Jousma J, Miyamoto H, Kwon Y, Jana A, Toth PT, Yan G, Ong SG, Rehman J (2023) Nuclear translocation of mitochondrial dehydrogenases as an adaptive cardioprotective mechanism. *Nat Commun* 14(1):4360. <https://doi.org/10.1038/s41467-023-40084-5>
- Stacpoole PW (2017) Therapeutic targeting of the pyruvate dehydrogenase complex/pyruvate dehydrogenase kinase (PDC/PDK) axis in cancer. *J Natl Cancer Inst* 109(11). <https://doi.org/10.1093/jnci/djx071>
- Staretz-Chacham O, Pode-Shakked B, Kristal E, Abraham SY, Porper K, Wormser O, Shelef I, Anikster Y (2021) The effects of a ketogenic diet on patients with dihydrolipoamide dehydrogenase deficiency. *Nutrients* 13(10). <https://doi.org/10.3390/nu13103523>
- Starkov AA, Chinopoulos C, Fiskum G (2004a) Mitochondrial calcium and oxidative stress as mediators of ischemic brain injury. *Cell Calcium* 36(3–4):257–264. <https://doi.org/10.1016/j.ceca.2004.02.012>
- Starkov AA, Fiskum G, Chinopoulos C, Lorenzo BJ, Browne SE, Patel MS, Beal MF (2004b) Mitochondrial alpha-ketoglutarate dehydrogenase complex generates reactive oxygen species. *J Neurosci* 24(36):7779–7788. <https://doi.org/10.1523/JNEUROSCI.1899-04.2004>

- Steussy CN, Popov KM, Bowker-Kinley MM, Sloan RB Jr, Harris RA, Hamilton JA (2001) Structure of pyruvate dehydrogenase kinase. Novel folding pattern for a serine protein kinase. *J Biol Chem* 276(40):37443–37450. <https://doi.org/10.1074/jbc.M104285200>
- Stiles AR, Venturoni L, Mucci G, Elbalalesy N, Woontner M, Goodman S, Abdenur JE (2016) New cases of DHTKD1 mutations in patients with 2-ketoadipic aciduria. *JIMD Rep* 25:15–19. https://doi.org/10.1007/8904_2015_462
- Stoops JK, Cheng RH, Yazdi MA, Maeng CY, Schroeter JP, Klueppelberg U, Kolodziej SJ, Baker TS, Reed LJ (1997) On the unique structural organization of the *Saccharomyces cerevisiae* pyruvate dehydrogenase complex. *J Biol Chem* 272(9):5757–5764
- Straub FB (1939) Isolation and properties of a flavoprotein from heart muscle tissue. *Biochem J* 33(5):787–792
- Strauss KA, Carson VJ, Soltys K, Young ME, Bowser LE, Puffenberger EG, Brigatti KW, Williams KB, Robinson DL, Hendrickson C, Beiler K, Taylor CM, Haas-Givler B, Chopko S, Hailey J, Muelly ER, Shellmer DA, Radcliff Z, Rodrigues A, Loeven K, Heaps AD, Mazariegos GV, Morton DH (2020) Branched-chain alpha-ketoacid dehydrogenase deficiency (maple syrup urine disease): treatment, biomarkers, and outcomes. *Mol Genet Metab* 129(3):193–206. <https://doi.org/10.1016/j.ymgme.2020.01.006>
- Strumilo S (2005) Short-term regulation of the alpha-ketoglutarate dehydrogenase complex by energy-linked and some other effectors. *Biochem-Moscow* 70(7):726–729. <https://doi.org/10.1007/s10541-005-0177-1>
- Sumegi B, Srere PA (1984) Complex I binds several mitochondrial NAD-coupled dehydrogenases. *J Biol Chem* 259(24):15040–15045
- Sun H, Olson KC, Gao C, Prosdocimo DA, Zhou M, Wang Z, Jeyaraj D, Youn JY, Ren S, Liu Y, Rau CD, Shah S, Ilkayeva O, Gui WJ, William NS, Wynn RM, Newgard CB, Cai H, Xiao X, Chuang DT, Schulze PC, Lynch C, Jain MK, Wang Y (2016) Catabolic defect of branched-chain amino acids promotes heart failure. *Circulation* 133(21):2038–2049. <https://doi.org/10.1161/CIRCULATIONAHA.115.020226>
- Supruniuk E, Zebrowska E, Chabowski A (2023) Branched chain amino acids-friend or foe in the control of energy substrate turnover and insulin sensitivity? *Crit Rev Food Sci Nutr* 63(15):2559–2597. <https://doi.org/10.1080/10408398.2021.1977910>
- Suryawan A, Hawes JW, Harris RA, Shimomura Y, Jenkins AE, Hutson SM (1998) A molecular model of human branched-chain amino acid metabolism. *Am J Clin Nutr* 68(1):72–81. <https://doi.org/10.1093/ajcn/68.1.72>
- Sutendra G, Kinnaird A, Dromparis P, Paulin R, Stenson TH, Haromy A, Hashimoto K, Zhang N, Flaim E, Michelakis ED (2014) A nuclear pyruvate dehydrogenase complex is important for the generation of acetyl-CoA and histone acetylation. *Cell* 158(1):84–97. <https://doi.org/10.1016/j.cell.2014.04.046>
- Sweatt AJ, Wood M, Suryawan A, Wallin R, Willingham MC, Hutson SM (2004) Branched-chain amino acid catabolism: unique segregation of pathway enzymes in organ systems and peripheral nerves. *Am J Physiol Endocrinol Metab* 286(1):E64–E76. <https://doi.org/10.1152/ajpendo.00276.2003>
- Szabo E, Ambrus A (2023) Lipoamide dehydrogenase (LADH) deficiency: medical perspectives of the structural and functional characterization of LADH and its pathogenic variants. *Biol Futur*. <https://doi.org/10.1007/s42977-023-00155-6>
- Szabo E, Mizsei R, Wilk P, Zambo Z, Torocsik B, Weiss MS, Adam-Vizi V, Ambrus A (2018) Crystal structures of the disease-causing D444V mutant and the relevant wild type human dihydrolipoamide dehydrogenase. *Free Radic Biol Med* 124:214–220. <https://doi.org/10.1016/j.freeradbiomed.2018.06.008>
- Szabo E, Wilk P, Nagy B, Zambo Z, Bui D, Weichsel A, Arjunan P, Torocsik B, Hubert A, Furey W, Montfort WR, Jordan F, Weiss MS, Adam-Vizi V, Ambrus A (2019) Underlying molecular alterations in human dihydrolipoamide dehydrogenase deficiency revealed by structural analyses of disease-causing enzyme variants. *Hum Mol Genet* 28(20):3339–3354. <https://doi.org/10.1093/hmg/ddz177>

- Szabo E, Nemes-Nikodem E, Vass KR, Zambo Z, Zrupko E, Torocsik B, Ozohanics O, Nagy B, Ambrus A (2023) Structural and biochemical investigation of selected pathogenic mutants of the human dihydrolipoamide dehydrogenase. *Int J Mol Sci* 24(13). <https://doi.org/10.3390/ijms241310826>
- Tahara EB, Barros MH, Oliveira GA, Netto LE, Kowaltowski AJ (2007) Dihydrolipoyl dehydrogenase as a source of reactive oxygen species inhibited by caloric restriction and involved in *Saccharomyces cerevisiae* aging. *FASEB J* 21(1):274–283. <https://doi.org/10.1096/fj.06-6686com>
- Tan HC, Hsu JW, Khoo CM, Tai ES, Yu S, Chacko S, Lai OF, Jahoor F (2018) Alterations in branched-chain amino acid kinetics in nonobese but insulin-resistant Asian men. *Am J Clin Nutr* 108(6):1220–1228. <https://doi.org/10.1093/ajcn/nqy208>
- Tanaka N, Koike K, Hamada M, Otsuka KI, Suematsu T, Koike M (1972) Mammalian α -keto acid dehydrogenase complexes. VII. Resolution and reconstitution of the pig heart 2-oxoglutarate dehydrogenase complex. *J Biol Chem* 247(12):4043–4049
- Tanner LM, Tynnenin O, Piippo K, Puhakka AM (2023) X-linked pyruvate dehydrogenase complex deficiency due to a novel PDHA1 variant associated with structural brain abnormalities in a fetus. *Prenat Diagn* 43(6):730–733. <https://doi.org/10.1002/pd.6349>
- Timmons JA, Atherton PJ, Larsson O, Sood S, Blokhin IO, Brogan RJ, Volmar CH, Josse AR, Slentz C, Wahlestedt C, Phillips SM, Phillips BE, Gallagher IJ, Kraus WE (2018) A coding and non-coding transcriptomic perspective on the genomics of human metabolic disease. *Nucleic Acids Res* 46(15):7772–7792. <https://doi.org/10.1093/nar/gky570>
- Tittmann K (2009) Reaction mechanisms of thiamin diphosphate enzymes: redox reactions. *FEBS J* 276(9):2454–2468. <https://doi.org/10.1111/j.1742-4658.2009.06966.x>
- Toyoda T, Suzuki K, Sekiguchi T, Reed LJ, Takenaka A (1998) Crystal structure of eucaryotic E3, lipoyamide dehydrogenase from yeast. *J Biochem* 123(4):668–674
- Tran TQ, Lowman XH, Kong M (2017) Molecular pathways: metabolic control of histone methylation and gene expression in cancer. *Clin Cancer Res* 23(15):4004–4009. <https://doi.org/10.1158/1078-0432.CCR-16-2506>
- Tretter L, Adam-Vizi V (2004) Generation of reactive oxygen species in the reaction catalyzed by α -ketoglutarate dehydrogenase. *J Neurosci* 24(36):7771–7778. <https://doi.org/10.1523/JNEUROSCI.1842-04.2004>
- Tretter L, Adam-Vizi V (2005) α -ketoglutarate dehydrogenase: a target and generator of oxidative stress. *Philos Trans R Soc Lond B Biol Sci* 360(1464):2335–2345. <https://doi.org/10.1098/rstb.2005.1764>
- Tso SC, Kato M, Chuang JL, Chuang DT (2006) Structural determinants for cross-talk between pyruvate dehydrogenase kinase 3 and lipoyl domain 2 of the human pyruvate dehydrogenase complex. *J Biol Chem* 281(37):27197–27204. <https://doi.org/10.1074/jbc.M604339200>
- Tso SC, Qi X, Gui WJ, Chuang JL, Morlock LK, Wallace AL, Ahmed K, Laxman S, Campeau PM, Lee BH, Hutson SM, Tu BP, Williams NS, Tambar UK, Wynn RM, Chuang DT (2013) Structure-based design and mechanisms of allosteric inhibitors for mitochondrial branched-chain α -ketoacid dehydrogenase kinase. *Proc Natl Acad Sci U S A* 110(24):9728–9733. <https://doi.org/10.1073/pnas.1303220110>
- Tso SC, Gui WJ, Wu CY, Chuang JL, Qi X, Skvora KJ, Dork K, Wallace AL, Morlock LK, Lee BH, Hutson SM, Strom SC, Williams NS, Tambar UK, Wynn RM, Chuang DT (2014) Benzothioephene carboxylate derivatives as novel allosteric inhibitors of branched-chain α -ketoacid dehydrogenase kinase. *J Biol Chem* 289(30):20583–20593. <https://doi.org/10.1074/jbc.M114.569251>
- Tuganova A, Yoder MD, Popov KM (2001) An essential role of Glu-243 and His-239 in the phosphotransfer reaction catalyzed by pyruvate dehydrogenase kinase. *J Biol Chem* 276(21):17994–17999. <https://doi.org/10.1074/jbc.M009327200>
- Tyagi TK, Ponnan P, Singh P, Bansal S, Batra A, Collin F, Guillonau F, Jore D, Patkar SA, Saxena RK, Parmar VS, Rastogi RC, Raj HG (2009) Moonlighting protein in *Starkeyomyces koorchalomoides*: characterization of dihydrolipoamide dehydrogenase as a protein acetyl-

- transferase utilizing acetoxycoumarin as the acetyl group donor. *Biochimie* 91(7):868–875. <https://doi.org/10.1016/j.biochi.2009.04.007>
- Tyx RE, Roche-Hakansson H, Hakansson AP (2011) Role of dihydrolipoamide dehydrogenase in regulation of raffinose transport in *Streptococcus pneumoniae*. *J Bacteriol* 193(14):3512–3524. <https://doi.org/10.1128/JB.01410-10>
- Uddin GM, Zhang L, Shah S, Fukushima A, Wagg CS, Gopal K, Al Batran R, Pherwani S, Ho KL, Boisvenue J, Karwi QG, Altamimi T, Wishart DS, Dyck JRB, Ussher JR, Oudit GY, Lopaschuk GD (2019) Impaired branched chain amino acid oxidation contributes to cardiac insulin resistance in heart failure. *Cardiovasc Diabetol* 18(1):86. <https://doi.org/10.1186/s12933-019-0892-3>
- Vassilyev DG, Symersky J (2007) Crystal structure of pyruvate dehydrogenase phosphatase 1 and its functional implications. *J Mol Biol* 370(3):417–426. <https://doi.org/10.1016/j.jmb.2007.05.002>
- Vatrinet R, Leone G, De Luise M, Girolimetti G, Vidone M, Gasparre G, Porcelli AM (2017) The alpha-ketoglutarate dehydrogenase complex in cancer metabolic plasticity. *Cancer Metab* 5:3. <https://doi.org/10.1186/s40170-017-0165-0>
- Vaubel RA, Rustin P, Isaya G (2011) Mutations in the dimer interface of dihydrolipoamide dehydrogenase promote site-specific oxidative damages in yeast and human cells. *J Biol Chem* 286(46):40232–40245. <https://doi.org/10.1074/jbc.M111.274415>
- Vaz R, Wincent J, Elfissi N, Rosengren Forsblad K, Pettersson M, Naess K, Wedell A, Wredenberg A, Lindstrand A, Ygberg S (2022) A missense variant in PDK1 associated with severe neurodevelopmental delay and epilepsy. *Biomedicines* 10(12). <https://doi.org/10.3390/biomedicines10123171>
- Vijayakrishnan S, Callow P, Nutley MA, McGow DP, Gilbert D, Kropholler P, Cooper A, Byron O, Lindsay JG (2011) Variation in the organization and subunit composition of the mammalian pyruvate dehydrogenase complex E2/E3BP core assembly. *Biochem J* 437:565–574. <https://doi.org/10.1042/bj20101784>
- Voronova V, Sokolov V, Morias Y, Boezelman MJ, Wagberg M, Henricsson M, Hansson K, Goltsov A, Peskov K, Sundqvist M (2022) Evaluation of therapeutic strategies targeting BCAA catabolism using a systems pharmacology model. *Front Pharmacol* 13:993422. <https://doi.org/10.3389/fphar.2022.993422>
- Wagenknecht T, Francis N, Derosier DJ (1983) Alpha-ketoglutarate dehydrogenase complex may be heterogeneous in quaternary structure. *J Mol Biol* 165(3):523–539. [https://doi.org/10.1016/s0022-2836\(83\)80217-4](https://doi.org/10.1016/s0022-2836(83)80217-4)
- Wagenknecht T, Francis N, DeRosier D (1986) Role of excess lipoyl dehydrogenase in reconstituted alpha-ketoglutarate dehydrogenase complex of *Escherichia coli*. *Biochem Biophys Res Commun* 135(3):802–807. [https://doi.org/10.1016/0006-291x\(86\)90999-x](https://doi.org/10.1016/0006-291x(86)90999-x)
- Wagenknecht T, Francis N, DeRosier DJ, Hainfeld JF, Wall JS (1987) Scanning transmission electron microscopic study of alpha-ketoglutarate dehydrogenase complex from *Escherichia coli*. *J Biol Chem* 262(2):877–882
- Wagenknecht T, Grassucci R, Schaak D (1990) Cryoelectron microscopy of frozen-hydrated alpha-ketoacid dehydrogenase complexes from *Escherichia coli*. *J Biol Chem* 265(36):22402–22408
- Wagner T, Bellinzoni M, Wehenkel A, O'Hare HM, Alzari PM (2011) Functional plasticity and allosteric regulation of α -ketoglutarate decarboxylase in central mycobacterial metabolism. *Chem Biol* 18(8):1011–1020. <https://doi.org/10.1016/j.chembiol.2011.06.004>
- Wallis NG, Perham RN (1994) Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J Mol Biol* 236(1):209–216. <https://doi.org/10.1006/jmbi.1994.1130>
- Wang J, Nemeria NS, Chandrasekhar K, Kumaran S, Arjunan P, Reynolds S, Calero G, Brukh R, Kakalis L, Furey W, Jordan F (2014a) Structure and function of the catalytic domain of the dihydrolipoyl acetyltransferase component in *Escherichia coli* pyruvate dehydrogenase complex. *J Biol Chem* 289(22):15215–15230. <https://doi.org/10.1074/jbc.M113.544080>

- Wang L, Wang FS, Chang C, Gershwin ME (2014b) Breach of tolerance: primary biliary cirrhosis. *Semin Liver Dis* 34(3):297–317. <https://doi.org/10.1055/s-0034-1383729>
- Wang JJ, Kumaran S, Zhou JY, Nerneria NS, Tao H, Kakalis L, Park YH, Birkaya B, Patel MS, Jordan F (2015) Elucidation of the interaction loci of the human pyruvate dehydrogenase complex E2.E3BP core with pyruvate dehydrogenase kinase 1 and kinase 2 by HDX exchange mass spectrometry and nuclear magnetic resonance. *Biochemistry* 54(1):69–82. <https://doi.org/10.1021/bi5013113>
- Wang W, Zhang F, Xia Y, Zhao S, Yan W, Wang H, Lee Y, Li C, Zhang L, Lian K, Gao E, Cheng H, Tao L (2016) Defective branched chain amino acid catabolism contributes to cardiac dysfunction and remodeling following myocardial infarction. *Am J Physiol Heart Circ Physiol* 311(5):H1160–H1169. <https://doi.org/10.1152/ajpheart.00114.2016>
- Wang Y, Guo YR, Liu K, Yin Z, Liu R, Xia Y, Tan L, Yang P, Lee JH, Li XJ, Hawke D, Zheng Y, Qian X, Lyu J, He J, Xing D, Tao YJ, Lu Z (2017) KAT2A coupled with the alpha-KGDH complex acts as a histone H3 succinyltransferase. *Nature* 552(7684):273–277. <https://doi.org/10.1038/nature25003>
- Wang C, Calcutt MW, Ferguson JF (2021) Knock-out of DHTKD1 alters mitochondrial respiration and function, and may represent a novel pathway in cardiometabolic disease risk. *Front Endocrinol (Lausanne)* 12:710698. <https://doi.org/10.3389/fendo.2021.710698>
- Wexler ID, Hemalatha SG, Patel MS (1991) Sequence conservation in the alpha and beta subunits of pyruvate dehydrogenase and its similarity to branched-chain alpha-keto acid dehydrogenase. *FEBS Lett* 282(1):209–213. [https://doi.org/10.1016/0014-5793\(91\)80479-m](https://doi.org/10.1016/0014-5793(91)80479-m)
- Wexler ID, Hemalatha SG, McConnell J, Buist NR, Dahl HH, Berry SA, Cederbaum SD, Patel MS, Kerr DS (1997) Outcome of pyruvate dehydrogenase deficiency treated with ketogenic diets. Studies in patients with identical mutations. *Neurology* 49(6):1655–1661. <https://doi.org/10.1212/wnl.49.6.1655>
- Whittle EF, Chilian M, Karimiani EG, Progre H, Buhas D, Kose M, Ganetzky RD, Toosi MB, Torbati PN, Badv RS, Sheliha I, Yang H, Elloumi HZ, Lee S, Jamshidi Y, Pittman AM, Houlden H, Ignatius E, Rahman S, Maroofian R, Yoon WH, Carroll CJ (2023) Biallelic variants in OGDH encoding oxoglutarate dehydrogenase lead to a neurodevelopmental disorder characterized by global developmental delay, movement disorder, and metabolic abnormalities. *Genet Med* 25(2):100332. <https://doi.org/10.1016/j.gim.2022.11.001>
- Wilkinson KD, Williams CH (1981) NADH inhibition and NAD activation of Escherichia coli lipamide dehydrogenase catalyzing the NADH-lipoamide reaction. *J Biol Chem* 256(5):2307–2314
- Williams CH Jr (1992) Lipoamide dehydrogenase, glutathione reductase, thioredoxin reductase, and mercuric ion reductase—a family of flavoenzyme transhydrogenases. In: Müller F (ed) *Chemistry and biochemistry of flavoenzymes*, vol 3. CRC Press, Boca Raton, FL, pp 121–211
- Williams CH Jr, Allison N, Russell GC, Prongay AJ, Arscott LD, Datta S, Sahlman L, Guest JR (1989) Properties of lipamide dehydrogenase and thioredoxin reductase from Escherichia coli altered by site-directed mutagenesis. *Ann N Y Acad Sci* 573:55–65. <https://doi.org/10.1111/j.1749-6632.1989.tb14986.x>
- Wolfson RL, Chantranupong L, Saxton RA, Shen K, Scaria SM, Cantor JR, Sabatini DM (2016) Sestrin2 is a leucine sensor for the mTORC1 pathway. *Science (New York, NY)* 351(6268):43–48. <https://doi.org/10.1126/science.aab2674>
- Wongkittichote P, Cuddapah SR, Master SR, Grange DK, Dietzen D, Roper SM, Ganetzky RD (2023) Biochemical characterization of patients with dihydrolipoamide dehydrogenase deficiency. *JIMD Rep* 64(5):367–374. <https://doi.org/10.1002/jmd2.12382>
- Woolbright BL, Rajendran G, Harris RA, Taylor JA III (2019) Metabolic flexibility in cancer: targeting the pyruvate dehydrogenase kinase:pyruvate dehydrogenase axis. *Mol Cancer Ther* 18(10):1673–1681. <https://doi.org/10.1158/1535-7163.MCT-19-0079>
- Wu Z, He X, Chen S (2021) Oncogenic circDHTKD1 promotes tumor growth and metastasis of oral squamous cell carcinoma in vitro and in vivo via upregulating miR-326-mediated GAB1. *Braz J Med Biol Res* 54(10):e10837. <https://doi.org/10.1590/1414-431X2020e10837>

- Wynn RM, Chuang JL, Davie JR, Fisher CW, Hale MA, Cox RP, Chuang DT (1992) Cloning and expression in *Escherichia coli* of mature E1 beta subunit of bovine mitochondrial branched-chain alpha-keto acid dehydrogenase complex. Mapping of the E1 beta-binding region on E2. *J Biol Chem* 267(3):1881–1887
- Wynn RM, Chuang JL, Cote CD, Chuang DT (2000) Tetrameric assembly and conservation in the ATP-binding domain of rat branched-chain alpha-ketoacid dehydrogenase kinase. *J Biol Chem* 275(39):30512–30519. <https://doi.org/10.1074/jbc.M005075200>
- Wynn RM, Ho R, Chuang JL, Chuang DT (2001) Roles of active site and novel K⁺ ion-binding site residues in human mitochondrial branched-chain alpha-ketoacid decarboxylase/dehydrogenase. *J Biol Chem* 276(6):4168–4174. <https://doi.org/10.1074/jbc.M008038200>
- Wynn RM, Kato M, Machius M, Chuang JL, Li J, Tomchick DR, Chuang DT (2004) Molecular mechanism for regulation of the human mitochondrial branched-chain alpha-ketoacid dehydrogenase complex by phosphorylation. *Structure* 12(12):2185–2196. <https://doi.org/10.1016/j.str.2004.09.013>
- Wynn RM, Kato M, Chuang JL, Tso SC, Li J, Chuang DT (2008) Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity. *J Biol Chem* 283(37):25305–25315. <https://doi.org/10.1074/jbc.M802249200>
- Wynn RM, Li J, Brautigam CA, Chuang JL, Chuang DT (2012) Structural and biochemical characterization of human mitochondrial branched-chain alpha-ketoacid dehydrogenase phosphatase. *J Biol Chem* 287(12):9178–9192. <https://doi.org/10.1074/jbc.M111.314963>
- Xiao F, Yu J, Guo Y, Deng J, Li K, Du Y, Chen S, Zhu J, Sheng H, Guo F (2014) Effects of individual branched-chain amino acids deprivation on insulin sensitivity and glucose metabolism in mice. *Metabolism* 63(6):841–850. <https://doi.org/10.1016/j.metabol.2014.03.006>
- Xing G, Ren M, Verma A (2018) Divergent induction of branched-chain aminotransferases and phosphorylation of branched chain keto-acid dehydrogenase is a potential mechanism coupling branched-chain keto-acid-mediated-astrocyte activation to branched-chain amino acid depletion-mediated cognitive deficit after traumatic brain injury. *J Neurotrauma* 35(20):2482–2494. <https://doi.org/10.1089/neu.2017.5496>
- Xu WY, Gu MM, Sun LH, Guo WT, Zhu HB, Ma JF, Yuan WT, Kuang Y, Ji BJ, Wu XL, Chen Y, Zhang HX, Sun FT, Huang W, Huang L, Chen SD, Wang ZG (2012) A nonsense mutation in DHTKD1 causes Charcot-Marie-Tooth disease type 2 in a large Chinese pedigree. *Am J Hum Genet* 91(6):1088–1094. <https://doi.org/10.1016/j.ajhg.2012.09.018>
- Xu W, Zhu H, Gu M, Luo Q, Ding J, Yao Y, Chen F, Wang Z (2013) DHTKD1 is essential for mitochondrial biogenesis and function maintenance. *FEBS Lett* 587(21):3587–3592. <https://doi.org/10.1016/j.febslet.2013.08.047>
- Xu WY, Zhu H, Shen Y, Wan YH, Tu XD, Wu WT, Tang L, Zhang HX, Lu SY, Jin XL, Fei J, Wang ZG (2018) DHTKD1 deficiency causes charcot-marie-tooth disease in mice. *Mol Cell Biol* 38(13). <https://doi.org/10.1128/MCB.00085-18>
- Xu J, Jakher Y, Ahrens-Nicklas RC (2020) Brain branched-chain amino acids in maple syrup urine disease: implications for neurological disorders. *Int J Mol Sci* 21(20). <https://doi.org/10.3390/ijms21207490>
- Xu E, Ji B, Jin K, Chen Y (2023) Branched-chain amino acids catabolism and cancer progression: focus on therapeutic interventions. *Front Oncol* 13:1220638. <https://doi.org/10.3389/fonc.2023.1220638>
- Yamagiwa S, Kamimura H, Takamura M, Aoyagi Y (2014) Autoantibodies in primary biliary cirrhosis: recent progress in research on the pathogenetic and clinical significance. *World J Gastroenterol* 20(10):2606–2612. <https://doi.org/10.3748/wjg.v20.i10.2606>
- Yang HC, Hainfeld JF, Wall JS, Frey PA (1985) Quaternary structure of pyruvate dehydrogenase complex from *Escherichia coli*. *J Biol Chem* 260(30):16049–16051
- Yang H, Frey PA, Hainfeld JF, Wall JS (1986) Pyruvate dehydrogenase complex of *Escherichia coli*: radial mass analysis of subcomplexes by scanning transmission electron microscopy. *Biophys J* 49(1):56–58. [https://doi.org/10.1016/S0006-3495\(86\)83591-3](https://doi.org/10.1016/S0006-3495(86)83591-3)

- Yang L, Shi Q, Ho DJ, Starkov AA, Wille EJ, Xu H, Chen HL, Zhang S, Stack CM, Calingasan NY, Gibson GE, Beal MF (2009) Mice deficient in dihydrolipoyl succinyl transferase show increased vulnerability to mitochondrial toxins. *Neurobiol Dis* 36(2):320–330. <https://doi.org/10.1016/j.nbd.2009.07.023>
- Yang N, Han L, Gu X, Ye J, Qiu W, Zhang H, Gong Z, Zhang Y (2012) Analysis of gene mutations in Chinese patients with maple syrup urine disease. *Mol Genet Metab* 106(4):412–418. <https://doi.org/10.1016/j.ymgme.2012.05.023>
- Yang L, Wagner T, Mechaly A, Boyko A, Bruch EM, Megrian D, Gubellini F, Alzari PM, Bellinzoni M (2023) High resolution cryo-EM and crystallographic snapshots of the actinobacterial two-in-one 2-oxoglutarate dehydrogenase. *Nat Commun* 14(1):4851. <https://doi.org/10.1038/s41467-023-40253-6>
- Yao L, Xu X, Xu Y, Li C, Xie F, Guo M, Liu Z, Liu X (2022) OGDHL ameliorates cognitive impairment and Alzheimer's disease-like pathology via activating Wnt/beta-catenin signaling in Alzheimer's disease mice. *Behav Brain Res* 418:113673. <https://doi.org/10.1016/j.bbr.2021.113673>
- Yap ZY, Strucinska K, Matsuzaki S, Lee S, Si Y, Humphries K, Tarnopolsky MA, Yoon WH (2021) A biallelic pathogenic variant in the OGDH gene results in a neurological disorder with features of a mitochondrial disease. *J Inher Metab Dis* 44(2):388–400. <https://doi.org/10.1002/jimd.12248>
- Yeaman SJ (1989) The 2-oxo acid dehydrogenase complexes: recent advances. *Biochem J* 257(3):625–632. <https://doi.org/10.1042/bj2570625>
- Yoshida T, Kido J, Mitsubuchi H, Matsumoto S, Endo F, Nakamura K (2017) Clinical manifestations in two patients with pyruvate dehydrogenase deficiency and long-term survival. *Hum Genome Var* 4:17020. <https://doi.org/10.1038/hgv.2017.20>
- Youn H, Kang SO (2000) Enhanced sensitivity of *Streptomyces seoulensis* to menadione by superfluous lipoamide dehydrogenase. *FEBS Lett* 472(1):57–61. [https://doi.org/10.1016/S0014-5793\(00\)01437-X](https://doi.org/10.1016/S0014-5793(00)01437-X)
- Yu XK, Hiromasa Y, Tsen H, Stoops JK, Roche TE, Zhou ZH (2008) Structures of the human pyruvate dehydrogenase complex cores: A highly conserved catalytic center with flexible N-terminal domains. *Structure* 16(1):104–114. <https://doi.org/10.1016/j.str.2007.10.024>
- Yu Y, Hao H, Kong L, Zhang J, Bai F, Guo F, Wei P, Chen R, Hu W (2023) A metabolomics-based analysis of the metabolic pathways associated with the regulation of branched-chain amino acids in rats fed a high-fructose diet. *Endocr. Connect* 12(10). <https://doi.org/10.1530/EC-23-0079>
- Yuan L, Kim H (2010) Characterization of a naturally occurring mutation (Arg-460 to Gly) close to FAD in human dihydrolipoamide dehydrogenase. *Bull Korean Chem Soc* 31(12):3511–3512. <https://doi.org/10.5012/bkcs.2010.31.12.3511>
- Yuan L, Kim H (2012) Characterization of a naturally occurring mutation (Ile-358 to Thr) in human dihydrolipoamide dehydrogenase of a patient with leigh syndrome. *Bull Korean Chem Soc* 33(5):1445–1446. <https://doi.org/10.5012/bkcs.2012.33.5.1445>
- Yuan L, Cho YJ, Kim H (2008) Characterization of two naturally occurring mutations close to cofactors in human dihydrolipoamide dehydrogenase. *Bull Korean Chem Soc* 29(12):2327–2328
- Yuan L, Cho YJ, Kim H (2009) Characterization of a naturally occurring mutation (Ile-12 to Thr) close to prosthetic group FAD in human dihydrolipoamide dehydrogenase. *Bull Korean Chem Soc* 30(4):777–778
- Zahonova K, Valach M, Tripathi P, Benz C, Oppendoes FR, Barath P, Lukacova V, Danchenko M, Faktorova D, Horvath A, Burger G, Lukes J, Skodova-Sverakova I (2023) Subunit composition of mitochondrial dehydrogenase complexes in diplomonid flagellates. *Biochim Biophys Acta Gen Subj* 1867(9):130419. <https://doi.org/10.1016/j.bbagen.2023.130419>
- Zervopoulos SD, Boukouris AE, Saleme B, Haromy A, Tejay S, Sutendra G, Michelakis ED (2022) MFN2-driven mitochondria-to-nucleus tethering allows a non-canonical nuclear entry pathway of the mitochondrial pyruvate dehydrogenase complex. *Mol Cell* 82(5):1066–1077. <https://doi.org/10.1016/j.molcel.2022.02.003>

- Zhang L, Han J (2017) Branched-chain amino acid transaminase 1 (BCAT1) promotes the growth of breast cancer cells through improving mTOR-mediated mitochondrial biogenesis and function. *Biochem Biophys Res Commun* 486(2):224–231. <https://doi.org/10.1016/j.bbrc.2017.02.101>
- Zhang J, Wang YT, Miller JH, Day MM, Munger JC, Brookes PS (2018) Accumulation of succinate in cardiac ischemia primarily occurs via canonical Krebs cycle activity. *Cell Rep* 23(9):2617–2628. <https://doi.org/10.1016/j.celrep.2018.04.104>
- Zhang X, Nemeria NS, Leandro J, Houten S, Lazarus M, Gerfen G, Ozohanics O, Ambrus A, Nagy B, Brukh R, Jordan F (2020) Structure-function analyses of the G729R 2-oxoadipate dehydrogenase genetic variant associated with a disorder of l-lysine metabolism. *J Biol Chem* 295(23):8078–8095. <https://doi.org/10.1074/jbc.RA120.012761>
- Zhao Y, Hawes J, Popov KM, Jaskiewicz J, Shimomura Y, Crabb DW, Harris RA (1994) Site-directed mutagenesis of phosphorylation sites of the branched chain alpha-ketoacid dehydrogenase complex. *J Biol Chem* 269(28):18583–18587
- Zhong Y, Gao Y, Zhou D, Ma X, Chen H, Xu Y, Yang W, Yu X (2022) Structural basis for the activity and regulation of human α -ketoglutarate dehydrogenase revealed by Cryo-EM. *Biochem Biophys Res Commun* 602:120–126. <https://doi.org/10.1016/j.bbrc.2022.02.093>
- Zhou ZH, McCarthy DB, O'Connor CM, Reed LJ, Stoops JK (2001) The remarkable structural and functional organization of the eukaryotic pyruvate dehydrogenase complexes. *Proc Natl Acad Sci U S A* 98(26):14802–14807. <https://doi.org/10.1073/pnas.011597698>
- Zhou M, Lu G, Gao C, Wang Y, Sun H (2012) Tissue-specific and nutrient regulation of the branched-chain alpha-keto acid dehydrogenase phosphatase, protein phosphatase 2Cm (PP2Cm). *J Biol Chem* 287(28):23397–23406. <https://doi.org/10.1074/jbc.M112.351031>
- Zhou J, Yang L, Ozohanics O, Zhang X, Wang J, Ambrus A, Arjunan P, Brukh R, Nemeria NS, Furey W (2018) A multipronged approach unravels unprecedented protein-protein interactions in the human 2-oxoglutarate dehydrogenase multienzyme complex. *J Biol Chem* 293(50):19213–19227. <https://doi.org/10.1074/jbc.RA118.005432>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 14

Structures and Electron Transport Paths in the Four Families of Flavin-Based Electron Bifurcation Enzymes



Xiang Feng , Gerrit J. Schut , Michael W. W. Adams , and Huilin Li

Abstract Oxidoreductases facilitating electron transfer between molecules are pivotal in metabolic pathways. Flavin-based electron bifurcation (FBEB), a recently discovered energy coupling mechanism in oxidoreductases, enables the reversible division of electron pairs into two acceptors, bridging exergonic and otherwise unfeasible endergonic reactions. This chapter explores the four distinct FBEB complex families and highlights a decade of structural insights into FBEB complexes. In this chapter, we discuss the architecture, electron transfer routes, and conformational changes across all FBEB families, revealing the structural foundation that facilitate these remarkable functions.

Keywords Electron bifurcation · Flavin · Iron-sulfur cluster · Electron transfer · Catalytic mechanism · cryoEM · Structural biology

X. Feng

Division of Chemistry and Chemical Engineering, Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA, USA

e-mail: xfeng2@caltech.edu

G. J. Schut · M. W. W. Adams

Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, USA

e-mail: gerti@uga.edu; adamsm@uga.edu

H. Li (✉)

Department of Structural Biology, Van Andel Institute, Grand Rapids, MI, USA

e-mail: Huilin.Li@vai.org

Introduction

Electron bifurcation is a biochemical process that involves the reversible division of an electron pair into two electrons that results in one electron with a more positive reduction potential and the second with a more negative reduction potential than the original electron pair (see Fig. 14.1a). Flavin-based electron bifurcation (FBEB) has emerged as a recent discovery, alongside the previously established quinone-based electron bifurcation (QBEB). These mechanisms contribute to the expanding array of examples within the electron bifurcation systems (Buckel and Thauer 2018a).

The QBEB was the only kind of electron bifurcation known for a long time (Li et al. 2008; Buckel and Thauer 2018b). The bifurcation reaction occurs in the electron transport chain complex III, which is the cytochrome bc1 complex embedded in the plasma membrane of bacteria or the mitochondrial inner membrane of eukaryotes. The bifurcation mechanism known as the Q-cycle was proposed by Nobel laureate Peter Mitchel in the 1970s (Mitchell 1976). Other early contributions to the experiments and theory of the QBEB mechanism include Baum (Baum et al. 1967) and Wikström (Wikström and Berden 1972). In QBEB, a quinone (Q) accepts two electrons to become quinol (QH₂), which then bifurcates the two electrons, directing

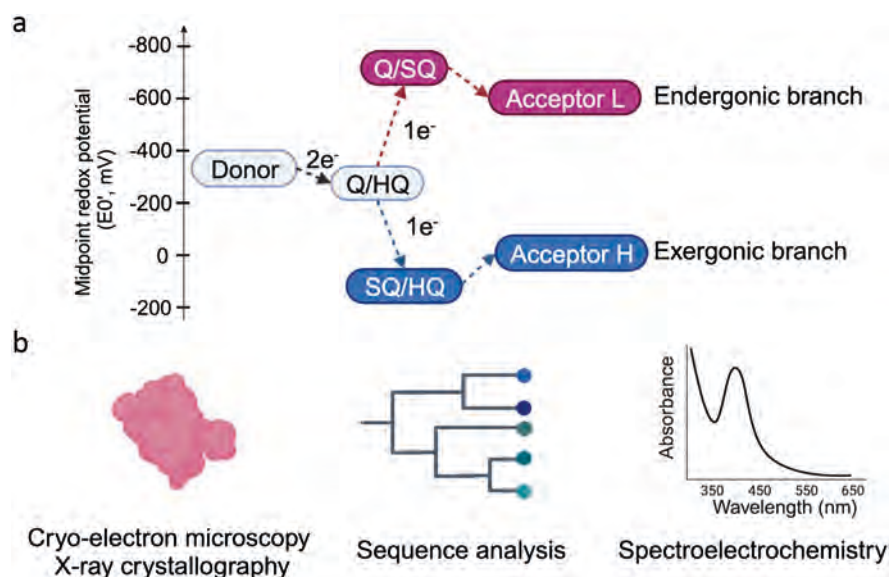


Fig. 14.1 Biophysical and bioinformatic approaches to studying the flavin-based electron bifurcation (FBEB) mechanism. (a) An electron pair is split into a high- and a low-potential level by FBEB. The final acceptor of high-energy, low-potential branch is usually a ferredoxin. This reaction is usually endergonic and is coupled with the other reaction in the low-energy, high-potential branch. The center of the reaction is a flavin molecule, either an FMN or an FAD. Q (Quinone), HQ (hydroquinone), and SQ (semiquinone) stand for three states of the flavin. (b) Three key methods in understanding the FBEB mechanism: structural biology, sequence analysis, and spectroelectrochemistry

one electron toward the “high-potential path,” passing through the iron-sulfur cluster containing Rieske protein to reach the cytochrome c1 and cytochrome c, and concomitantly the other electron of the semiquinone (SQ) state toward the “low-potential path.” This consists of two hemes that reduce quinone or semiquinone that is located at the second Q binding site in complex III. During the process, four protons are pumped into the intermembrane space to generate the proton gradient. Several reviews have nicely summarized the biochemistry and structural features of this system (Sarewicz et al. 2021; Crofts 2004; Crofts et al. 2013; Xia et al. 2013).

Around 15 years ago, a second kind of electron bifurcation reaction was discovered by Wolfgang Buckel and Rolf Thauer, which is based on a flavin-containing central molecule: either a flavin adenine dinucleotide (FAD) or a flavin mononucleotide (FMN) (Buckel and Thauer 2018b). The protein complexes catalyzing such reactions were subsequently found in numerous anaerobes, spanning a wide range of metabolic processes. Moreover, they are not integral transmembrane proteins like complex III and are instead water soluble or in a few cases loosely associated with the cell membrane. Remarkably, the FBEB appears to have evolved at least four times independently, leading to the discoveries of four phylogenetically unrelated families of the FBEB enzyme complexes so far (Buckel and Thauer 2018b). Three of the four families are named after their conserved electron bifurcating modules: the electron transfer flavoprotein EtfAB-containing family (EtfAB), the bifurcating ABC-containing family (BfuABC, previously termed HydABC), and the heterodisulfide reductase (Hdr) family (HdrABC). The fourth family is the NADH-dependent ferredoxin NADP+ oxidoreductase (NfnSL), named after the function rather than its relatively simple bifurcating structural module. It is noted that the Nfn-like Stn enzyme found in *Sporomusa ovata* catalyzes the same reaction as NfnSL (Kremp et al. 2020), but it is a member of the BfuABC family (Nfn-BfuABC) (Schut et al. 2022).

As of now, representative structures have been reported for all four families (Table 14.1). The rapid expansion of the available FBEB complex structures over the past decade can be largely attributed to the advent of high-resolution cryogenic electron microscopy (cryoEM). There are several reasons why cryoEM is an ideal method for obtaining structural insights into FBEB complexes. First, FBEB complexes are multi-subunit complexes, and usually they further form even larger oligomers with high molecular weights, making them excellent candidates for cryoEM, which excels in visualizing protein complexes exceeding 100 kDa. Second, cryoEM requires a relatively small amount of sample compared to X-ray crystallography, which is advantageous for studying novel proteins present in relatively low concentrations in poorly studied anaerobic bacteria for which genetic tools (for overexpression) are not available. Third, cryoEM obviates the need for crystal packing and can elucidate protein complex structures in various conformations as well as providing information on oligomeric mixtures that typically prohibit crystallography. This is particularly beneficial when dealing with proteins that undergo substantial conformational changes during reaction processes. In three out of the four families, significant domain movements have been observed during their functional processes, as discussed below, reminiscent of the behavior of the Rieske protein in the QBEB

Table 14.1 A list of available structures of the four families of FBEB enzymes

Class	PDB code	Enzyme	States	Donor	Low-pot. acceptor	High-pot. acceptor
EtfAB	4L2I	EtfAB-Bcd	EtfAB	2 NADH	Fd/Fld	Crotonyl-CoA
	5OL2	EtfAB-Bcd	D-state	2 NADH	Fd/Fld	Crotonyl-CoA
	6FAH	CarCDE	D-state	2 NADH	Fd	Caffeyl-CoA
	7KOE	Fix/ EtfABCX	D-state	2 NADH	Fd	Menaquinone
	7QH2	Ldh-EtfAB	D-state	2 NADH	Fd	Pyruvate
NfnSL	4YRY	NfnI		2 NADPH	Fd	NAD ⁺
	5JFC	NfnI	Without NADP bound	2 NADPH	Fd	NAD ⁺
	5JCA	NfnI	With NADP bound			
	5VJ7	NfnII		2 NADPH	Fd	?
HdrABC	5ODC	HdrABC-MvhAGD	With FAD	2 H ₂	Fd	CoM-S-S-coB
	5ODI		FAD and HS-CoM			
	5ODQ		FAD and bromoethanesulfonate			
	5ODR		FAD and 2 min with CoM-S-S-CoB			
	5ODH		FAD and 3.5 min with CoM-S-S-CoB			
	7BKB	Fdh-Hdr-Fmd	Hexamer	F ₄₂₀ /Formate	CO ₂	CoM-S-S-coB
	7BKC		Dimer			
	7BKD		Mobile arm state 1			
	7BKE		Mobile arm state 2			
BfuABC	7T2R	[NiFe]-BfuABCSL	Without FMN/NAD	2 H ₂	Fd	NAD ⁺
	7T30		With FMN/NAD			
	8A6T	[FeFe]-BfuABC (<i>T. kiuvi</i>)	Reduced state	2 H ₂	Fd	NADP ⁺
	8BEW		Oxidized state			
	8A5E	[FeFe]-BfuABC (<i>A. woodii</i>)	Reduced state	2 H ₂	Fd	NAD ⁺
	7Q4W		Oxidized state, symmetric			
	7Q4V		Oxidized state, unsymmetric			
	7P5H	[FeFe]-BfuABC (<i>T. maritima</i>)	D2-tetramer	2 H ₂	Fd	NAD ⁺
	7P8N		“Bridge” closed forward			
	7P91		“Bridge” closed reverse			
	7P92		“Bridge” open			
	8OH5	Nfn-BfuABC (StnABC)	State2	2 NADPH	Fd	NAD ⁺
	8OH9		State1			

system. For instance, the recent cryoEM study of a HdrABC-containing complex successfully resolved two distinct states, one resembled the previous crystal structure, while the other is a novel state (Watanabe et al. 2021).

Structural studies have revealed that all FBEB enzymes contain three components: (1) a flavin-containing region (three families use FAD, and one, the Bfu family, uses FMN) that bifurcates electrons, (2) the electron donor binding site, and (3) the electron acceptor binding sites that are at different levels of redox potential. The low-potential, high-energy acceptor is typically a ferredoxin, although flavodoxin or [FeS] clusters connecting to another active site (Chowdhury et al. 2016; Watanabe et al. 2021) are also used. The high-potential, low-energy acceptor could be a short-chain CoA derivative, menaquinone, NAD⁺, and heterodisulfide (see Table 14.1). By simply connecting the electron transporting cofactors identified in the enzyme structures within the range of direct electron transfer, plausible electron transfer pathways emerge. However, the actual electron paths utilized by the electron bifurcation reaction are far from clear. Nevertheless, the structural models and the derived potential electron transfer paths are indispensable, as they provide a physical basis as well as a testable hypothesis for elucidating the underlying bifurcation mechanisms. These include identifying residues responsible for coordinating cofactors, fine-tuning the cofactor redox potentials, mediating the enzyme conformational changes, and those involved in electron transport and proton transfer pathways. Substituting these residues allows for an examination of their impact on the FBEB activity.

The vast dataset of microbial genomic information has been critical to our expanding knowledge on FBEB and allows for rapid identification of microbial gene clusters that may form the FBEB complexes (Garcia Costas et al. 2017; Liang et al. 2019; Schut et al. 2022; Appel et al. 2021). Comparing the sequences of the bifurcating and non-bifurcating homologs can reveal potentially critical sequence motifs responsible for the FBEB functions. Despite the importance of the structural and bioinformatic knowledge, the electron bifurcation mechanism of an FBEB enzyme cannot be understood without experimental measurement of the redox potentials and redox states of the cofactors mediating the process. Spectroscopic and electrochemical tools such as infrared (FTIR) and electron paramagnetic resonance (EPR) can detect signature absorbance and resonance signals of the electron transporting cofactors, such as iron sulfur clusters and flavin semiquinones. However, interpreting such data can be an enormous challenge due to the presence of numerous cofactors in a typical FBEB complex leading to overlapping signals. The combination of the genomics and bioinformatics, structure characterization, and spectroscopic and electrochemical measurements will continue to deepen our understanding of the FBEB mechanism (Fig. 14.1b).

The evolution of the FBEB system has likely been driven by the need for reduced ferredoxin to power many energy-demanding microbial pathways in predominantly anaerobic metabolism (Müller et al. 2018). Low-potential high-energy electrons can only be produced from high-potential electrons by coupling an endergonic reaction with an exergonic reaction. However, there are cases in which both acceptors are high-potential acceptors at different redox levels (Schut et al. 2021), the

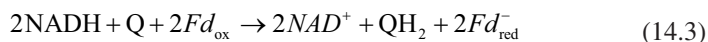
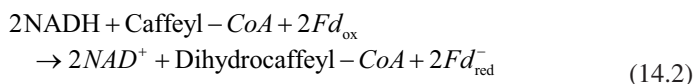
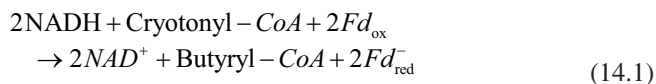
physiological significance of such a bifurcation reaction is unclear. By analogy to semiconductor physics, the enzymatic FBEB process may resemble a “transistor” that controls and amplifies the electric potential. The fundamental mechanism of the FBEB or even QBEB is not fully understood. In addition, at least one of the four FBEB families use a different mechanism for electron bifurcation, as discussed below in this chapter.

The FBEB systems possess two important features. One feature is the reversibility of the electron bifurcation reactions where, by coupling exergonic and endergonic reactions, the net free energy change is close to zero, particularly under physiological conditions (Baymann et al. 2018). The reversal of electron bifurcation is termed flavin-based electron confurcation (FBEC) (Yan and Ferry 2018; Müller et al. 2018; Kayastha et al. 2022), although FBEB is typically used with no implied physiological direction. In many cases, the same reactions can occur in both directions *in vivo* just like the Q-cycle in the QBEB, depending on the redox state of the organism (Osyczka et al. 2004), but in most cases, only one direction occurs physiologically (Kayastha et al. 2022). The second feature is the presence of a “gating” mechanism (Crofts et al. 2006). The gating mechanism prevents the two single electrons resulting from FBEB from entering the same high-potential branch, thereby creating a “short-circuiting” scenario in which the redox energy is wasted without producing reduced ferredoxin. The gating mechanism likely relies on protein conformational changes in the electron bifurcating process. Indeed, significant domain movement has been observed in three of the four families of the FBEB complexes, as well as in the Rieske protein in the QBEB complex III. It was recently proposed that there is a steep free-energy (reduction potential) landscape between the high- and low-potential branches, and such landscape suppresses short-circuiting and ensures efficient electron bifurcation (Yuly et al. 2020). This proposal relies on spatial separation of the high- and low-potential branches. We suggest that the conformational dynamics of the FBEB complexes is key to maintain such spatial separation in the changing energy landscape during the electron bifurcation process. Several excellent reviews have summarized the development, chemistry, and genetics of the field (Poudel et al. 2018; Baymann et al. 2018; Buckel and Thauer 2018a; Peters et al. 2018; Müller et al. 2018). This chapter is focused on the structures and conformational changes of select FBEB/FBEC enzyme complexes from each of the four families.

The ETF-Containing Family

The ETF-containing family of the FBEB complexes is the largest and most diverse (Buckel and Thauer 2018a). This family is involved in multiple metabolic pathways including amino acid degradation, lactate utilization, and short-chain fatty acid formation (Li et al. 2008; Chowdhury et al. 2014). The enzyme complexes use electrons from NADH to generate reduced ferredoxin in the high-energy, low-potential pathway, coupled with an exogenic high-potential pathway of reducing either

menaquinone (Schut et al. 2019), caffeyl-CoA (Demmer et al. 2018), crotonyl-CoA (Chowdhury et al. 2014), or other acceptor substrates (four examples of different high-potential acceptor are shown in Eqs. 14.1–14.4). A key feature of the family is that all members contain a homolog of the electron transfer flavoprotein (ETF), which is a conserved protein that shuttle electrons from various acyl-CoA dehydrogenases to the quinone pool in mitochondria (Henriques et al. 2021).



Architecture of the Electron Bifurcating ETF Family

ETF enzymes are a widely distributed family of proteins containing at least two flavin cofactors, typically FAD. In the canonical non-bifurcating ETFs, the cofactors are FAD and adenosine monophosphate (AMP). In the FBEB complexes, the AMP is replaced with an FAD, and this is the flavin that mediates electron bifurcation. Structures have been determined for four enzymes of the ETF-containing family that carry out the FBEB reactions described in Eqs. (14.1–14.4) (Demmer et al. 2017; Feng et al. 2021; Kayastha et al. 2022; Demmer et al. 2018) as well as one structure for the ETF module of the complex (Chowdhury et al. 2014). The structures of the holoenzymes can be divided into two parts: an EtfAB bifurcating module that accepts mid-potential electrons from NADH and reduces ferredoxin and an oxidoreductase module (e.g., EtfCX in the EtfABCX complex) that transfers high-potential electrons to the high-potential acceptor. The diversity of the oxidoreductase modules is reminiscent of the eukaryotic ETF proteins that interact with multiple proteins to transfer electrons. For example, in humans, there are more than 14 different protein modules that are coupled to the ETF proteins (Henriques et al. 2021).

The EtfABs have a dimension of 60 Å by 50 Å (Fig. 14.2a). The gene order of the two subunits is not always consistent in different genomes. For example, the gene order of the two chains is swapped in the homologous FixAB complex (Schut et al. 2019; Edgren and Nordlund 2004). To make it more consistent, we name the homologs of the EtfA gene as Etf-α and the homologs of the EtfB gene as Etf-β. The Etf-α subunit contains two domains (domain I and II), while the β-subunit contains only one (domain III). The three domains are similar, all having the Rossmann-like α/β/α sandwich fold (Fig. 14.2a). The two Etf-α domains are connected by a loop. The main interface between the two subunits is between domain I and III (Fig. 14.2a).

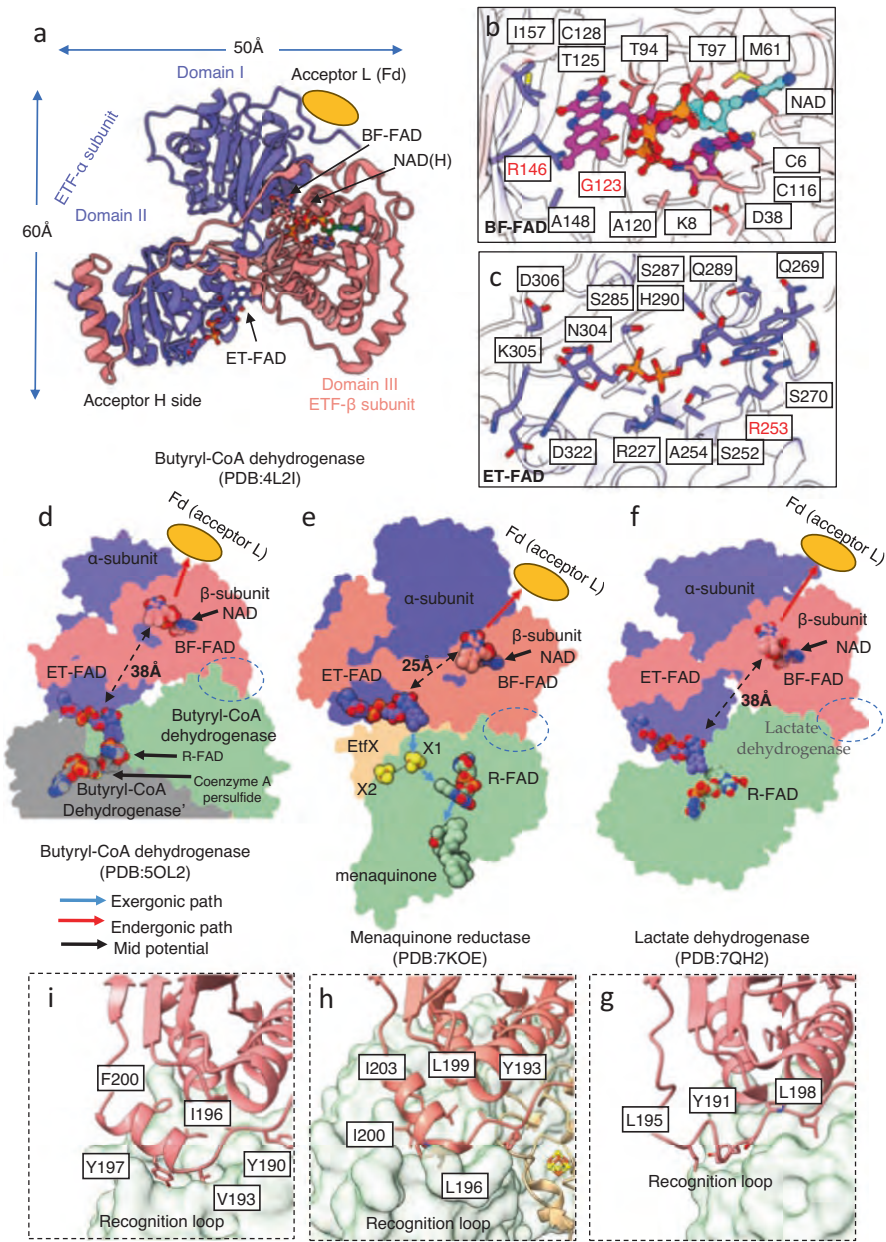


Fig. 14.2 The ETF-containing complexes. (a) The consensus bifurcating module is composed of ETF- α and β subunits with three domains (Domains I-III), using butyryl-CoA dehydrogenase (PDB: 4L2I), the only structure containing NAD, as an example. The module contains two FADs: one electron bifurcating (BF) and the other electron transferring (ET). The high-potential electron transfer direction (acceptor H) and low-potential electron transfer direction (acceptor L) are shown. (b, c) The local environment of the two FADs. Arg-146 highlighted in red is proposed to play an

(continued)

important role in bifurcation; the ET-FAD lacks such a chelating residue. **(d–f)** Three enzymes with known structures in this class. The distance between BF-FAD and ET-FAD is too far for direct electron transfer in all three enzymes. Domain II is proposed to move toward domain III to enable the ETF-FAD to receive high-potential electron from the BF-FAD. All complexes contain a third FAD (reducing, R-FAD) to transfer electron to a final high-potential acceptor. The Etf chain A and B are α and β subunits, respectively, except in panel **(e)**, where the Etf chain A is subunit β and Etf chain B is subunit α . **(i, g)** The recognition loop of subunit B interacts hydrophobically with subunit A in all three structures

The C-terminal extension of domain III, known as the β subunit arm, forms a β -strand and an α -helix to tightly interact with domain II.

Cofactors and Electron Donor Coordination

The cofactors in the ETF-contain FBEB family members form a relatively simple network. Two FADs are found in EtfAB with one FAD in each chain. Both FADs bind at the canonical sites located at the top of the β -sheet of the Rossmann-fold. The FAD in domain II is the same as in the non-bifurcating ETF, while the FAD in domain III replaces the AMP in non-bifurcating ETFs. The FAD in domain III is referred to as the bifurcating FAD (BF-FAD), because this FAD bifurcates two electrons to two different pathways. The non-bifurcating FADs simply transfer electrons to a downstream functional module and are therefore called electron transfer FAD (ET-FAD). The FBEB complexes also contain one FAD in the oxidoreductase core that transfers electrons to a final high-potential electron acceptor and therefore are referred to as reducing FAD (R-FAD, see Fig. 14.2d–f).

The adenine of the BF-FAD is coordinated in a similar fashion to AMP in the non-bifurcating ETFs. What differentiates a bifurcating ETF from a non-bifurcating ETF is the presence of amino acid residues that coordinate the flavin ring system of the BF-FAD. Important residues include those with small side chains to avoid steric clash with the flavin moiety (Garcia Costas et al. 2017) and an arginine that interacts with the isoalloxazine ring to stabilize the transition state of the BF-FAD (Gonzalez-Viegas et al. 2023) (Fig. 14.2b). Positively charged lysine and arginine are also present at the equivalent location in the NfnLS-like and HdrABC-containing families. Mutation of these key residues abolishes the function of the FBEB complexes (Demmer et al. 2018). Unlike the BF-FAD that is fully buried inside the complexes, the ET-FAD flavin moiety is partially exposed, allowing it to interact with either the R-FAD or the BF-FAD to transfer electrons (Fig. 14.2c). The residues surrounding the ET-FAD are well conserved in butyryl-CoA dehydrogenase, menaquinone reductase, and lactate dehydrogenase (Fig. 14.2d–f), and indeed, substitution of the interacting arginine to glutamine (R243Q) in the butyryl-CoA dehydrogenase abolished the FBEB activity (Demmer et al. 2017), indicating that an intact ET-FAD coordination environment is critical to the FBEB function.

The Interfaces with the Effector Subunits Are Diverse

The effector oxidoreductase cores can assemble into oligomers, e.g., the lactate dehydrogenase, the butyryl-CoA dehydrogenase, and the EtfABCX all form a dimer (Feng et al. 2021; Kayastha et al. 2022; Demmer et al. 2017), and the CarB forms a tetramer (Demmer et al. 2018). The EtfAB module and the oxidoreductase core have two interfaces. The main interface is mediated by an elongated loop with a short α -helix. This structural feature is conserved in both bifurcating and non-bifurcating ETFs and is referred to as the “recognition loop” (Toogood et al. 2007). The loop interacts with a diversity of structures through a hydrophobic patch on the short α -helix as well as a conserved tyrosine (Fig. 14.2g–i). The second interface involves domain II and is mediated by the ET-FAD. Interestingly, the ET-FAD directly interacts with the R-FAD in the oxidoreductase core in butyryl CoA dehydrogenase and lactate dehydrogenase (Fig. 14.2d, f), but the ET-FAD and R-FAD have no direct interaction, and electron transfer between them is mediated by the EtfX, a ferredoxin-like protein with two iron sulfur clusters, in the menaquinone reductase EtfABCX (FixABCX) (Feng et al. 2021) (Fig. 14.2e). The EtfCX subunits are conserved but merged into a single polypeptide non-bifurcating ETF-QO in mammals (Zhang et al. 2006). The coordinating residues of the two iron sulfur clusters in EtfX are also present in these non-bifurcating systems, indicating a conserved high-potential electron transport path. The structure of the EtfAB-EtfQO complex is not yet known. Based on the sequence homology, the EtfAB-QO structure should be very similar to that of EtfABCX. Bioinformatic analysis has also indicated that the ETF-containing family likely has even more members and have diverse but as yet uncharacterized functions (Garcia Costas et al. 2017).

The Domain II Movement Is Linked to the FBEB Function

The EtfAB domains I and III are closely packed and largely stationary in a single conformation in all known complex structures, but domain II is highly mobile and exists in multiple conformations or states. The domain II movement has been proposed to be involved in the electron bifurcation process because the distance between ET-FAD and BF-FAD is too far away for electron transfer in all the observed conformations. For example, the distance between the two FADs is 18 Å in the structure of the EtfAB module determined in the absence of the function module: the butyryl-CoA dehydrogenase (PDB ID 4L2I) (Fig. 14.2a). The distance reaches nearly 40 Å in the intact butyryl-CoA dehydrogenase as well as in the lactate dehydrogenase structure (Fig. 14.2d, f). In the EtfABCX menaquinone reductase structure, the ET-FAD is 25 Å away from the BF-FAD (Fig. 14.2e), still too far away for direct electron transfer. The ET-FADs in these enzymes can transfer electrons to the downstream R-FAD (Fig. 14.2d–f). Thus, these states were termed the dehydrogenase-like states or D-states. These structures also indicate that domain II must move toward domain III to bring the two FADs closer than the 14 Å electron

transfer distance to enable bifurcation. However, the closely packed bifurcation or B-state (ETF-FAD and BF-FAD within 14 Å) has not been experimentally observed due to its short-lived nature.

The Ferredoxin Binding Site

No EtfAB-containing structures published so far have captured a bound ferredoxin, again likely due to the transient nature of the binding. The ferredoxin binding site is expected to be on domain I at the opposite side of where the ET-FAD binds, such that the low-potential electrons from the BF-FAD can reach their exclusive acceptor ferredoxin (Fig. 14.2d–f). The CarCDE complex contains an N-terminal ferredoxin-like domain that was thought to suggest how the redox protein might bind (Demmer et al. 2018). However, truncation of this domain did not affect the bifurcating functions (Demmer et al. 2018), indicating that this domain mediates the electron transfer to a downstream site where the ferredoxin binds. Therefore, the truncation merely shortened the low-potential electron transfer path but did not abolish electron transfer to ferredoxin.

The Nfn-Like Family

The transhydrogenase Nfn-like family is the only family with a simple and largely stationary FBEB architecture. The family is also widely distributed in the microbial world (Poudel et al. 2018). Three X-ray crystallographic structures have been reported in this family, and they are essentially the same with an Cα RMSD of 0.8 Å (Demmer et al. 2015; Lubner et al. 2017; Nguyen et al. 2017). As the name suggests, all enzymes in the family transfer electrons from NADPH to Fd_{ox} and NAD^+ (Eq. 14.5). Three types of Nfn complexes have been characterized (Nfn I, II, and III), two with known structures (Nfn I and II) (Nguyen et al. 2017; Liang et al. 2019). Nfn III has no reported structure, and the electron bifurcation function is predicted by sequence analysis (Nguyen et al. 2017).



A typical Nfn-like complex structure (Nfn I) is 90 Å by 65 Å and composed of two polypeptides (Fig. 14.3a, b). Because of the notable size difference, they are named as the large subunit (NfnL) and the small subunit (NfnS). NfnL consists of a helical domain that coordinates two iron-sulfur clusters (assigned L1 and L2 by their distance to the BF-FAD) and two Rossmann-fold domains that bind electron-bifurcating FAD (BF-FAD) and the electron donor NADPH. The NfnL structure is similar to that of NADPH-dependent glutamine synthase (Swuec et al. 2019) (PDB ID 6S6S, chain E). The BF-FAD binds in a similar manner as it does in glutamine

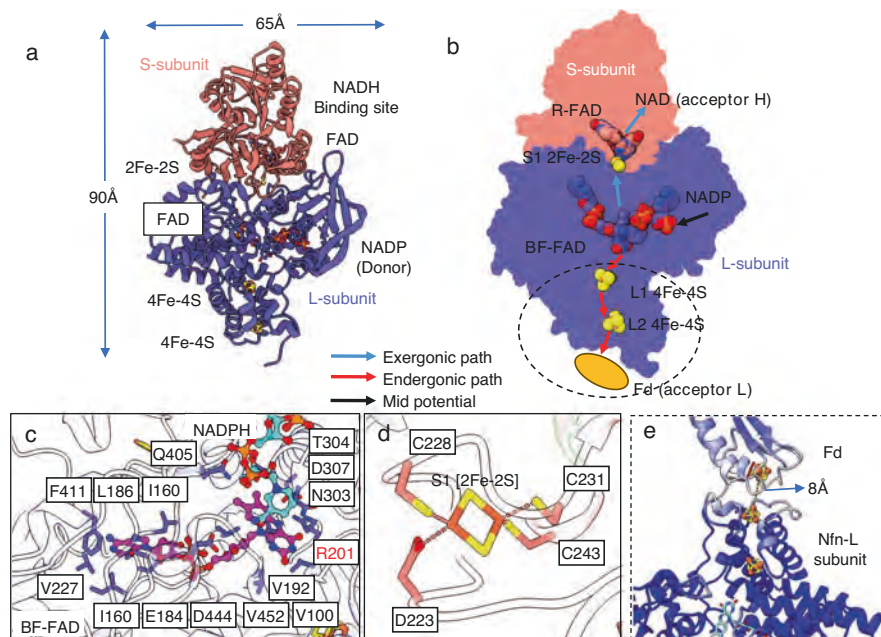


Fig. 14.3 The NfnSL complex. **(a, b)** The consensus bifurcating module contains a large (L) and a small (S) subunit, as typified by the transhydrogenase NfnI structure **(a)**, PDB ID 5JCA). The high-potential electron path leading to NAD (acceptor H) and the low-potential electron transfer path leading to ferredoxin (acceptor L) are shown in **(b)**. **(c)** The local environment of the electron bifurcating FAD. Arg-201 highlighted with red label is proposed to play an important role in bifurcation. **(d)** Cofactor S1 is coordinated by three Cys and one Asp. **(e)** An AlphaFold-Multimer predicted binding mode of the low-potential electron acceptor Fd on the bottom tip of the NfnL subunit. The structure is colored by the prediction confidence level with blue the highest confidence and red the lowest confidence

synthase. A notable distinction is the presence of a highly conserved arginine (Arg-201 in Nfn I) that forms a hydrogen bonding network with the 2'-O and N5 of the BF-FAD isoalloxazine ring (Fig. 14.3c). The structure of NfnS consists of an antiparallel six-stranded β -barrel, an Rossmann fold domain, a [2Fe-2S] cluster binding domain, and a C-terminal α -helix. The NfnS structure is similar to the dihydroorotate dehydrogenase B (Rowland et al. 2000) (PDB ID 1EP1, chain A) except for an extra α -helix at the heterodimer interface. The FAD in the NfnS Rossmann fold domain transfers electrons to the NAD functioning as the ET-FAD. The residues interacting with the S1 cluster [2Fe-2S] contain an aspartic acid residue (Asp-223), which is atypical for iron-sulfur cluster coordination but is conserved in the family (Fig. 14.3d). The unique Asp(Cys)₃ coordination was found to increase the redox potential of the S1 cluster compared to a typical (Cys)₄ ligated cluster (Hagen et al. 2000).

Spectroscopic and electrochemical studies have demonstrated that the BF-FAD is indeed the site of electron bifurcation (Fig. 14.3b) (Lubner et al. 2017). The

bifurcation reaction is started by the oxidation of NADPH at the BF-FAD, after which one electron is transferred to the [2Fe-2S] S1 cluster along the exergonic branch and then to NAD. The low-potential electron is transferred to the ferredoxin through the L1 and L2 [4Fe-4S] clusters. This family of enzyme complexes does not appear to undergo large conformational changes or domain movements. However, the distance between the [2Fe-2S] S1 cluster and the BF-FAD is 14–15 Å, which is at the upper limit for electron transfer as a distance at or below 14 Å is usually required (Page et al. 1999). A hydrogen-deuterium exchange mass spectrometry analysis showed that NfnAB rearranges upon NADPH binding to reduce the distance between the [2Fe-2S] and B-FAD to 13 Å (Demmer et al. 2016), which allows efficient electron transfer to the [2Fe-2S] S1 cluster and then to the NADH. The structural rearrangement may further allow the S1 cluster to move away from the BF-FAD to avoid short-circuiting. The structure of *Pyrococcus furiosus* NfnI was determined with ferredoxin bound as the enzyme was purified from native biomass (rather than heterologously produced) and the ferredoxin co-purified bound to the enzyme (Nguyen et al. 2017). Here we used the powerful AlphaFold (Jumper et al. 2021; Mirdita et al. 2022) to predict how ferredoxin binds to the Nfn I. We found the predicted NfnL structure stays very close to the experimental structure (<1 Å Cα RMSD) in which the Fd [4Fe-4S] cluster is only 8 Å away from the [4Fe-4S] L2 cluster in NfnL subunit, allowing for direct electron transfer (Fig. 14.3e).

The HdrABC-Containing Family

Structures of two HdrABC-containing family members have been reported so far (Wagner et al. 2017; Watanabe et al. 2021). Both enzyme complexes are from hydrogenotrophic and methanogenic archaea, which play critical roles in the global methane cycle (Thauer et al. 2008). The two complexes share the core of a (HdrABC-MvhD)₂ dimer of heterotetramers, which is the electron bifurcation module (Table 14.1, Fig. 14.4a). Additional subunits that interact with the electron donors and acceptors are attached to the periphery of the dimer core (Fig. 14.4b). In the HdrABC-MvhD-MvhAG complex, the peripheral MvhAG oxidizes hydrogen gas to exergonically reduce the high-potential heterodisulfide CoM-S-S-CoB to coenzyme M (CoM-SH) and coenzyme B (CoB-SH) and couples this to the endergonic reduction of ferredoxin (Kaster et al. 2011) (Eq. 14.6). In the HdrABC-MvD-FmdABCDFG-FdhAB (Fdh-Hdr-Fmd) complex, the FdhAB oxidizes either the midpotential methanogenic coenzyme 8-hydroxy-5-deazaflavin (F₄₂₀) by FdhB (Eq. 14.7) or formate by FdA (Eq. 14.8). The electrons are then transferred to the BF-FAD for exergonic reduction of the heterodisulfide, and in the Hdr-Fmd-Fdh, the low-potential electrons are not used to reduce ferredoxin but instead are linked to FmdABCDG through polyferredoxin FmdF to fix CO₂ to formylmethanofuran (formyl-MFR) (Watanabe et al. 2021). A recent study showed that the HdrABC core module can have diverse functions and is found in both *Archaea* and *Bacteria* (Appel et al. 2021).

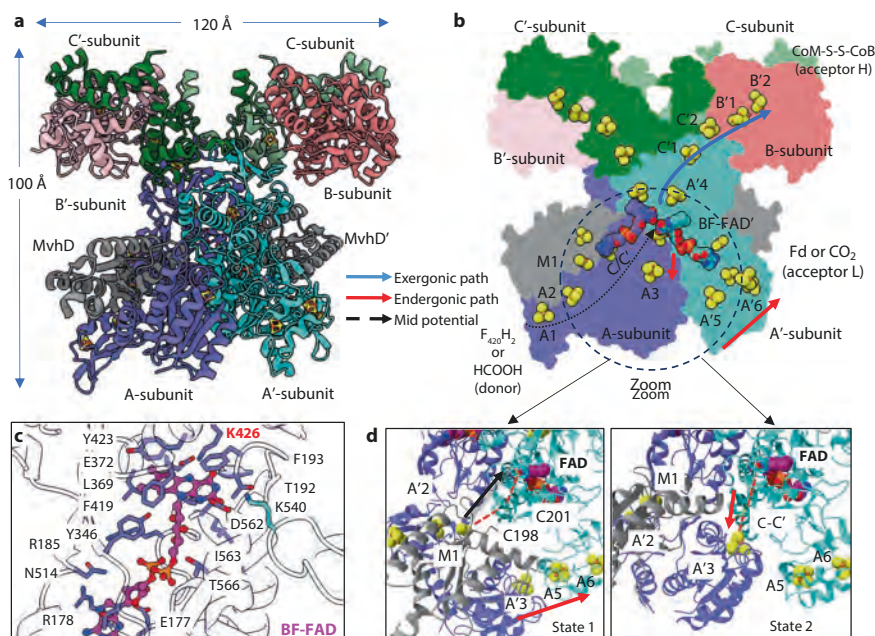
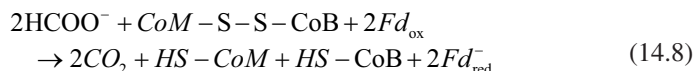
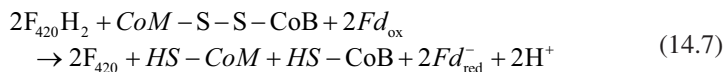
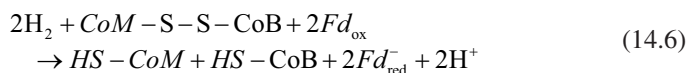


Fig. 14.4 The HdrABC-containing complex. (a, b) The consensus bifurcating module contains HdrABC and MvhD subunits (a, PDB ID 7bkc). Both protomers are thought to be required for electron bifurcation. The high-potential electron transfer direction to CoM-S-S-CoB (acceptor H) and the low-potential electron transfer direction to Fd or CO₂ (acceptor L) are shown in (b). The electron donor can be either F₄₂₀H₂ or HCOOH. (c) The local environment of the electron bifurcating FAD. Lys-426 highlighted in red label is proposed to play an important role in bifurcation. (d) The movement enabling the bifurcation process was observed in the cryoEM structures. In state 1, two mid-potential electrons are transferred to the FAD through a two-cysteine system, and the electron pair bifurcate at FMN. The A3' cofactor stores a low-potential electron from FMN in state 2 after a conformation change. Conversion from state 2 to state 1 allows the low-potential electron to reach the final acceptor through the A5 and A6 clusters

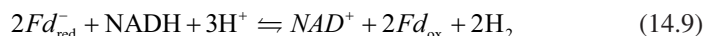


The HdrABC-MvhD dimer core is densely packed, with dimensions of 100 Å by 120 Å (Fig. 14.4a). The two HdrABC family complexes are among the largest known FBEB complexes. The Fdh-Hdr-Fmd complex assembles into a giant hexamer with a mass of 3 MDa (Watanabe et al. 2021) (Fig. 14.4a, b). The HdrA subunit contains four domains: an N-terminal domain (NTD) featuring an unusual five cysteine linked A3 [4Fe-4S] cluster, a thioredoxin-reductase (TrxR) domain that coordinates the BF-FAD and the A4 [4Fe-4S] cluster, an inserted ferredoxin domain coordinating the A5 and A6 [4Fe-4S] clusters, and a C-terminal ferredoxin domain coordinating the A1 and A2 [4Fe-4S] clusters. The MvhD subunit has a fold similar to the HdrA NTD but coordinates an M1 [2Fe-2S] cluster rather than a [4Fe-4S] cluster. The HdrB subunit adopts an aspartate racemase-like fold and coordinates the two closely related B1 and B2 [4Fe-4S] clusters with an unusual non-cubane structure. The HdrC subunit is located between HdrA and HdrB and has the ferredoxin-like fold that coordinates the C1 and C2 [4Fe-4S] clusters. As seen in the other two FBEB families, a positively charged lysine (K426) of the HdrA thioredoxin reductase domain ligates the N5 atom of the BF-FAD and may contribute to the stabilization of negatively charged intermediates, such as the fully reduced FADH⁻ (Fig. 14.4c). The importance of this lysine and the interacting glutamate (E372) is underscored by their conservation in all electron bifurcating HdrAs.

Unlike the oligomeric complexes in other FBEB families, in which individual protomers function independently, both protomers of the HdrA dimer are involved in the same electron transfer path (Fig. 14.4b). In HdrABC-MvhAGD complex, the MvhA and MvhG are homologous to the large and small subunits of [NiFe]-hydrogenases, respectively. The mid-potential electron transfer branch starts at the [NiFe] cofactor in the MvhA active site and proceeds through five [4Fe-4S] clusters in the MvhG and the A1 and A2 clusters in the HdrA C-terminal domain to reach the M1 [2Fe-2S] cluster in the MvhD adaptor. In the Fdh-Hdr-Fmd complex, the mid-potential electron transfer path branches before HdrA with a pterin-based formate oxidation site in FdhA and a F₄₂₀ binding site in FdhB. The BF-FAD in the HdrA' subunit then bifurcates two electrons, with one electron moving along the exergonic branch through the A4 cluster in the HdrA', the C1 and C2 clusters in the HdrC, and the two non-cubane B1 and B2 clusters in the HdrB to finally reduce the CoM-S-S-CoB in the HdrC. The assignment of HdrA, but not HdrA', belonging to the same protomer complex with HdrBC, is based on their interactions (Fig. 14.4b). The complex exists in two states (states 1 and 2) linked by a large conformational change (Fig. 14.4d). In state 2, the A3 cluster in the MvhA NTD is closer to the BF-FAD in the HdrA', enabling the bifurcation reaction. Presumably, the complex returns to state 1 post bifurcation to allow the electron in the low-potential branch to continue from the A3 cluster to the A5' and A6' clusters in the inserted ferredoxin domain of the HdrA'.

The BfuABC Family

The bifurcating (Bfu)-ABC family features a heterotrimer first identified in an unusually complex [FeFe]-hydrogenase (HydABC) (Schut and Adams 2009). The [FeFe]-hydrogenase couples the endergonic oxidation of NADH to the exergonic oxidation of a reduced ferredoxin to reduce protons, the mid-potential acceptor, to generate H₂ gas (Eq. 14.9). The physiological reaction is an electron confurcation process, the reverse of electron bifurcation. Similar confurcating/bifurcating hydrogenases have been identified in other species. Recently, a comprehensive bio-informatics study revealed that the HydABC family is ubiquitous in the microbial world, but only about half of the members are hydrogenases. The remainder are proposed to catalyze diverse metabolic reactions with a myriad of metabolic functions well beyond H₂ production (Schut et al. 2022). For this reason, this family was renamed as the bifurcating BfuABC family.



Since the structural reports of the first two family members in 2022 (Feng et al. 2022; Furlan et al. 2022), this family has rapidly become one with the largest pool of structural information (Katsyv et al. 2023; Kumar et al. 2023) (Table 14.1). It is also worth noting that all were cryo-EM structures, highlighting the importance of the method in understanding the nature and function of FBEB complexes. So far, the structures of three types of the BfuABC complexes have been determined: the [FeFe]-hydrogenase ([FeFe]-HydABC), the five-subunit [NiFe]-hydrogenase ([NiFe]-HydABCSL), and the most recent a transhydrogenase (StnABC), referred to here as Nfn-BfuABC. This Nfn-BfuABC carries out the same function as the NfnSL-family members in which electrons are transferred from NADPH to NAD⁺ and to Fd_{ox} (Eq. 14.5).

The Consensus Bifurcating Module Structure and Cofactors

The [FeFe]-HydABC hydrogenase is the first characterized member of the BfuABC family to be identified (Schut and Adams 2009), but the [NiFe]-HydABCSL, which includes two additional subunits ([NiFe]-SL) attached to the electron bifurcating HydABC hydrogenase core, was the first member of the family to be structurally characterized (Feng et al. 2022). Note that in the Nfn-BfuABC (StnABC) complex, the order of the genes encoding subunits A and C is swapped compared to other FeFe- and NiFe-BfuABC members. To avoid confusion and to make the names consistent, we refer in Fig. 14.5a to the A, B, and C-subunits as defined in the hydrogenases. Just like the EtfAB-containing and HdrABC-containing classes of FBEB complex described above, the Bfu-ABC complexes also assemble into larger oligomers. The [FeFe]-BfuABC complex of *Thermotoga maritima* and the Nfn-BfuABC complexes both form tetramers, while the [FeFe]-BfuABC from

Thermoanaerobacter kivui and *Acetobacterium woodii* and the [NiFe]-BfuABCSL from *Acetomicrobium mobile* all form dimers.

The bifurcating core module of the BfuABC family is around 80 Å by 70 Å. The BfuA subunit contains an N-terminal domain that coordinates one [2Fe-2S] and three [4Fe-4S] clusters that are arranged into a “Y-shape” (Zuchan et al. 2021). Specifically, the A1 [2Fe-2S] cluster coordinating subdomain and the A2 and A4 [4Fe-4S] clusters containing Fd-like domain are linked by the median domain that coordinates the A3 [4Fe-4S] cluster on the third branch. The A3 cluster is ligated in part with a histidine. The C-terminal regions of all three experimental structures contain a domain with two intertwined $\alpha/\beta/\alpha$ sandwich folds. In the [FeFe]-BfuABC, this domain coordinates the A5 [4Fe-4S] cluster and the [Fe-Fe] “H-cluster,” which is the active site of the hydrogenase (Fig. 14.5b). In the [NiFe]-BfuABC and Nfn-BfuABC, this domain only coordinates the A5 cluster (Fig. 14.5c, d). In the bifurcating [NiFe]-BfuABCSL enzyme, the [NiFe] cluster is in the hydrogenase module HydSL, and this is coupled to the other two electron transfer pathways via the A3 cluster (Fig. 14.5c). In the Nfn-BfuABC complex, the A3 cluster coordinating subdomain in BfuA is larger and resembles the NfnL of the NfnLS-containing family with a C α RMSD of 1.1 Å. Indeed, this domain contains all cofactors found in the NfnL subunit of the NfnLS family, including the L1 and L2 [4Fe-4S] clusters and the BF-FAD equivalents (Kumar et al. 2023).

The BfuB-subunit contains an N-terminal thioredoxin-like domain that coordinates the B5 [2Fe2S] cluster, a Rossmann-like domain, a soluble ligand binding β -grasp (SLBB) domain, a small four-helix bundle, and a C-terminal Fd-like domain. The electron-bifurcating FMN and the high-potential electron acceptor NAD are tightly packed together at the interface between the Rossmann-like domain and the SLBB domain. The four-helix bundle domain coordinates the B1 and B2 clusters; B1 is clearly a [4Fe-4S] cluster, but the identity of the B2 is somewhat ambiguous and has two assignments in the literature. The EM density fits well with a [2Fe-2S] cluster in the [NiFe]-HydABC (Feng et al. 2022) but fits only a single cation that was assigned to a zinc atom in the [FeFe]- and Nfn-BfuBs (Katsyv et al. 2023; Furlan et al. 2022; Kumar et al. 2023). The C-terminal ferredoxin domain coordinates the B3 and B4 [4Fe-4S] clusters and is connected to the four-helix bundle domain by a flexible loop. The flexible link leads to either the low resolution in the EM map (Feng et al. 2022) or the ferredoxin domain in several orientations (Furlan et al. 2022). The BfuC-subunit consists of an N-terminal domain and a C-terminal domain linked by a flexible loop. The N-terminal domain has a four-helix bundle fold and is located near the BfuA/BfuB interface. The C-terminal domain of BfuC has a thioredoxin fold that coordinates the C1 [2Fe–2S] cluster.

The combination of the “Y-shaped” arrangement of iron-sulfur clusters in the BfuA subunit with the NADH-oxidizing module in the BfuB/BfuC-subunits is prevalent in many non-bifurcating enzymes, such as complex I (Berrisford and Sazanov 2009), formate dehydrogenase (Radon et al. 2020), [NiFe]-hydrogenase (Shomura et al. 2017), and the ferredoxin dependent single subunit [FeFe]-hydrogenase (HydA) (Peters et al. 1998). The prevalence of such an arrangement is likely due to the ancient evolutionary root that dates to the last universal common ancestor

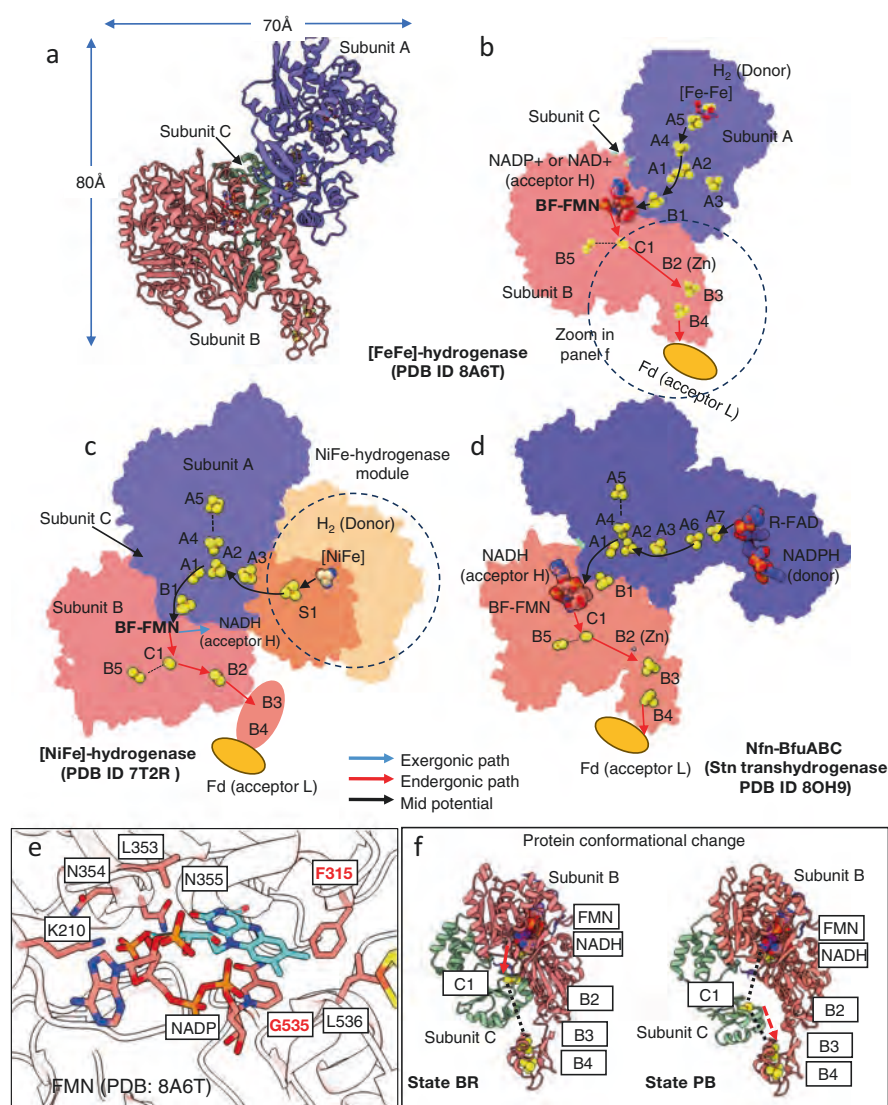


Fig. 14.5 The BfuABC-containing complexes. **(a, b)** The consensus bifurcating module contains A, B, and C subunits as typified by the [FeFe]-hydrogenase structure **(a)**, PDB ID 8A6T). The paths for middle-potential electron donor, high-potential electron transfer to NAD/NADP (acceptor H), and low-potential electron transfer to ferredoxin (acceptor L) are shown in **(b)**. **(b–d)** Three enzymes with known structures so far: [FeFe]-Hydrogenase HydABC, chain ABC represent subunits $\alpha\beta\gamma$, respectively **(b)**; [NiFe]-Hydrogenase ABCSL, chain ABC are subunit $\alpha/\beta/\gamma$, respectively; **(c)** and the *Sporomusa* transhydrogenase Nfn-BfuABC (StnABC) **(d)**. The subunits in the Nfn-BfuABC enzyme are not labeled as in the description of the StnABC enzyme (Kumar et al. 2023); its A and C subunits are homologous to BfuC and BfuA, respectively, as indicated in the color scheme of **(d)**. The mid-potential path could use either the A2 or A4 path to transfer electron to bifurcating FMN. **(e)** The local environment of the BF-FMN. No Lys or Arg is interacting with the isoalloxazine ring, in contrast to the other two classes. The distance from C1 to B3/B4 is too long for direct electron transfer in the reduced state. A conformational change by the C-subunit is proposed to bring C1 close to B3/B4, transitioning from the state before bifurcation (BR) to the state post bifurcation (PB). **(f)** The required conformational change was indeed observed in both [FeFe]- and [NiFe]-Bfu hydrogenases

(Zuchan et al. 2021). As an example of such evolution, the equivalent of the BfuA subunit of the non-bifurcating [NiFe]-hydrogenase lacks the A4 cluster as well as the C-terminal domain of the A-subunit (Shomura et al. 2017). The BfuA subunit of the bifurcating [NiFe]-hydrogenase has evolved to include the A4 and A5 clusters in the A-subunit, although it is not clear if these two clusters are part of the electron transfer pathways.

The Cofactor Network and the Electron Transfer Paths in the BfuABC Family

In the BfuABC family of FBEB enzymes, the “third” substrate, in addition to NAD and ferredoxin, feeds into the mid-potential path through BfuA to the B1 cluster in BfuB, e.g., the pathway starts from the H-cluster that oxidizes H₂ gas in the [FeFe]-hydrogenase. The electrons are then transferred along the path of A5-A4-A2-A1-B1 to reach the BF-FMN in BfuB (Fig. 14.5b). The A3 cluster is not utilized in this process. Interestingly, in the [FeFe]-hydrogenases, the A3 cluster is only 9 Å away from the A3' of the partner protomer in the oligomeric complexes, well within the range of direct electron transfer, raising the possibility that the clusters from other protomers are involved in electron transfer (Furlan et al. 2022; Katsyv et al. 2023). However, a mutation that abolishes the A3 coordination did not affect the FBEB function, indicating the A3 cluster is not utilized in the oligomeric complexes (Katsyv et al. 2023). The middle potential path in the [NiFe]-ABCSL hydrogenase is along the [NiFe]-S1-A3-A2-A1-B1 clusters (Fig. 14.5c). The Nfn-BfuABC complex also skips the A4 and A5 clusters and utilizes the equivalent route of the NADPH-R-FAD-A7-A6-A3-A2-A1-B1-FMN clusters to transfer electrons from NADPH to the bifurcating center focused at FMN (Fig. 14.5d). Indeed, the A5 cluster is too far from other clusters for electron transfer except for the A4 cluster in both [NiFe]- and Nfn-BfuABC complexes. Therefore, A5 likely does not participate in the electron transfer pathway. In Nfn-BfuABC, the ET-FAD is coordinated similarly to the BF-FAD in the NfnSL-family: a conserved Arg ligates the ET-FAD, but it is too far to H-bond with the N5 atom. Another difference is the A7 cluster ligation by a lysine rather than the glutamic acid coordinating the L1 cluster in the NfnSL-family (Fig. 14.3b). Different coordinating residues may alter the cofactor redox potential and explain why a similar cofactor arrangement architecture ends up in a very different electron transfer (potential) path. Indeed, mutation of the A7-coordinating lysine affected the function of Nfn-BfuABC (Kumar et al. 2023).

All BfuABC-type structures reported so far are consistent with the FMN functioning as part of the bifurcation center (Fig. 14.5b–d). However, the FMN interacts with the low-potential electron carrier, NAD/H, not the mid-potential carrier, H₂, whereas in the other three families of FBEB enzymes, the flavin (FAD) interacts directly or indirectly with the mid-potential carrier. Hence, the FMN in the Bfu enzymes does not carry out a “simple” electron bifurcation reaction in which two mid-potential electrons are split and passed to low- and high-potential pathways. It

was therefore proposed that the bifurcation site in the BfuABC family is a combination of FMN and at least two iron sulfur clusters, and bifurcation occurs by an as yet unknown mechanism (Schut et al. 2022; Feng et al. 2022). Consequently, the FMN coordination lacks an Arg/Lys that ligates the N5 atom of the BF-FAD in the other FBEB families, consistent with a different electron bifurcation mechanism (Fig. 14.5e). In the BfuABC family, the FMN and NAD are stabilized by the ⁴³⁰GGPSG⁴³⁴ motif and the F³¹⁵ (amino acid sequence as in the [FeFe]-Hydrogenase) present in all three member complexes, and these two motifs distinguish the bifurcating from the non-bifurcating B/C-subunits (Losey et al. 2020). The high-potential electron acceptor NAD⁺ (or NADP⁺) stacks against the FMN to facilitate direct electron transfer. The low-potential electron path is not entirely clear, as none of the available structures contain a bound ferredoxin (Fig. 14.5b–d). The FMN-C1-B5 clusters are within the electron transfer distance of each other, but the B5 and C1 are too far from the B3/B4 clusters. Nevertheless, mutations that abolish the B5 or B3/B4 binding greatly affected the function of the [FeFe]-BfuABC, supporting an involvement of the B5 and B3/B4 clusters in the low-potential path (Katsyv et al. 2023).

Apparently, the BfuC subunit's C-terminal domain (CTD) needs to move to bring the C1 cluster of BfuC closer to the B3/B4 clusters to complete the high-potential electron path and, at the same time, to move the C1 cluster away from the BF-FMN in BfuB to prevent backtracking or short-circuiting of the low-potential electrons. Indeed, such CTD movement was observed in the structure of the [NiFe]-BfuABCSL bound to both FMN and NAD⁺ (Feng et al. 2022). In one state, the CTD of BfuC packed closely with the BfuB subunit just as in the FMN/NAD-free state, allowing the electron to transfer from the FMN to the C1 cluster, and this state was referred to as a bifurcation ready (BR) state. In the other state, the CTD of BfuB shifted toward the B2 cluster, allowing low-potential electrons to transfer to the B3/B4 clusters via B2. This state was referred to as the post-bifurcation (PB) state. A similar movement was also observed in the oxidized [FeFe]-BfuABC (Fig. 14.5f). However, the B2 site in BfuB was reported to be a zinc atom rather than a [2Fe-2S] cluster in the [FeFe]-HydABC. Therefore, it was proposed, based on an MD simulation, that low-potential electrons are transferred from the C1 cluster directly to the B3/B4 clusters in the CTD of BfuB and then go on to reach the final low-potential acceptor, ferredoxin (Fig. 14.5d).

Current Challenges on Understanding the Bifurcation Mechanism in the BfuABC Family

Although the conformational changes revealed by cryoEM in the BfuABC family of enzyme complexes can satisfy the physical constraints on electron transfer in terms of distance, the exact mechanism of electron bifurcation in these complexes is still unclear. Because NAD⁺ is a two-electron acceptor on the high-potential electron

path, the hydrogenase needs to extract four electrons by oxidizing two H_2 molecules in the mid-potential electron path. The difficult question to address is how the four electrons are stored before the bifurcation reaction occurs and how the FMN coordinates the bifurcation of two pairs of electrons, if indeed that is what it does. What we do know is that the mechanism of the Bfu enzymes is different from that of the FAD (only)-based electron bifurcation in the three other FBEB families, which appear to share at least at one level a common mechanism. It is possible that in BfuABC complexes, some combination of the two iron-sulfur clusters, C1, B1 and B5, along with FMN, function together as a single unit to facilitate the bifurcation of four electrons (Feng et al. 2022; Schut et al. 2022). A recent paper based on experimental and simulation studies suggested that kinetic gating can be achieved by conformational changes on the BfuB/C-subunits (Katsyv et al. 2023). In that scenario, the C1 and B5 clusters store the first pair of electrons from H_2 oxidation and promote NAD(P)⁺ binding. As the second pair of electrons from H_2 are transferred to NAD(P)⁺ through FMN, the NAD(P)⁺ dissociates, and the CTD of C-subunit moves to enable electron transfer to the B3/B4 cluster and finally to ferredoxin. The caveat of this mechanism is that the “BF-FMN” does not really bifurcate the electrons to different potentials because the electron pair move in the same direction and the flavin does not reach the SQ state in the proposed mechanism. Clearly, the bifurcation mechanism of the enigmatic BfuABC family awaits further investigation.

Conclusion and Outlook

Structure-based studies have greatly advanced our understanding of many fascinating FBEB enzyme complexes. In this chapter, we have reviewed the structural information of several key FBEB complexes representing the four major families, including subunit composition, redox cofactor network, and the structure dynamics. However, very important questions remain about the details of the gating mechanisms in each of the four identified FBEB families. We do not yet know the minimal requirements of the sequence/structure motifs that allow flavin-based electron bifurcation to occur and distinguish a bifurcating enzyme from its non-bifurcating counterpart. A chimeric complex was recently made by joining the bifurcating modules of an FBEB complex with the non-bifurcating homologs of the remainder of the complex, but the chimeric complex had only low electron bifurcation activity (Vogt et al. 2019), indicating that either the interface between different modules or careful tuning of the cofactor redox potentials is important for FBEB function. Ultimately, it should be possible to expand the structure/function space of such oxidoreductases by combining novel metabolic modules with the FBEB modules, as nature has successfully done over the course of evolution (Raanan et al. 2020). The Bfu enzymes may represent the ultimate with regard to using an electron bifurcating module (BfuBC) to affect a wide array of metabolic reactions. Recently, time-resolved serial femtosecond crystallography has been successfully applied to understanding

photochemistry (Poddar et al. 2022). This method could be useful to understand the FBEB mechanism in detail, particularly in the NfnSL family, which represent the simplest of the bifurcating systems yet known, and the only one for which a viable mechanism has been proposed based on experimental data (Lubber et al. 2017). Finally, recent advances in cryoEM technology have led to the determination at atomic resolution of protein structures in which certain protons can be observed forming an H-bonding network (Nakane et al. 2020; Yip et al. 2020), raising the hope that we may be able to resolve the redox states along with the atomic configuration of the cofactors involved in the FBEB reactions. The role of protons in the proposed mechanisms of FBEB enzymes is rarely considered and clearly needs to be in the future.

Acknowledgments This work was funded by grants from the Division of Chemical Sciences, Geosciences and Biosciences, Office of Basic Energy Sciences of the US Department of Energy (DE-SC0020085 to H.L. and DE-FG02-95ER20175 to M.W.W.A.).

References

- Appel L, Willstein M, Dahl C, Ermiler U, Boll M (2021) Functional diversity of prokaryotic HdrA(BC) modules: role in flavin-based electron bifurcation processes and beyond. *Biochim Biophys Acta Bioenerg* 1862(4):148379. <https://doi.org/10.1016/j.bbambio.2021.148379>
- Baum H, Rieske JS, Silman HI, Lipton SH (1967) On the mechanism of electron transfer in complex iii of the electron transfer chain. *Proc Natl Acad Sci U S A* 57(3):798–805. <https://doi.org/10.1073/pnas.57.3.798>
- Baymann F, Schoepp-Cothenet B, Duval S, Guiral M, Brugna M, Baffert C, Russell MJ, Nitschke W (2018) On the natural history of flavin-based electron bifurcation. *Front Microbiol* 9:1357. <https://doi.org/10.3389/fmicb.2018.01357>
- Berrisford JM, Sazanov LA (2009) Structural basis for the mechanism of respiratory complex I. *J Biol Chem* 284(43):29773–29783. <https://doi.org/10.1074/jbc.M109.032144>
- Buckel W, Thauer RK (2018a) Flavin-based electron bifurcation, a new mechanism of biological energy coupling. *Chem Rev* 118(7):3862–3886. <https://doi.org/10.1021/acs.chemrev.7b00707>
- Buckel W, Thauer RK (2018b) Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD(+) (Rnf) as electron acceptors: a historical review. *Front Microbiol* 9:401. <https://doi.org/10.3389/fmicb.2018.00401>
- Chowdhury NP, Kломann K, Seubert A, Buckel W (2016) Reduction of flavodoxin by electron bifurcation and sodium ion-dependent reoxidation by NAD+ catalyzed by ferredoxin-NAD+ reductase (Rnf). *J Biol Chem* 291(23):11993–12002. <https://doi.org/10.1074/jbc.M116.726299>
- Chowdhury NP, Mowafy AM, Demmer JK, Upadhyay V, Koelzer S, Jayamani E, Kahnt J, Hornung M, Demmer U, Ermiler U, Buckel W (2014) Studies on the mechanism of electron bifurcation catalyzed by electron transferring flavoprotein (Etf) and butyryl-CoA dehydrogenase (Bcd) of *Acidaminococcus fermentans*. *J Biol Chem* 289(8):5145–5157. <https://doi.org/10.1074/jbc.M113.521013>
- Crofts AR (2004) The cytochrome bc1 complex: function in the context of structure. *Annu Rev Physiol* 66:689–733. <https://doi.org/10.1146/annurev.physiol.66.032102.150251>
- Crofts AR, Hong S, Wilson C, Burton R, Victoria D, Harrison C, Schulten K (2013) The mechanism of ubihydroquinone oxidation at the Qo-site of the cytochrome bc1 complex. *Biochim Biophys Acta* 1827(11–12):1362–1377. <https://doi.org/10.1016/j.bbambio.2013.01.009>

- Crofts AR, Lhee S, Crofts SB, Cheng J, Rose S (2006) Proton pumping in the bc₁ complex: a new gating mechanism that prevents short circuits. *Biochim Biophys Acta* 1757(8):1019–1034. <https://doi.org/10.1016/j.bbabi.2006.02.009>
- Demmer JK, Bertsch J, Oppinger C, Wohlers H, Kayastha K, Demmer U, Ermler U, Muller V (2018) Molecular basis of the flavin-based electron-bifurcating caffeoyl-CoA reductase reaction. *FEBS Lett* 592(3):332–342. <https://doi.org/10.1002/1873-3468.12971>
- Demmer JK, Huang H, Wang S, Demmer U, Thauer RK, Ermler U (2015) Insights into flavin-based electron bifurcation via the NADH-dependent reduced ferredoxin:NADP oxidoreductase structure. *J Biol Chem* 290(36):21985–21995. <https://doi.org/10.1074/jbc.M115.656520>
- Demmer JK, Rupprecht FA, Eisinger ML, Ermler U, Langer JD (2016) Ligand binding and conformational dynamics in a flavinbased electron-bifurcating enzyme complex revealed by Hydrogen-Deuterium Exchange Mass Spectrometry. *FEBS letters* 590(24):4472–4479. <https://doi.org/10.1002/1873-3468.12489>
- Demmer JK, Pal Chowdhury N, Selmer T, Ermler U, Buckel W (2017) The semiquinone swing in the bifurcating electron transferring flavoprotein/butyryl-CoA dehydrogenase complex from *Clostridium difficile*. *Nat Commun* 8(1):1577. <https://doi.org/10.1038/s41467-017-01746-3>
- Edgren T, Nordlund S (2004) The fixABCX genes in *Rhodospirillum rubrum* encode a putative membrane complex participating in electron transfer to nitrogenase. *J Bacteriol* 186(7):2052–2060. <https://doi.org/10.1128/jb.186.7.2052-2060.2004>
- Feng X, Schut GJ, Haja DK, Adams MWW, Li H (2022) Structure and electron transfer pathways of an electron-bifurcating NiFe-hydrogenase. *Sci Adv* 8(8):eabm7546. <https://doi.org/10.1126/sciadv.abm7546>
- Feng X, Schut GJ, Lipscomb GL, Li H, Adams MWW (2021) Cryoelectron microscopy structure and mechanism of the membrane-associated electron-bifurcating flavoprotein Fix/EtfABCX. *Proc Natl Acad Sci U S A* 118(2). <https://doi.org/10.1073/pnas.2016978118>
- Furlan C, Chongdar N, Gupta P, Lubitz W, Ogata H, Blaza JN, Birrell JA (2022) Structural insight on the mechanism of an electron-bifurcating [FeFe] hydrogenase. *eLife* 11. <https://doi.org/10.7554/eLife.79361>
- Garcia Costas AM, Poudel S, Miller AF, Schut GJ, Ledbetter RN, Fixen KR, Seefeldt LC, Adams MWW, Harwood CS, Boyd ES, Peters JW (2017) Defining electron bifurcation in the electron-transferring flavoprotein family. *J Bacteriol* 199(21). <https://doi.org/10.1128/jb.00440-17>
- Gonzalez-Viegas M, Kar RK, Miller AF, Mroginiski MA (2023) Noncovalent interactions that tune the reactivities of the flavins in bifurcating electron transferring flavoprotein. *J Biol Chem* 299(6):104762. <https://doi.org/10.1016/j.jbc.2023.104762>
- Hagen WR, Silva PJ, Amorim MA, Hagedoorn PL, Wassink H, Haaker H, Robb FT (2000) Novel structure and redox chemistry of the prosthetic groups of the iron-sulfur flavoprotein sulfide dehydrogenase from *Pyrococcus furiosus*; evidence for a [2Fe-2S] cluster with Asp(Cys)3 ligands. *J Biol Inorg Chem* 5(4):527–534. <https://doi.org/10.1007/pl00021452>
- Henriques BJ, Katrine Jentoft Olsen R, Gomes CM, Bross P (2021) Electron transfer flavoprotein and its role in mitochondrial energy metabolism in health and disease. *Gene* 776:145407. <https://doi.org/10.1016/j.gene.2021.145407>
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstern S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873):583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kaster AK, Moll J, Parey K, Thauer RK (2011) Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proc Natl Acad Sci U S A* 108(7):2981–2986. <https://doi.org/10.1073/pnas.1016761108>
- Katsy A, Kumar A, Saura P, Poverlein MC, Freibert SA, Jain S, Gamiz-Hernandez AP, VRI K, Muller V, Schuller JM (2023) Molecular basis of the electron bifurcation mechanism in the

- [FeFe]-hydrogenase complex HydABC. *J Am Chem Soc* 145(10):5696–5709. <https://doi.org/10.1021/jacs.2c11683>
- Kayastha K, Katsyov A, Himmrich C, Welsch S, Schuller JM, Ermler U, Muller V (2022) Structure-based electron-conformation mechanism of the Ldh-EtfAB complex. *eLife*:11. <https://doi.org/10.7554/eLife.77095>
- Kremp F, Roth J, Müller V (2020) The *Sporomusa* type Nfn is a novel type of electron-bifurcating transhydrogenase that links the redox pools in acetogenic bacteria. *Sci Rep* 10(1):14872. <https://doi.org/10.1038/s41598-020-71038-2>
- Kumar A, Kremp F, Roth J, Freibert SA, Muller V, Schuller JM (2023) Molecular architecture and electron transfer pathway of the Stn family transhydrogenase. *Nat Commun* 14(1):5484. <https://doi.org/10.1038/s41467-023-41212-x>
- Li F, Hinderberger J, Seedorf H, Zhang J, Buckel W, Thauer RK (2008) Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*. *J Bacteriol* 190(3):843–850. <https://doi.org/10.1128/jb.01417-07>
- Liang J, Huang H, Wang S (2019) Distribution, evolution, catalytic mechanism, and physiological functions of the flavin-based electron-bifurcating NADH-dependent reduced ferredoxin: NADP(+) oxidoreductase. *Front Microbiol* 10:373. <https://doi.org/10.3389/fmicb.2019.00373>
- Losey NA, Poudel S, Boyd ES, McInerney MJ (2020) The beta subunit of non-bifurcating NADH-dependent [FeFe]-hydrogenases differs from those of multimeric electron-bifurcating [FeFe]-hydrogenases. *Front Microbiol* 11. <https://doi.org/10.3389/fmicb.2020.01109>
- Lubner CE, Jennings DP, Mulder DW, Schut GJ, Zadovnyy OA, Hoben JP, Tokmina-Lukaszewska M, Berry L, Nguyen DM, Lipscomb GL, Bothner B, Jones AK, Miller AF, King PW, Adams MWW, Peters JW (2017) Mechanistic insights into energy conservation by flavin-based electron bifurcation. *Nat Chem Biol* 13(6):655–659. <https://doi.org/10.1038/nchembio.2348>
- Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) ColabFold: making protein folding accessible to all. *Nat Methods* 19(6):679–682. <https://doi.org/10.1038/s41592-022-01488-1>
- Mitchell P (1976) Possible molecular mechanisms of the protonmotive function of cytochrome systems. *J Theor Biol* 62(2):327–367. [https://doi.org/10.1016/0022-5193\(76\)90124-7](https://doi.org/10.1016/0022-5193(76)90124-7)
- Müller V, Chowdhury NP, Basen M (2018) Electron bifurcation: a long-hidden energy-coupling mechanism. *Annu Rev Microbiol* 72:331–353. <https://doi.org/10.1146/annurev-micro-090816-093440>
- Nakane T, Kotecha A, Sente A, McMullan G, Masiulis S, Brown P, Grigoras IT, Malinauskaitė L, Malinauskas T, Miehl J, Uchański T, Yu L, Karia D, Pechnikova EV, de Jong E, Keizer J, Bischoff M, McCormack J, Tiemeijer P, Hardwick SW, Chirgadze DY, Murshudov G, Aricescu AR, Scheres SHW (2020) Single-particle cryo-EM at atomic resolution. *Nature* 587(7832):152–156. <https://doi.org/10.1038/s41586-020-2829-0>
- Nguyen DMN, Schut GJ, Zadovnyy OA, Tokmina-Lukaszewska M, Poudel S, Lipscomb GL, Adams LA, Dinsmore JT, Nixon WJ, Boyd ES, Bothner B, Peters JW, Adams MWW (2017) Two functionally distinct NADP(+)-dependent ferredoxin oxidoreductases maintain the primary redox balance of *Pyrococcus furiosus*. *J Biol Chem* 292(35):14603–14616. <https://doi.org/10.1074/jbc.M117.794172>
- Oszycza A, Moser CC, Daldal F, Dutton PL (2004) Reversible redox energy coupling in electron transfer chains. *Nature* 427(6975):607–612. <https://doi.org/10.1038/nature02242>
- Page CC, Moser CC, Chen X, Dutton PL (1999) Natural engineering principles of electron tunneling in biological oxidation-reduction. *Nature* 402(6757):47–52. <https://doi.org/10.1038/46972>
- Peters JW, Beratan DN, Bothner B, Dyer RB, Harwood CS, Heiden ZM, Hille R, Jones AK, King PW, Lu Y, Lubner CE, Minteer SD, Mulder DW, Raugei S, Schut GJ, Seefeldt LC, Tokmina-Lukaszewska M, Zadovnyy OA, Zhang P, Adams MW (2018) A new era for electron bifurcation. *Curr Opin Chem Biol* 47:32–38. <https://doi.org/10.1016/j.cbpa.2018.07.026>

- Peters JW, Lanzilotta WN, Lemon BJ, Seefeldt LC (1998) X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. *Science* 282(5395):1853–1858. <https://doi.org/10.1126/science.282.5395.1853>
- Poddar H, Heyes DJ, Schirò G, Weik M, Leys D, Scrutton NS (2022) A guide to time-resolved structural analysis of light-activated proteins. *FEBS J* 289(3):576–595. <https://doi.org/10.1111/febs.15880>
- Poudel S, Dunham EC, Lindsay MR, Amenabar MJ, Fones EM, Colman DR, Boyd ES (2018) Origin and evolution of flavin-based electron bifurcating enzymes. *Front Microbiol* 9:1762. <https://doi.org/10.3389/fmicb.2018.01762>
- Raanan H, Poudel S, Pike DH, Nanda V, Falkowski PG (2020) Small protein folds at the root of an ancient metabolic network. *Proc Natl Acad Sci U S A* 117(13):7193–7199. <https://doi.org/10.1073/pnas.1914982117>
- Radon C, Mittelstadt G, Duffus BR, Burger J, Hartmann T, Mielke T, Teutloff C, Leimkuhler S, Wendler P (2020) Cryo-EM structures reveal intricate Fe-S cluster arrangement and charging in *Rhodobacter capsulatus* formate dehydrogenase. *Nat Commun* 11(1):1912. <https://doi.org/10.1038/s41467-020-15614-0>
- Rowland P, Nørager S, Jensen KF, Larsen S (2000) Structure of dihydroorotate dehydrogenase B: electron transfer between two flavin groups bridged by an iron-sulphur cluster. *Structure* 8(12):1227–1238. [https://doi.org/10.1016/s0969-2126\(00\)00530-x](https://doi.org/10.1016/s0969-2126(00)00530-x)
- Sarewicz M, Pintscher S, Pietras R, Borek A, Bujnowicz L, Hanke G, Cramer WA, Finazzi G, Osyczka A (2021) Catalytic reactions and energy conservation in the cytochrome bc(1) and b(6)f complexes of energy-transducing membranes. *Chem Rev* 121(4):2020–2108. <https://doi.org/10.1021/acs.chemrev.0c00712>
- Schut GJ, Adams MW (2009) The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. *J Bacteriol* 191(13):4451–4457. <https://doi.org/10.1128/jb.01582-08>
- Schut GJ, Haja DK, Feng X, Poole FL, Li H, Adams MWW (2022) An abundant and diverse new family of electron bifurcating enzymes with a non-canonical catalytic mechanism. *Front Microbiol* 13:946711. <https://doi.org/10.3389/fmicb.2022.946711>
- Schut GJ, Mohamed-Raseek N, Tokmina-Lukaszewska M, Mulder DW, Nguyen DMN, Lipscomb GL, Hoben JP, Patterson A, Lubner CE, King PW, Peters JW, Bothner B, Miller AF, Adams MWW (2019) The catalytic mechanism of electron-bifurcating electron transfer flavoproteins (ETFs) involves an intermediary complex with NAD. *J Biol Chem* 294(9):3271–3283. <https://doi.org/10.1074/jbc.RA118.005653>
- Schut GJ, Thorgersen MP, Poole FL 2nd, Haja DK, Putumbaka S, Adams MWW (2021) Tungsten enzymes play a role in detoxifying food and antimicrobial aldehydes in the human gut microbiome. *Proc Natl Acad Sci U S A* 118(43). <https://doi.org/10.1073/pnas.2109008118>
- Shomura Y, Taketa M, Nakashima H, Tai H, Nakagawa H, Ikeda Y, Ishii M, Igarashi Y, Nishihara H, Yoon KS, Ogo S, Hirota S, Higuchi Y (2017) Structural basis of the redox switches in the NAD(+)-reducing soluble [NiFe]-hydrogenase. *Science* 357(6354):928–932. <https://doi.org/10.1126/science.aan4497>
- Swuec P, Chaves-Sanjuan A, Camilloni C, Vanoni MA, Bolognesi M (2019) Cryo-EM structures of *Azospirillum brasilense* glutamate synthase in its oligomeric assemblies. *J Mol Biol* 431(22):4523–4526. <https://doi.org/10.1016/j.jmb.2019.08.011>
- Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R (2008) Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat Rev Microbiol* 6(8):579–591. <https://doi.org/10.1038/nrmicro1931>
- Toogood HS, Leys D, Scrutton NS (2007) Dynamics driving function: new insights from electron transferring flavoproteins and partner complexes. *FEBS J* 274(21):5481–5504. <https://doi.org/10.1111/j.1742-4658.2007.06107.x>
- Vogt MS, Schuhle K, Kolzer S, Peschke P, Chowdhury NP, Kleinsorge D, Buckel W, Essen LO, Heider J (2019) Structural and functional characterization of an electron transfer flavopro-

- tein involved in toluene degradation in strictly anaerobic bacteria. *J Bacteriol.* <https://doi.org/10.1128/jb.00326-19>
- Wagner T, Koch J, Ermiler U, Shima S (2017) Methanogenic heterodisulfide reductase (HdrABC-MvhAGD) uses two noncubane [4Fe-4S] clusters for reduction. *Science* 357(6352):699–703. <https://doi.org/10.1126/science.aan0425>
- Watanabe T, Pfeil-Gardiner O, Kahnt J, Koch J, Shima S, Murphy BJ (2021) Three-megadalton complex of methanogenic electron-bifurcating and CO(2)-fixing enzymes. *Science* 373(6559):1151–1156. <https://doi.org/10.1126/science.abg5550>
- Wikström MK, Berden JA (1972) Oxidoreduction of cytochrome b in the presence of antimycin. *Biochim Biophys Acta* 283(3):403–420. [https://doi.org/10.1016/0005-2728\(72\)90258-7](https://doi.org/10.1016/0005-2728(72)90258-7)
- Xia D, Esser L, Tang WK, Zhou F, Zhou Y, Yu L, Yu CA (2013) Structural analysis of cytochrome bc1 complexes: implications to the mechanism of function. *Biochim Biophys Acta* 1827(11–12):1278–1294. <https://doi.org/10.1016/j.bbabi.2012.11.008>
- Yan Z, Ferry JG (2018) Electron bifurcation and confurcation in methanogenesis and reverse methanogenesis. *Front Microbiol* 9:1322. <https://doi.org/10.3389/fmicb.2018.01322>
- Yip KM, Fischer N, Paknia E, Chari A, Stark H (2020) Atomic-resolution protein structure determination by cryo-EM. *Nature* 587(7832):157–161. <https://doi.org/10.1038/s41586-020-2833-4>
- Yuly JL, Zhang P, Lubner CE, Peters JW, Beratan DN (2020) Universal free-energy landscape produces efficient and reversible electron bifurcation. *Proc Natl Acad Sci U S A* 117(35):21045–21051. <https://doi.org/10.1073/pnas.2010815117>
- Zhang J, Frerman FE, Kim JJ (2006) Structure of electron transfer flavoprotein-ubiquinone oxidoreductase and electron transfer to the mitochondrial ubiquinone pool. *Proc Natl Acad Sci U S A* 103(44):16212–16217. <https://doi.org/10.1073/pnas.0604567103>
- Zuchan K, Baymann F, Baffert C, Brugna M, Nitschke W (2021) The dyad of the Y-junction- and a flavin module unites diverse redox enzymes. *Biochim Biophys Acta Bioenerg* 1862(6):148401. <https://doi.org/10.1016/j.bbabi.2021.148401>

Chapter 15

The Prothrombin-Prothrombinase Interaction



Bosko M. Stojanovski, Bassem M. Mohammed, and Enrico Di Cera

Abstract The hemostatic response to vascular injury entails a sequence of proteolytic events where several inactive zymogens of the trypsin family are converted to active proteases. The cascade starts with exposure of tissue factor from the damaged endothelium and culminates with conversion of prothrombin to thrombin in a reaction catalyzed by the prothrombinase complex composed of the enzyme factor Xa, cofactor Va, Ca^{2+} , and phospholipids. This cofactor-dependent activation is paradigmatic of analogous reactions of the blood coagulation and complement cascades, which makes elucidation of its molecular mechanism of broad significance to the large class of trypsin-like zymogens to which prothrombin belongs. Because of its relevance as the most important reaction in the physiological response to vascular injury, as well as the main trigger of pathological thrombotic complications, the mechanism of prothrombin activation has been studied extensively. However, a molecular interpretation of this mechanism has become available only recently from important developments in structural biology. Here we review current knowledge on the prothrombin-prothrombinase interaction and outline future directions for the study of this key reaction of the coagulation cascade.

Keywords Blood coagulation · Factor V · Factor Va · Prothrombin · Vitamin K-dependent clotting factors

B. M. Stojanovski · B. M. Mohammed · E. Di Cera (✉)

Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University
School of Medicine, St. Louis, MO, USA

e-mail: bosko.stojanovski@health.slu.edu; bassem.mohammed@health.slu.edu;
enrico@slu.edu

Introduction

The hemostatic response to vascular injury entails a cascade of proteolytic events where several inactive zymogens of the trypsin family are converted to active proteases (Davie et al. 1991). Clot formation is initiated along the extrinsic pathway by exposure of tissue factor at the site of injury and complex formation with factor (f) VIIa, which in turn activates small quantities of fXa sufficient to convert trace amounts of prothrombin into thrombin (Mann 2003). Thrombin then initiates a series of reactions that accelerate its own generation including activation of cofactors V and VIII (Camire and Bos 2009) and zymogen fXI (Gailani and Broze 1991). The cascade culminates with the explosive conversion of prothrombin into thrombin in a reaction catalyzed by the prothrombinase complex composed of fXa, cofactor fVa, Ca^{2+} , and phospholipids (Mann et al. 2003). Blood coagulation can also be triggered along the intrinsic pathway, where fXIa converts fIX to fIXa and the fIXa-fVIIIa complex efficiently activates fX to fXa. Attenuation of the coagulation response averts excessive clotting and is mediated by various anticoagulant proteins including tissue factor pathway inhibitor (TFPI), antithrombin III, and activated protein C (APC) (Mann et al. 2003).

Because of its relevance as the most important reaction in the physiological response to vascular injury, as well as the main trigger of pathological thrombotic complications, the mechanism of prothrombin activation has been studied extensively (Mann et al. 2003; Krishnaswamy 2013). However, a molecular interpretation of this mechanism has been difficult to formulate with the support of structural biology. Controversial models of prothrombin activation have been proposed from standard biochemical approaches, envisioning conformational transitions of prothrombinase (Brufatto and Nesheim 2003; Kim and Nesheim 2007) or prothrombin (Orcutt and Krishnaswamy 2004; Bianchini et al. 2005). Significant developments have relied on more direct biophysical approaches based on single molecule Förster resonance energy transfer (smFRET) measurements (Pozzi et al. 2016a), X-ray crystallography (Pozzi et al. 2016b; Chinnaraj et al. 2018), and cryogenic electron microscopy (cryo-EM) (Ruben et al. 2022; Di Cera et al. 2022). A molecular picture of the prothrombin-prothrombinase interaction is finally emerging and, with that, some new lines of investigation of this and related reactions of the coagulation cascade.

Prothrombin, Factor V, and Factor X

Prothrombin is a vitamin K-dependent coagulation factor synthesized in the liver and circulates in the blood as an inactive zymogen (Degen and Davie 1987). The modular structure of prothrombin comprises an N-terminal Gla domain that promotes binding to membrane surfaces, kringle 1 and kringle 2, which serve as spacers and a protease domain that hosts the active site (Mann et al. 1990). Prothrombin

is converted into thrombin following cleavage at R271, located in the linker that connects kringle 2 to the protease domain, and R320 in the A chain. Cleavage at R271 sheds the Gla and kringle domains, while cleavage at R320 promotes folding of the catalytic B chain according to the Huber-Bode mechanism of zymogen activation (Huber and Bode 1978). Thrombin then performs various procoagulant, prothrombotic, and anticoagulant roles that are essential for the effective amplification and attenuation of the coagulation response (Di Cera 2008). The prothrombotic role of thrombin primarily stems from the ability to activate platelets through cleavage of the protease-activated receptor (PAR) 1 (Vu et al. 1991), and the procoagulant roles involves the conversion of fibrinogen into an insoluble clot to prevent bleeding at the site of injury (Di Cera 2008). Paradoxically, thrombin also functions as an anticoagulant by activating protein C with the assistance of the endothelial receptor thrombomodulin, resulting in inactivation of fVa and fVIIIa and attenuation of the coagulation response (Esmon 2003).

Coagulation fV is a large glycoprotein that circulates in the blood at a concentration of 20 nM and features an A1-A2-B-A3-C1-C2 domain assembly analogous to that of fVIII (Mann and Kalafatis 2003; Rosing and Tans 1997). The A domains share sequence homology with ceruloplasmin, and the C domains display low sequence homology with the discoidin domain of lectins (Jenny et al. 1987). The large B domain is poorly conserved across vertebrates and shares no significant sequence homology with any known proteins. Thrombin activates fV by cleaving at R709, R1018, and R1545 and removing the B domain in its entirety (Camire and Bos 2009). However, an intramolecular interaction within the B domain between a basic region (BR) and acidic region (AR) locks fV in its inactive form (Bos and Camire 2012) because disruption of this interaction produces fVa-like activity in the absence of complete proteolytic removal of the B domain (Toso and Camire 2004). The splice variant fV short carries a large deletion of the B domain that includes the BR (Vincent et al. 2013). Interestingly, the variant has fVa-like activity, but the unmasked AR provides a high affinity binding site for the C-terminal basic region of TFPI α that together with protein S (PS) hijacks fXa in a resulting fV short/TFPI α /PS/fXa complex and produces a bleeding phenotype (Dahlback 2023). Inactivation of fVa is catalyzed by APC upon cleavage at R306 and R506 (Kalafatis et al. 1994). Among the two cleavage sites, APC shows kinetic preference for R506, but the second cleavage at R306 promotes dissociation of the A2 domain and is needed for complete loss of cofactor function (Rosing and Tans 1997). Interestingly, incorporation of fVa into the prothrombinase complex offers protection against the APC-mediated proteolytic cleavage at R506 but not at R306 (Norstrom et al. 2006). The importance of the proteolytic downregulation of fVa is evident from the fV^{R506Q} Leiden mutation associated with APC resistance and increased thrombotic phenotype (Bertina et al. 1994).

Coagulation fX is a two-chain, vitamin K-dependent glycoprotein (Camire 2021) that shares a domain organization with fVII, fIX, PC, and protein Z, indicating that these proteins all evolved from an archetypal protein in the common vertebrate ancestor (Stojanovski and Di Cera 2022). The multidomain structure comprises a membrane binding Gla domain, two epidermal growth factor (EGF1 and EGF2)

domains, and a protease domain hosting the active site. When coagulation is initiated along the extrinsic pathway, fX is activated to fXa by the TF-fVIIa and fIXa-fVIIIa complexes. The latter complex activates fX along the intrinsic pathway (Camire 2021). Incorporation of fXa into the prothrombinase complex enhances activity toward prothrombin and protects fXa from inactivation by antithrombin (Brufatto et al. 2003). The activity of fXa is also regulated by TFPI, especially during the initiation phase of clotting (Mast and Ruf 2022).

Kinetics of the Prothrombin-Prothrombinase Interaction

The conversion of prothrombin into thrombin proceeds along two alternative pathways, depending on whether cleavage is initiated at R271 or R320 (Mann et al. 2003; Mann 2003), two residues that are absolutely conserved in vertebrates (Stojanovski and Di Cera 2022). Cleavage at R271 sheds the auxiliary Gla and kringle domains and generates the inactive intermediate prethrombin-2. The alternative cleavage at R320 separates the A and B chains, which remain connected through a disulfide bond, and generates the active intermediate meizothrombin (Fig. 15.1). Selection of the pathway is context-dependent *in vivo*. Activation proceeds along the prethrombin-2 pathway on the surface of platelets (Haynes et al. 2012; Wood et al. 2010), but along the meizothrombin pathway on red blood cells (Whelihan et al. 2012). *In vitro*, the presence of fVa drives activation along the meizothrombin

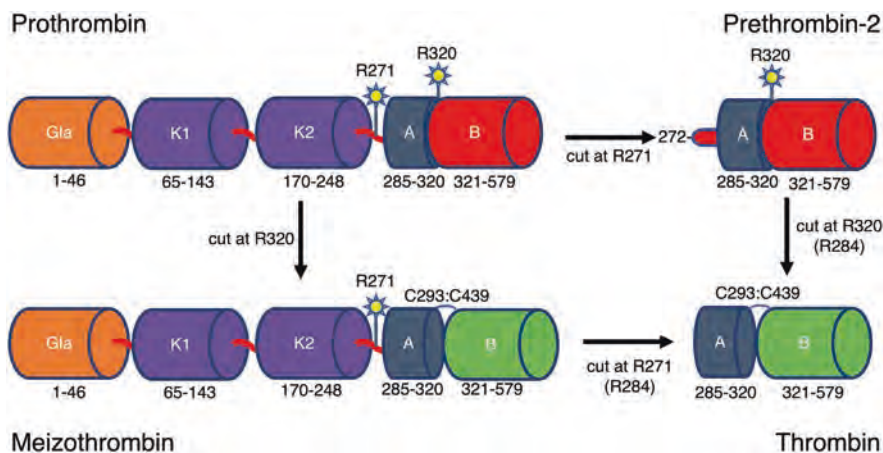


Fig. 15.1 Prothrombin activation. Schematic representation of prothrombin activation by the prothrombinase complex. The conversion to thrombin entails cleavage at R271 in Lnk3 and R320 in the A chain, along two alternative pathways generating the intermediates prethrombin-2 (cleavage at R271 first) or meizothrombin (cleavage at R320 first). An additional autoproteolytic cleavage at R284 produces thrombin. The presence of fVa, *in vitro*, promotes activation along the meizothrombin pathway. In the absence of fVa, activation proceeds along the prethrombin-2 pathway. Modified from Di Cera et al. (2022)

pathway (Rosing et al. 1986; Krishnaswamy et al. 1986), with and without phospholipids (Rosing et al. 1986; Boskovic et al. 1990; Tans et al. 1991). In the absence of fVa, activation proceeds along the prethrombin-2 pathway (Esmon et al. 1974) or with cleavage at R155 to generate prethrombin-1 (Rosing et al. 1980).

Several models have been proposed to account for how prothrombinase can preferentially cleave at R271 or R320. These models are based on studies using liposomes and fail to recapitulate the physiologically relevant platelet surface, where prothrombinase cleaves preferentially at R271 instead of R320. Preferential cleavage at R320 observed in the presence of liposomes has been explained in terms of more favorable docking (Lee et al. 2010) and alternative conformations of prothrombinase (Brufatto and Nesheim 2003; Kim and Nesheim 2007) or prothrombin (Orcutt and Krishnaswamy 2004; Bianchini et al. 2005). Direct evidence of alternative conformations of prothrombinase has been difficult to capture by structural biology or biophysical methods. In the case of fXa, there is evidence of alternative conformations preferentially populated upon interaction with fVa and prothrombin (Brandstetter et al. 1996). Despite significant level of disorder in the EGF1 domain in several crystal structures of fXa, it was argued that the EGF1 and EGF2 domains assume an extended conformation in the original tetragonal crystal structure of fXa, whereas in the inhibitor-bound orthorhombic structure, fXa populates a compact conformation with the EGF1 domain coming proximally to the protease domain (Brandstetter et al. 1996). Further evidence that fXa undergoes conformational changes upon binding to fVa have been provided by fluorescence resonance energy transfer studies, revealing that the distance of the active site of fXa relative to the membrane surface changes upon cofactor binding (Husten et al. 1987). On the other hand, single molecule studies have shown that prothrombin exists in equilibrium between open and closed forms (Pozzi et al. 2016a) linked to breaking and making of an intramolecular contact between kringle 1 and the protease domain (Chinnaraj et al. 2018). This equilibrium has direct relevance to selection of the pathway of activation. Prothrombin variants stabilized in the closed form are preferentially cleaved at R320, while variants stabilized in the open form are preferentially cleaved at R271 (Pozzi et al. 2016a; Chinnaraj et al. 2018). An additional factor directing fXa cleavage of prothrombin is the extended sequence surrounding R271 and R320 (Stojanovski and Di Cera 2021). In the absence of fVa, fXa cleaves preferentially the sequence at R271, regardless of its position. On the other hand, the presence of fVa always directs cleavage at the R320 position, regardless of its sequence. This implies that prothrombinase has an intrinsic preference for the position of R320, which has important implications for how prothrombin is recognized along the two alternative pathways of activation.

Incorporation of fVa in the prothrombinase complex produces a profound enhancement in the rate of prothrombin activation, which is more pronounced for cleavage of R320 (~40,000-fold increase) compared to R271 in prothrombin (450-fold increase) or meizothrombin (34-fold increase) (Brufatto and Nesheim 2003). The molecular origin of the poor specificity of fXa for prothrombin when fVa is absent likely stems from factors that adversely affect the rate of productive diffusion of substrate into the active site (effect on second order rate constant k_{on}) and efficient

turnover of substrate by the enzyme-substrate complex (effect on k_{cat}) (Di Cera et al. 2022). The role of fVa is to correct these defects by increasing the rate of productive diffusion and increasing the turnover rate. The structural details of how fVa enhances k_{cat} may be difficult to capture. Small rearrangements of the catalytic residues translate in large perturbations of k_{cat} as recently shown by replacement of the active site Ser with the nearly equivalent nucleophile Thr. The substitution in fXa greatly diminishes the fVa-mediated enhancement of prothrombin activation (Pelc et al. 2015), proving that a large component of the effect of fVa on k_{cat} is achieved by allosterically optimizing the rotamer of the OH group of the catalytic Ser of fXa. On the other hand, the effect of fVa on k_{on} would require pre-organization of prothrombinase for a rapid, “rigid-body” association with prothrombin (Di Cera 2020). These conclusions derived from the kinetics of prothrombin activation await validation by structural studies.

Structural Biology of Coagulation Factors and Their Complexes

Solving the structure of coagulation factors and their complexes remains a high priority for the field of hemostasis and thrombosis but constitutes a significant challenge. Most of these proteins have a molecular weight too high for NMR investigation, which has been limited to thrombin and prethrombin-2 (Lechtenberg et al. 2010; Fuglestad et al. 2012; Ruben et al. 2020) composed of the protease domain only (Fig. 15.1). X-ray crystallography has succeeded with thrombin (Bode et al. 1992; Pineda et al. 2004); proteases bearing EGF domains like fVIIa (Banner et al. 1996), fIXa (Hopfner et al. 1999; Brandstetter et al. 1995), and fXa (Kamata et al. 1998; Shi et al. 2009; Padmanabhan et al. 1993); zymogens like fXI (Papagrigoriou et al. 2006) and prothrombin (Chinnaraj et al. 2018; Pozzi et al. 2013, 2014, 2016b); and cofactors like fVIII (Shen et al. 2008) and fVa in a truncated form (Adams et al. 2004). A few complexes have been crystallized, like thrombin bound to the EGF456 domains of thrombomodulin (Fuentes-Prior et al. 2000; Barranco-Medina et al. 2017), and fVIIa bound to soluble tissue factor (Banner et al. 1996). Though remarkable, these successes have also shown the limits of X-ray crystallography. For example, the structures of important anticoagulant proteins like PC, PS, and TFPI have been pursued for decades without success. Among proteins directly involved in prothrombin activation (Fig. 15.1), meizothrombin has been crystallized without Gla and kringle 1 domains (Martin et al. 1997; Papaconstantinou et al. 2008), fXa without the Gla domain (Kamata et al. 1998; Shi et al. 2009; Padmanabhan et al. 1993), and fVa in its truncated APC-inactivated form (Adams et al. 2004). Prothrombin has been crystallized without the Gla domain (Pozzi et al. 2013), portions of the linker connecting the two kringles (Pozzi et al. 2014, 2016b), or with an engineered disulfide bond to artificially lock its conformation in the closed form (Chinnaraj et al. 2018). In all these cases, conformational flexibility or intrinsic

disorder has made it necessary to remove large segments of the protein or to introduce significant structural constraints. There is no X-ray structure of prothrombinase in complex with prothrombin or meizothrombin and prothrombinase itself has been crystallized only as a surrogate from the snake *Pseudonaja textilis* (Lechtenberg et al. 2013). Alternative approaches are needed to advance our structural understanding of the mechanism of prothrombin activation (Fig. 15.1). A revolution in structural biology has recently come from cryo-EM, a technique capable of solving the structure of high molecular weight proteins, unlike NMR, and under conditions not requiring alignment in a crystal for diffraction, unlike X-ray crystallography (Bonomi and Vendruscolo 2018; Cheng 2018; Frank 2018; Di Cera et al. 2022). The unique ability of cryo-EM to handle large macromolecules in a native-like environment has been key to its recent success in the analysis of coagulation factors and their complexes (Di Cera et al. 2022; Ruben et al. 2021, 2022; Fuller et al. 2021; Gish et al. 2021; Childers et al. 2022; Anderson et al. 2022; Javitt et al. 2020; Feng et al. 2019; Alvadia et al. 2019; Irons et al. 2013; Mohammed et al. 2023).

The Prothrombin-Prothrombinase Complex

We have determined the structure of the prothrombin-prothrombinase complex that offers a view of how prothrombinase initiates prothrombin activation along the meizothrombin pathway (Di Cera et al. 2022; Ruben et al. 2022). In this structure, prothrombin and fXa face each other in curved conformations against the fVa scaffold. Prothrombin aligns along the A2, A1, and C2 domains in a conformation similar to the closed form identified by X-ray crystallography (Chinnaraj et al. 2018), and fXa is positioned along the A2, A3, and C1 domains with a kink between the Gla and EGF1 domains that puts the active site 70 Å above the plane of the membrane (Fig. 15.2). Importantly, prothrombin engages prothrombinase only through its protease domain that binds to the gate segment (residues 696–702) in the A2 domain of fVa and to the active site region of fXa. The paucity of extensive contacts between prothrombin and prothrombinase accounts for an interaction limited only by diffusion between pre-organized components (Kim and Nesheim 2007; Mann et al. 1990; Di Cera 2020). There are no contacts between exosite I and the autolysis loop of prothrombin with either fXa or fVa, contrary to the suggestions of molecular models (Lechtenberg et al. 2013; Lee et al. 2011; Shim et al. 2015) and biochemical studies (Bock et al. 2007; Krishnaswamy 2005, 2013), reviewed recently (Schreuder et al. 2019). The mechanism of prothrombin activation involves a drop of the lid on fXa and a conformational change of the gate that sequesters R271 on one side and provides a surface on the other side to present R320 to the active site of fXa (Di Cera et al. 2022). Importantly, the sequence ⁶⁹⁵DYDYQ⁶⁹⁹ within the gate region was identified in previous studies as a competitive inhibitor of prothrombin activation by prothrombinase, capable of shifting cleavage from R320 to R271 (Beck et al. 2004; Kalafatis 2005; Kalafatis et al. 2003; Bukys et al. 2006). The cryo-EM structure of the prothrombin-prothrombinase complex shows how fVa directs R320 to the active

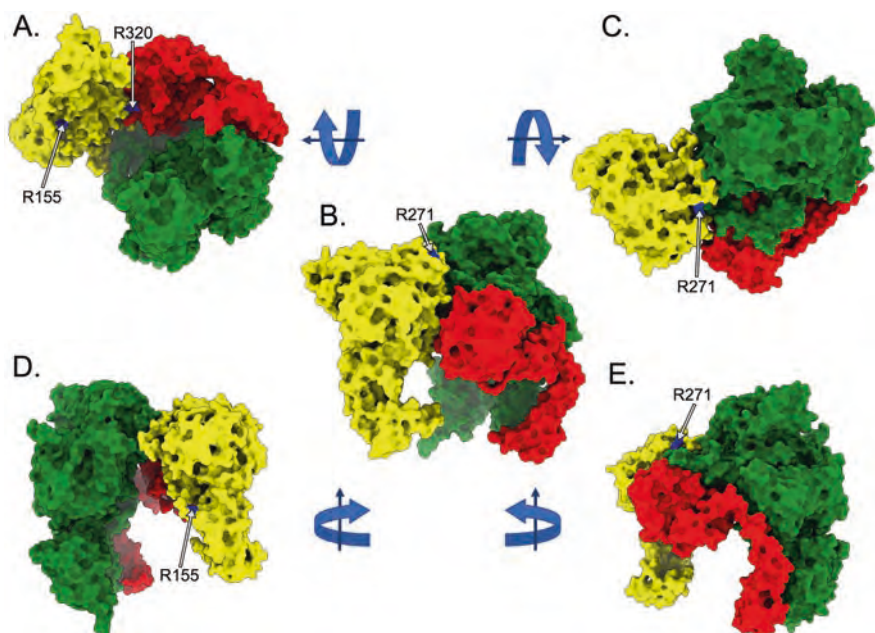


Fig. 15.2 (a–e) Cryo-EM structure of the prothrombin-prothrombinase complex (PDB 7TTP). Molecular surface representation of the ternary complex of prothrombin (yellow), fVa (green), and fXa (red). The complex is shown in different orientations obtained by 90° rotation of the view in the center (b). Prothrombin contacts prothrombinase only through the protease domain (b,d,e), that binds to the A2 domain of fVa (b,c,d) and the protease domain of fXa (a,b,c). The sites of prothrombin activation by prothrombinase at R271 (b,c,e), R320 (a), and R155 (d) are in blue

site of fXa and prevents alternative cleavages at R271 and R155, consistent with biochemical evidence that the presence of fVa drives activation along the meizothrombin pathway, *in vitro* (Rosing et al. 1980, 1986; Krishnaswamy et al. 1986), with or without phospholipids (Rosing et al. 1986; Boskovic et al. 1990; Tans et al. 1991), and in a position-dependent manner (Stojanovski and Di Cera 2021).

It is possible that the absence of a membrane surface may have biased the arrangement of the three components in the complex. To this end, we have determined a preliminary structure of the prothrombin-prothrombinase complex on nanodiscs (Fig. 15.3). The structure shows an arrangement of the complex that is essentially similar to that reported in the absence of nanodiscs, except for the Gla domain of prothrombin that shifts slightly toward the Gla domain of fXa but remain separate from prothrombinase. The structure also confirms that prothrombin contacts prothrombinase only through its protease domain and that prothrombinase is largely pre-organized to bind prothrombin in a conformation that presents R320 to the active site of fXa. That explains the large enhancement effect of fVa on the k_{on} of the interaction-, with the membrane assisting in this rapid interaction by reducing the dimensionality of the system (Di Cera 2020).

The structure of the prothrombin-prothrombinase complex also offers context for interpreting the intriguing properties of fV short, a splice isoform of fV associated with the East Texas bleeding disorder and featuring a large deletion of the disordered B domain (Dahlbäck 2016; Vincent et al. 2013; Broze and Girard 2013) (Fig. 15.4a). Most of the differences between fV short and fV reside in the A2 domain (Mohammed et al. 2023) that provides a suitable scaffold for the interaction with fXa and prothrombin that could explain the fVa-like activity of this variant (Bos and Camire 2012; Camire 2011; Toso and Camire 2004; Camire and Bos 2009; Cheng 2018). A possible complex with prothrombin and fXa is obtained by replacing fVa with fV short (Fig. 15.4b). Interestingly, the shorter B domain running in the front of fV short from side to side provides no hindrance to the binding of prothrombin and fXa. Small rearrangements of the protease domain of prothrombin and the EGF2 domain of fXa would provide optimal docking of the two components into a productive complex with fV short (Mohammed et al. 2023).

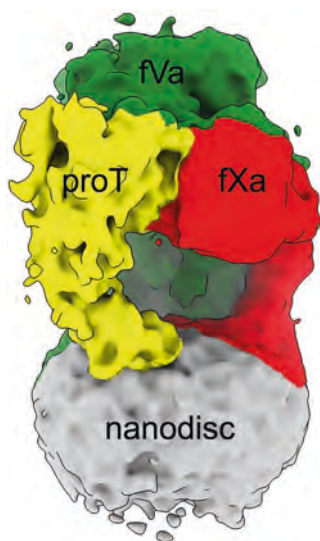


Fig. 15.3 Preliminary cryo-EM structure of the prothrombin-prothrombinase complex on nanodiscs. A de novo analysis of the original micrographs from which the EMD-26061 map (PDB 7TPQ) was generated yields a 3D reconstruction in which prothrombin and prothrombinase come together on phospholipid nanodiscs. Further classification and refinement with cryoSPARC v4.3.1 generated a new 3D map (EMD-42405) from a total of 4988 particles showing an envelope of the complex at 6.5 Å resolution. Importantly, the arrangement of the three proteins on nanodiscs is very similar to that reported in the absence of nanodiscs (Fig. 15.2), confirming that prothrombin contacts prothrombinase only through its protease domain. The curved conformations of prothrombin and fXa are retained. The presence of nanodiscs causes the Gla domain of prothrombin to shift closer but not in contact to fXa

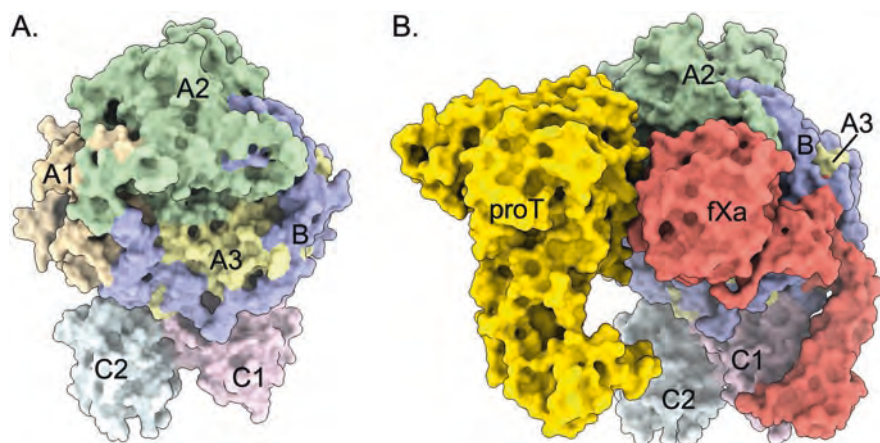


Fig. 15.4 (a, b) Cryo-EM structure of fV short. (a) The cryo-EM structure of fV short reveals the entire A1-A2-B-A3-C1-C2 assembly. The protein is rendered in surface representation with the constitutive domains colored in wheat (A1), pale green (A2), light blue (B), pale yellow (A3), light pink (C1), pale cyan (C2). The B domain (133 residues total) is resolved in its entirety and stretches across the entire width of the protein, from R709 through R1545, making contacts with the A1, A2, and A3 domains and suspended over the C1 and C2 domains. (b) A complex of prothrombin (yellow) and fXa (red) bound to fV short (colored as in panel a) is generated by replacing fVa in the prothrombin-prothrombinase complex (Fig. 15.2) with fV short (see panel a), without further optimization. The B domain in the front of fV short (a) poses no steric hindrance to the binding of prothrombin and fXa (b)

Summary

The structure of the prothrombin-prothrombinase complex is an important step toward future cryo-EM studies of this and related interactions relevant to blood coagulation. We are at the start of an exciting new chapter in the study of coagulation factors and their complexes. Ongoing efforts on the complex of meizothrombin bound to prothrombinase are unraveling the structural underpinnings of the interaction of R271 with the active site of fXa and the conformational rearrangements needed in prothrombinase to accommodate a substrate similar but not identical to prothrombin. Likewise, ongoing studies of fV short bound to fXa will elucidate similarities and differences between this complex and prothrombinase. In turn, this will enable the analysis of more challenging complexes involving prothrombin and the TFPI α /PS pair, thereby revealing the structural basis of the procoagulant and anticoagulant functions of fV short. Success with the prothrombin-prothrombinase complex will motivate cryo-EM investigation of the PC-thrombin-thrombomodulin complex, as well as of similar macromolecular interactions in the intrinsic pathway like the fVIIIa/fIXa/fX and fXIa/fIX complexes.

Acknowledgments We gratefully acknowledge Tracey Baird for her assistance with illustrations. Katherine Basore and Brock Summers at Washington University Center for Cellular Imaging

provided expert assistance with the cryo-EM work. Cassandra Kukla and Sheila Rahman were involved in early developments of this article. This study was supported in part by the National Institutes of Health Research Grants HL049413, HL139554, and HL147821.

References

- Adams TE, Hockin MF, Mann KG, Everse SJ (2004) The crystal structure of activated protein C-inactivated bovine factor Va: implications for cofactor function. *Proc Natl Acad Sci U S A* 101(24):8918–8923. <https://doi.org/10.1073/pnas.0403072101>
- Alvadia C, Lim NK, Clerico Mosina V, Oostergetel GT, Dutzler R, Paulino C (2019) Cryo-EM structures and functional characterization of the murine lipid scramblase TMEM16F. *Elife* 8. <https://doi.org/10.7554/eLife.44365>
- Anderson JR, Li J, Springer TA, Brown A (2022) Structures of VWF tubules before and after concatemerization reveal a mechanism of disulfide bond exchange. *Blood* 140(12):1419–1430. <https://doi.org/10.1182/blood.2022016467>
- Banner DW, D'Arcy A, Chene C, Winkler FK, Guha A, Konigsberg WH, Nemerson Y, Kirchhofer D (1996) The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature* 380(6569):41–46. <https://doi.org/10.1038/380041a0>
- Barranco-Medina S, Murphy M, Pelc L, Chen Z, Di Cera E, Pozzi N (2017) Rational design of protein C activators. *Sci Rep* 7:44596. <https://doi.org/10.1038/srep44596>
- Beck DO, Bukys MA, Singh LS, Szabo KA, Kalafatis M (2004) The contribution of amino acid region ASP695-TYR698 of factor V to procofactor activation and factor Va function. *J Biol Chem* 279(4):3084–3095. <https://doi.org/10.1074/jbc.M306850200>
- Bertina RM, Koелеman BP, Koster T, Rosendaal FR, Dirven RJ, de Ronde H, van der Velden PA, Reitsma PH (1994) Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 369(6475):64–67
- Bianchini EP, Orcutt SJ, Panizzi P, Bock PE, Krishnaswamy S (2005) Ratcheting of the substrate from the zymogen to proteinase conformations directs the sequential cleavage of prothrombin by prothrombinase. *Proc Natl Acad Sci U S A* 102(29):10099–10104
- Bock PE, Panizzi P, Verhamme IM (2007) Exosites in the substrate specificity of blood coagulation reactions. *J Thromb Haemost* 5(Suppl 1):81–94
- Bode W, Turk D, Karshikov A (1992) The refined 1.9-Å X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human alpha-thrombin: structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships. *Protein Sci* 1(4):426–471
- Bonomi M, Vendruscolo M (2018) Determination of protein structural ensembles using cryo-electron microscopy. *Curr Opin Struct Biol* 56:37–45. <https://doi.org/10.1016/j.sbi.2018.10.006>
- Bos MH, Camire RM (2012) A bipartite autoinhibitory region within the B-domain suppresses function in factor V. *J Biol Chem* 287(31):26342–26351. <https://doi.org/10.1074/jbc.M112.377168>
- Boskovic DS, Giles AR, Nesheim ME (1990) Studies of the role of factor Va in the factor Xa-catalyzed activation of prothrombin, fragment 1.2-prethrombin-2, and dansyl-L-glutamyl-glycyl-L-arginine-meizothrombin in the absence of phospholipid. *J Biol Chem* 265(18):10497–10505
- Brandstetter H, Bauer M, Huber R, Lollar P, Bode W (1995) X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and hemophilia B. *Proc Natl Acad Sci U S A* 92(21):9796–9800. <https://doi.org/10.1073/pnas.92.21.9796>
- Brandstetter H, Kuhne A, Bode W, Huber R, von der Saal W, Wirthensohn K, Engh RA (1996) X-ray structure of active site-inhibited clotting factor Xa. Implications for drug design and substrate recognition. *J Biol Chem* 271(47):29988–29992

- Broze GJ Jr, Girard TJ (2013) Factor V, tissue factor pathway inhibitor, and east Texas bleeding disorder. *J Clin Invest* 123(9):3710–3712. <https://doi.org/10.1172/JCI71220>
- Brufatto N, Nesheim ME (2003) Analysis of the kinetics of prothrombin activation and evidence that two equilibrating forms of prothrombinase are involved in the process. *J Biol Chem* 278(9):6755–6764. <https://doi.org/10.1074/jbc.M206413200>
- Brufatto N, Ward A, Nesheim ME (2003) Factor Xa is highly protected from antithrombin-fondaparinux and antithrombin-enoxaparin when incorporated into the prothrombinase complex. *J Thromb Haemost* 1(6):1258–1263. <https://doi.org/10.1046/j.1538-7836.2003.00254.x>
- Bukys MA, Kim PY, Nesheim ME, Kalafatis M (2006) A control switch for prothrombinase: characterization of a hirudin-like pentapeptide from the COOH terminus of factor Va heavy chain that regulates the rate and pathway for prothrombin activation. *J Biol Chem* 281(51):39194–39204. <https://doi.org/10.1074/jbc.M604482200>
- Camire RM (2011) A new look at blood coagulation factor V. *Curr Opin Hematol* 18(5):338–342. <https://doi.org/10.1097/MOH.0b013e3283497ebc>
- Camire RM (2021) Blood coagulation factor X: molecular biology, inherited disease, and engineered therapeutics. *J Thromb Thrombolysis* 52(2):383–390. <https://doi.org/10.1007/s11239-021-02456-w>
- Camire RM, Bos MH (2009) The molecular basis of factor V and VIII procofactor activation. *J Thromb Haemost* 7(12):1951–1961. <https://doi.org/10.1111/j.1538-7836.2009.03622.x>
- Cheng Y (2018) Single-particle cryo-EM-How did it get here and where will it go. *Science* 361(6405):876–880. <https://doi.org/10.1126/science.aat4346>
- Childers KC, Peters SC, Spiegel PC Jr (2022) Structural insights into blood coagulation factor VIII: procoagulant complexes, membrane binding, and antibody inhibition. *J Thromb Haemost* 20(9):1957–1970. <https://doi.org/10.1111/jth.15793>
- Chinnaraj M, Chen Z, Pelc LA, Grese Z, Bystranowska D, Di Cera E, Pozzi N (2018) Structure of prothrombin in the closed form reveals new details on the mechanism of activation. *Sci Rep* 8(1):2945. <https://doi.org/10.1038/s41598-018-21304-1>
- Dahlback B (2023) Natural anticoagulant discovery, the gift that keeps on giving: finding FV-Short. *J Thromb Haemost* 21(4):716–727. <https://doi.org/10.1016/j.jtha.2023.01.033>
- Dahlbäck B (2016) Pro- and anticoagulant properties of factor V in pathogenesis of thrombosis and bleeding disorders. *Int J Lab Hematol* 38(Suppl 1):4–11. <https://doi.org/10.1111/ijlh.12508>
- Davie EW, Fujikawa K, Kisiel W (1991) The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry* 30(43):10363–10370
- Degen SJ, Davie EW (1987) Nucleotide sequence of the gene for human prothrombin. *Biochemistry* 26(19):6165–6177
- Di Cera E (2008) Thrombin. *Mol Aspects Med* 29(4):203–254
- Di Cera E (2020) Mechanisms of ligand binding. *Biophys Rev* 1(1):011303. <https://doi.org/10.1063/5.0020997>
- Di Cera E, Mohammed BM, Pelc LA, Stojanovski BM (2022) Cryo-EM structures of coagulation factors. *Res Pract Thromb Haemost* 6(7):e12830. <https://doi.org/10.1002/rth2.12830>
- Esmon CT (2003) The protein C pathway. *Chest* 124(3 Suppl):26S–32S
- Esmon CT, Owen WG, Jackson CM (1974) The conversion of prothrombin to thrombin. V. The activation of prothrombin by factor Xa in the presence of phospholipid. *J Biol Chem* 249(24):7798–7807
- Feng S, Dang S, Han TW, Ye W, Jin P, Cheng T, Li J, Jan YN, Jan LY, Cheng Y (2019) Cryo-EM studies of TMEM16F calcium-activated ion channel suggest features important for lipid scrambling. *Cell Rep* 28(2):567–579 e564. <https://doi.org/10.1016/j.celrep.2019.06.023>
- Frank J (2018) Single-particle reconstruction of biological molecules-story in a sample (Nobel Lecture). *Angew Chem Int Ed Engl* 57(34):10826–10841. <https://doi.org/10.1002/anie.201802770>
- Fuentes-Prior P, Iwanaga Y, Huber R, Pagila R, Rumennik G, Seto M, Morser J, Light DR, Bode W (2000) Structural basis for the anticoagulant activity of the thrombin-thrombomodulin complex. *Nature* 404(6777):518–525

- Fuglestad B, Gasper PM, Tonelli M, McCammon JA, Markwick PR, Komives EA (2012) The dynamic structure of thrombin in solution. *Biophys J* 103(1):79–88. <https://doi.org/10.1016/j.bpj.2012.05.047>. S0006-3495(12)00627-3 [pii]
- Fuller JR, Knockenhauer KE, Leksa NC, Peters RT, Batchelor J (2021) Molecular determinants of the factor VIII/von Willebrand factor complex revealed by BIVV001 cryo-electron microscopy. *Blood* 137(21):2970–2980. <https://doi.org/10.1182/blood.2020009197>
- Gailani D, Broze GJ Jr (1991) Factor XI activation in a revised model of blood coagulation. *Science* 253(5022):909–912
- Gish JS, Jarvis L, Childers KC, Peters SC, Garrels CS, Smith I, Spencer HT, Doering CB, Lollar P, Spiegel PC (2021) Structure of blood coagulation factor VIII in complex with an anti-C1 domain pathogenic antibody inhibitor. *Blood* 137(21):2981–2986. <https://doi.org/10.1182/blood.2020008940>
- Haynes LM, Bouchard BA, Tracy PB, Mann KG (2012) Prothrombin activation by platelet-associated prothrombinase proceeds through the prethrombin-2 pathway via a concerted mechanism. *J Biol Chem* 287(46):38647–38655. <https://doi.org/10.1074/jbc.M112.407791>. M112.407791 [pii]
- Hopfner KP, Lang A, Karcher A, Sichler K, Kopetzki E, Brandstetter H, Huber R, Bode W, Engh RA (1999) Coagulation factor IXa: the relaxed conformation of Tyr99 blocks substrate binding. *Structure* 7(8):989–996. [https://doi.org/10.1016/s0969-2126\(99\)80125-7](https://doi.org/10.1016/s0969-2126(99)80125-7)
- Huber R, Bode W (1978) Structural basis of the activation and action of trypsin. *Acc Chem Res* 11:114–122
- Husten EJ, Esmen CT, Johnson AE (1987) The active site of blood coagulation factor Xa. Its distance from the phospholipid surface and its conformational sensitivity to components of the prothrombinase complex. *J Biol Chem* 262(27):12953–12961
- Irons EE, Flatt JW, Doronin K, Fox TL, Acchione M, Stewart PL, Shayakhmetov DM (2013) Coagulation factor binding orientation and dimerization may influence infectivity of adenovirus-coagulation factor complexes. *J Virol* 87(17):9610–9619. <https://doi.org/10.1128/JVI.01070-13>
- Javitt G, Khmelnitsky L, Albert L, Bigman LS, Elad N, Morgenstern D, Ilani T, Levy Y, Diskin R, Fass D (2020) Assembly mechanism of mucin and von Willebrand factor polymers. *Cell* 183(3):717–729 e716. <https://doi.org/10.1016/j.cell.2020.09.021>
- Jenny RJ, Pittman DD, Toole JJ, Kriz RW, Aldape RA, Hewick RM, Kaufman RJ, Mann KG (1987) Complete cDNA and derived amino acid sequence of human factor V. *Proc Natl Acad Sci U S A* 84(14):4846–4850. <https://doi.org/10.1073/pnas.84.14.4846>
- Kalafatis M (2005) Coagulation factor V: a plethora of anticoagulant molecules. *Curr Opin Hematol* 12(2):141–148. <https://doi.org/10.1097/01.moh.0000155016.30296.90>
- Kalafatis M, Beck DO, Mann KG (2003) Structural requirements for expression of factor Va activity. *J Biol Chem* 278(35):33550–33561. <https://doi.org/10.1074/jbc.M303153200>
- Kalafatis M, Rand MD, Mann KG (1994) The mechanism of inactivation of human factor V and human factor Va by activated protein C. *J Biol Chem* 269(50):31869–31880
- Kamata K, Kawamoto H, Honma T, Iwama T, Kim SH (1998) Structural basis for chemical inhibition of human blood coagulation factor Xa. *Proc Natl Acad Sci U S A* 95(12):6630–6635. <https://doi.org/10.1073/pnas.95.12.6630>
- Kim PY, Nesheim ME (2007) Further evidence for two functional forms of prothrombinase each specific for either of the two prothrombin activation cleavages. *J Biol Chem* 282(45):32568–32581
- Krishnaswamy S (2005) Exosite-driven substrate specificity and function in coagulation. *J Thromb Haemost* 3(1):54–67
- Krishnaswamy S (2013) The transition of prothrombin to thrombin. *J Thromb Haemost* 11(Suppl 1):265–276. <https://doi.org/10.1111/jth.12217>
- Krishnaswamy S, Mann KG, Nesheim ME (1986) The prothrombinase-catalyzed activation of prothrombin proceeds through the intermediate meizothrombin in an ordered, sequential reaction. *J Biol Chem* 261(19):8977–8984

- Lechtenberg BC, Johnson DJ, Freund SM, Huntington JA (2010) NMR resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc Natl Acad Sci U S A* 107:14087–14092. <https://doi.org/10.1073/pnas.1005255107>. 1005255107 [pii]
- Lechtenberg BC, Murray-Rust TA, Johnson DJ, Adams TE, Krishnaswamy S, Camire RM, Huntington JA (2013) Crystal structure of the prothrombinase complex from the venom of *Pseudonaja textilis*. *Blood* 122(16):2777–2783. <https://doi.org/10.1182/blood-2013-06-511733>
- Lee CJ, Wu S, Eun C, Pedersen LG (2010) A revisit to the one form kinetic model of prothrombinase. *Biophys Chem* 149(1–2):28–33. <https://doi.org/10.1016/j.bpc.2010.03.011>. S0301-4622(10)00070-0 [pii]
- Lee CJ, Wu S, Pedersen LG (2011) A proposed ternary complex model of prothrombinase with prothrombin: protein-protein docking and molecular dynamics simulations. *J Thromb Haemost* 9(10):2123–2126. <https://doi.org/10.1111/j.1538-7836.2011.04463.x>
- Mann KG (2003) Thrombin formation. *Chest* 124(3 Suppl):4S–10S
- Mann KG, Butenas S, Brummel K (2003) The dynamics of thrombin formation. *Arterioscler Thromb Vasc Biol* 23(1):17–25
- Mann KG, Kalafatis M (2003) Factor V: a combination of Dr Jekyll and Mr Hyde. *Blood* 101(1):20–30. <https://doi.org/10.1182/blood-2002-01-0290>
- Mann KG, Nesheim ME, Church WR, Haley P, Krishnaswamy S (1990) Surface-dependent reactions of the vitamin K-dependent enzyme complexes. *Blood* 76(1):1–16
- Martin PD, Malkowski MG, Box J, Esmen CT, Edwards BF (1997) New insights into the regulation of the blood clotting cascade derived from the X-ray crystal structure of bovine meizothrombin des F1 in complex with PPACK. *Structure* 5(12):1681–1693
- Mast AE, Ruf W (2022) Regulation of coagulation by tissue factor pathway inhibitor: implications for hemophilia therapy. *J Thromb Haemost* 20(6):1290–1300. <https://doi.org/10.1111/jth.15697>
- Mohammed BM, Pelc LA, Rau MJ, Di Cera E (2023) Cryo-EM structure of coagulation factor V short. *Blood* 141(26):3215–3225. <https://doi.org/10.1182/blood.2022019486>
- Norstrom EA, Tran S, Steen M, Dahlback B (2006) Effects of factor Xa and protein S on the individual activated protein C-mediated cleavages of coagulation factor Va. *J Biol Chem* 281(42):31486–31494. <https://doi.org/10.1074/jbc.M606441200>
- Orcutt SJ, Krishnaswamy S (2004) Binding of substrate in two conformations to human prothrombinase drives consecutive cleavage at two sites in prothrombin. *J Biol Chem* 279:54927–54936
- Padmanabhan K, Padmanabhan KP, Tulinsky A, Park CH, Bode W, Huber R, Blankenship DT, Cardin AD, Kisiel W (1993) Structure of human des(1–45) factor Xa at 2.2 Å resolution. *J Mol Biol* 232(3):947–966
- Papaconstantinou ME, Gandhi PS, Chen Z, Bah A, Di Cera E (2008) Na(+) binding to meizothrombin desF1. *Cell Mol Life Sci* 65(22):3688–3697
- Papagrigoriou E, McEwan PA, Walsh PN, Emsley J (2006) Crystal structure of the factor XI zymogen reveals a pathway for transactivation. *Nat Struct Mol Biol* 13(6):557–558
- Pelc LA, Chen Z, Gohara DW, Vogt AD, Pozzi N, Di Cera E (2015) Why Ser and not Thr brokers catalysis in the trypsin fold. *Biochemistry* 54:1457–1464. <https://doi.org/10.1021/acs.biochem.5b00014>
- Pineda AO, Carrell CJ, Bush LA, Prasad S, Caccia S, Chen ZW, Mathews FS, Di Cera E (2004) Molecular dissection of Na+ binding to thrombin. *J Biol Chem* 279(30):31842–31853
- Pozzi N, Bystranowska D, Zuo X, Di Cera E (2016a) Structural architecture of prothrombin in solution revealed by single molecule spectroscopy. *J Biol Chem* 291(35):18107–18116. <https://doi.org/10.1074/jbc.M116.738310>
- Pozzi N, Chen Z, Di Cera E (2016b) How the linker connecting the two kringles influences activation and conformational plasticity of prothrombin. *J Biol Chem* 291:6071–6082. <https://doi.org/10.1074/jbc.M115.700401>
- Pozzi N, Chen Z, Gohara DW, Niu W, Heyduk T, Di Cera E (2013) Crystal structure of prothrombin reveals conformational flexibility and mechanism of activation. *J Biol Chem* 288(31):22734–22744. <https://doi.org/10.1074/jbc.M113.466946>. M113.466946 [pii]

- Pozzi N, Chen Z, Pelc LA, Shropshire DB, Di Cera E (2014) The linker connecting the two kringles plays a key role in prothrombin activation. *Proc Natl Acad Sci U S A* 111(21):7630–7635. <https://doi.org/10.1073/pnas.1403779111>
- Rosing J, Tans G (1997) Factor V. *Int J Biochem Cell Biol* 29(10):1123–1126. [https://doi.org/10.1016/s1357-2725\(97\)00040-x](https://doi.org/10.1016/s1357-2725(97)00040-x)
- Rosing J, Tans G, Govers-Riemslog JW, Zwaal RF, Hemker HC (1980) The role of phospholipids and factor Va in the prothrombinase complex. *J Biol Chem* 255(1):274–283
- Rosing J, Zwaal RF, Tans G (1986) Formation of meizothrombin as intermediate in factor Xa-catalyzed prothrombin activation. *J Biol Chem* 261(9):4224–4228
- Ruben EA, Gandhi PS, Chen Z, Koester SK, DeKoster GT, Frieden C, Di Cera E (2020) (19) F NMR reveals the conformational properties of free thrombin and its zymogen precursor prethrombin-2. *J Biol Chem* 295(24):8227–8235. <https://doi.org/10.1074/jbc.RA120.013419>
- Ruben EA, Rau MJ, Fitzpatrick J, Di Cera E (2021) Cryo-EM structures of human coagulation factors V and Va. *Blood* 137(22):3137–3144. <https://doi.org/10.1182/blood.2021010684>
- Ruben EA, Summers B, Rau MJ, Fitzpatrick J, Di Cera E (2022) Cryo-EM structure of the prothrombin-prothrombinase complex. *Blood* 139(24):3463–3473. <https://doi.org/10.1182/blood.2022015807>
- Schreuder M, Reitsma PH, Bos MHA (2019) Blood coagulation factor Va's key interactive residues and regions for prothrombinase assembly and prothrombin binding. *J Thromb Haemost* 17(8):1229–1239. <https://doi.org/10.1111/jth.14487>
- Shen BW, Spiegel PC, Chang CH, Huh JW, Lee JS, Kim J, Kim YH, Stoddard BL (2008) The tertiary structure and domain organization of coagulation factor VIII. *Blood* 111(3):1240–1247. <https://doi.org/10.1182/blood-2007-08-109918>
- Shi Y, Li C, O'Connor SP, Zhang J, Shi M, Bisaha SN, Wang Y, Sitkoff D, Pudzianowski AT, Huang C, Klei HE, Kish K, Yanchunas J Jr, Liu EC, Hartl KS, Seiler SM, Steinbacher TE, Schumacher WA, Atwal KS, Stein PD (2009) Aroylguanidine-based factor Xa inhibitors: the discovery of BMS-344577. *Bioorg Med Chem Lett* 19(24):6882–6889. <https://doi.org/10.1016/j.bmcl.2009.10.084>
- Shim JY, Lee CJ, Wu S, Pedersen LG (2015) A model for the unique role of factor Va A2 domain extension in the human ternary thrombin-generating complex. *Biophys Chem* 199:46–50. <https://doi.org/10.1016/j.bpc.2015.02.003>
- Stojanovski BM, Di Cera E (2021) Role of sequence and position of the cleavage sites in prothrombin activation. *J Biol Chem* 297(2):100955. <https://doi.org/10.1016/j.jbc.2021.100955>
- Stojanovski BM, Di Cera E (2022) Comparative sequence analysis of vitamin K-dependent coagulation factors. *J Thromb Haemost* 20(12):2837–2849. <https://doi.org/10.1111/jth.15897>
- Tans G, Janssen-Claessen T, Hemker HC, Zwaal RF, Rosing J (1991) Meizothrombin formation during factor Xa-catalyzed prothrombin activation. Formation in a purified system and in plasma. *J Biol Chem* 266(32):21864–21873
- Toso R, Camire RM (2004) Removal of B-domain sequences from factor V rather than specific proteolysis underlies the mechanism by which cofactor function is realized. *J Biol Chem* 279(20):21643–21650. <https://doi.org/10.1074/jbc.M402107200>
- Vincent LM, Tran S, Livaja R, Bensend TA, Milewicz DM, Dahlback B (2013) Coagulation factor V(A2440G) causes east Texas bleeding disorder via TFPIalpha. *J Clin Invest* 123(9):3777–3787. <https://doi.org/10.1172/JCI69091>
- Vu TK, Hung DT, Wheaton VI, Coughlin SR (1991) Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell* 64(6):1057–1068. 0092-8674(91)90261-V [pii]
- Whelihan MF, Zachary V, Orfeo T, Mann KG (2012) Prothrombin activation in blood coagulation: the erythrocyte contribution to thrombin generation. *Blood* 120(18):3837–3845. <https://doi.org/10.1182/blood-2012-05-427856>. blood-2012-05-427856 [pii]
- Wood JP, Silveira JR, Maille NM, Haynes LM, Tracy PB (2010) Prothrombin activation on the activated platelet surface optimizes expression of procoagulant activity. *Blood* 117:1710–1718. <https://doi.org/10.1182/blood-2010-09-311035>. blood-2010-09-311035 [pii]

Chapter 16

Structure and Function of the Glycosylphosphatidylinositol Transamidase, a Transmembrane Complex Catalyzing GPI Anchoring of Proteins



Dianfan Li

Abstract Glycosylphosphatidylinositol (GPI) anchoring of proteins is a ubiquitous posttranslational modification in eukaryotic cells. GPI-anchored proteins (GPI-APs) play critical roles in enzymatic, signaling, regulatory, and adhesion processes. Over 20 enzymes are involved in GPI synthesis, attachment to client proteins, and remodeling after attachment. The GPI transamidase (GPI-T), a large complex located in the endoplasmic reticulum membrane, catalyzes the attachment step by replacing a C-terminal signal peptide of proproteins with GPI. In the last three decades, extensive research has been conducted on the mechanism of the transamidation reaction, the components of the GPI-T complex, the role of each subunit, and the substrate specificity. Two recent studies have reported the three-dimensional architecture of GPI-T, which represent the first structures of the pathway. The structures provide detailed mechanisms for assembly that rationalizes previous biochemical results and subunit-dependent stability data. While the structural data confirm the catalytic role of PIGK, which likely uses a caspase-like mechanism to cleave the proproteins, they suggest that unlike previously proposed, GPAA1 is not a catalytic subunit. The structures also reveal a shared cavity for GPI binding. Somewhat unexpectedly, PIGT, a single-pass membrane protein, plays a crucial role in GPI recognition. Consistent with the assembly mechanisms and the active site architecture, most of the disease mutations occur near the active site or the subunit interfaces. Finally, the catalytic dyad is located ~ 22 Å away from the membrane interface of the GPI-binding site, and this architecture may confer substrate specificity through topological matching between the substrates and the elongated active site. The research conducted thus far sheds light on the intricate processes involved in GPI anchoring and paves the way for further mechanistic studies of GPI-T.

D. Li (✉)

CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences (CAS), Shanghai, China

e-mail: dianfan.li@sibcb.ac.cn

Keywords Glycosylphosphatidylinositol anchoring · GPI-AP biosynthesis · GPI attachment · GPI transamidase · GPI-T structure · Membrane complex

Biological and Pharmaceutical Importance of GPI-APs and the Biosynthesis Pathway

Glycosylphosphatidylinositol (GPI) anchoring of proteins is a posttranslational modification conserved among all eukaryotes (Beihammer et al. 2020; Hong and Kinoshita 2009; Kinoshita 2020; Pittet and Conzelmann 2007). Being tethered at the cell surface, GPI-anchored proteins (GPI-APs) execute crucial enzymatic, signaling, and adhesion functions (Orlean and Menon 2007; Pittet and Conzelmann 2007). In humans, there are over 150 GPI-APs (Kinoshita 2020), including well-known examples such as alkaline phosphatase that serve as a marker for hepatic diseases, CD59 and CD55 as complement inhibitors, and prion that causes transmissible and fatal neurodegenerative diseases (summarized in Kinoshita 2020). Although some human cell lines tolerate the lack of GPI-APs (Liu et al. 2021), disruption of the GPI-AP biosynthesis leads to embryonic lethality in animals (McKean and Niswander 2012). Further, genetic defects in GPI-AP biosynthesis cause severe developmental disorders (summarized in Kinoshita (2014, 2020)), and causative mutations in new loci are expanding (Ben Ayed et al. 2023; Cash et al. 2020; Castle et al. 2021; Duval et al. 2021; Holtz et al. 2020; Jezela-Stanek et al. 2020a, 2020b; Johnstone et al. 2020; Li et al. 2020; Loong et al. 2023; Nguyen et al. 2020; Peron et al. 2020; Salian et al. 2021, 2022; Thompson et al. 2020, 2023; Tremblay-Laganière et al. 2021; Vetro et al. 2020; Zanni et al. 2022). Moreover, some members of the biosynthesis pathway are oncogenes (Wu et al. 2006; Zhang et al. 2019) that serve as biomarkers for related cancers (Gamage and Hendrickson 2013).

Appropriate GPI-AP biogenesis is also essential for lower eukaryotes, including pathogenic fungi and parasites (Hong and Kinoshita 2009; Pittet and Conzelmann 2007). In *Saccharomyces cerevisiae*, there are two classes of GPI-APs. The minor class is retained in the plasma membrane. The majority of GPI-APs are transferred to the cell wall, a process that involves the removal of the glucosaminylphosphatidylinositol moiety followed by the covalent link between the mannose residues of the remnant GPI-AP and cell wall β 1,6-glucan (Pittet and Conzelmann 2007). Genetic disruption of the GPI-AP biosynthesis pathway in *S. cerevisiae* leads to lethality or conditional lethality, including temperature and chemical sensitivity. In *Candida albicans*, the leading cause of fungal infections worldwide, genetic impairment of GPI-AP synthesis reduces its virulence (Liu et al. 2016; Richard et al. 2002; Yadav et al. 2014), and chemical inhibition of the pathway suppresses its growth (Watanabe et al. 2012) and enhances immunogenicity (Huang et al. 2020). In animal models, chemical inhibitors can prevent mice from infection by *Candida auris* (Wiederhold et al. 2019) and *Aspergillus fumigatus* (Gebremariam et al.

2019), both of which are classified as critical priority for antifungal research by the World Health Organization. Moreover, GPI-AP synthesis inhibitors can prevent mice from infection by *Coccidioides immitis* (Viriyakosol et al. 2019) and *Rhizopus deleamar* (Fu et al. 2020), which also cause fatal infectious diseases.

GPI-AP biosynthesis is also of critical importance for single-cell parasites (Hong and Kinoshita 2009). In *Trypanosoma brucei*, the pathogen responsible for the fatal African sleeping sickness, a GPI-AP known as the variant surface glycoprotein (VSG) makes up 90% of all its surface protein. By densely coating the parasite and constant changing its sequence and glycosylation pattern, VSG helps evade host immune surveillance. Therefore, the GPI-AP biogenesis pathway is a chemically validated drug target against *T. brucei* (Smith et al. 2004). In *Plasmodium falciparum*, the parasite that causes more than half a million deaths each year (mostly children), GPI-APs are critical for malaria pathogenesis by acting as endotoxins (Naik et al. 2000). Treating *P. falciparum* with known GPI-AP synthesis inhibitors such as manogepix, gepinacin, and salicylic hydroxamic acid can arrest cell growth during the trophozoite stage of intraerythrocytic asexual development (Fenollar et al. 2022). Taken together, the study of GPI-AP biosynthesis is of both biological and pharmaceutical importance.

Unique Characteristics of GPI-APs

The chemical structure of GPI was first elucidated in the 1980s using GPI-APs from *T. brucei* (Ferguson et al. 1988). Since then, structures of GPI from other species (Güther et al. 1992; Schneider et al. 1990), including humans (Homans et al. 1988), have also been determined. These findings reveal a common backbone consisting of a phosphatidylinositol group linked to a polysaccharide core α -Man3-(1 $\rightarrow$ 2)- α -Man2-(1 $\rightarrow$ 6)- α -Man1-(1 $\rightarrow$ 4)- α -GlcN (Man1/2/3, the three mannoses; GlcN, glucosamine; italic numbers indicate carbon numbering) (Fig. 16.1a). The glycan core contains a Man3-linked ethanolamine phosphate (EtNP) group that bridges GPI to client proteins in most cases. Beyond this minimal requirement, GPI can have diverse modifications. Depending on the maturation stages, species, and tissues, the glycan core can be further decorated with EtNPs and other phosphates, as well as complex saccharides such as acetylgalactosaminyl-galactosyl sialic acid (McConville and Ferguson 1993; Tsai et al. 2012). A Man4 residue linking to the C2 of Man3 is essential for GPI anchoring in some species such as *S. cerevisiae* (Benachour et al. 1999). Similarly, the fatty chains can exist in diverse forms of diacyl- and diradylglycerols (Kanzawa et al. 2009), and ceramides (Conzelmann et al. 1992) with varying length and degree of saturation (Kinoshita and Fujita 2016). The inositol ring undergoes palmitoylation. While this third fatty chain is removed in most mature GPI-APs, it stays in the procyclic form of *T. brucei* (Hong and Kinoshita 2009) and in most (Richier et al. 1992; Walter et al. 1990), but not all (Becker Kojić and Terness 2002), GPI-APs of the human erythrocyte cells.

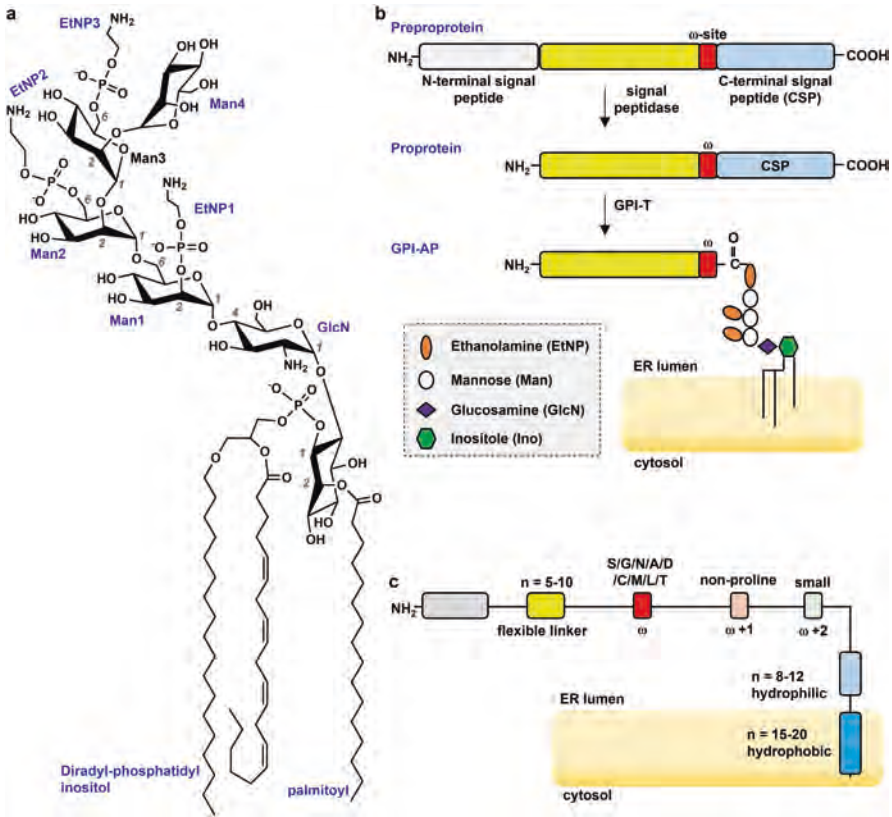


Fig. 16.1 Chemical structure of GPI and scheme of GPI-T reaction. **(a)** Chemical structure of a typical GPI. Various parts are labeled in blue. Italic numbers indicate carbon atoms of the sugar residues. The lipid chains are drawn as the mostly frequent ones from mammal cells. *EtNP* ethanolamine phosphate, *GlcN* glucosamine, *Man* mannose. **(b)** Scheme of GPI anchoring. Client proteins are processed by signal peptidases to generate proproteins. GPI-T then replaces the C-terminal signal peptide with GPI. **(c)** Scheme of a proprotein. Characteristics of residues at various positions are indicated

The unique chemical structure of GPI gives client proteins with features that differ from those of transmembrane proteins. The fatty chains are usually long and saturated, enabling GPI-APs to enrich in lipid rafts (Brown and Rose 1992; Maeda et al. 2007; Zurzolo et al. 1994), which are microdomains with higher membrane thickness. Lipid rafts, also known as the signaling hub, are enriched with cholesterol and other bioactive lipids such sphingolipids that recruit and regulate signaling complexes. Concentrated in these microdomains, GPI-APs form low-specificity and transient homo- or hetero-oligomers through lipid-, glycan-, protein-, or metal ion-mediated interactions (Caiola et al. 2007; Lebreton et al. 2021; Paladino et al. 2017; Seong et al. 2013; Sezgin et al. 2017; Sharma et al. 2004; Suzuki et al. 2012). Such clustering would increase the apparent affinity with their interacting proteins. For signaling GPI-APs, this may add another level of regulation. Furthermore,

GPI-APs can be released by GPI-specific enzymes or proteases in a regulated manner. This process, known as shedding, participates in important biological activities such as fertilization and neurogenesis (reviewed in Fujihara and Ikawa 2016).

A Brief Overview of the Biochemical Steps in GPI-AP Biosynthesis

The biosynthesis of GPI-AP involves over 20 enzymatic steps. Because the catalysis steps occur in different compartments, namely, the endoplasmic reticulum (ER) and Golgi apparatus, vesicle transport is also involved in shuttling the lipidated proteins during the biosynthesis (Kinoshita et al. 2013). Comprehensive reviews of the biochemical synthesis of GPI-AP can be found elsewhere (Liu and Fujita 2020; Pittet and Conzelmann 2007), and this chapter will only provide a brief overview.

The synthesis of GPI-AP begins with the addition of the acetylated glucosamine to phosphatidylinositol (PI) on the cytosolic leaflet of the ER membrane. This reaction is catalyzed by a transmembrane enzyme called GPI-GnT (Murakami et al. 2005), which uses a uridine diphosphate-activated sugar donor (UDP-N-acetylglucosamine). Current knowledge suggests that GPI-GnT contains seven subunits, five of which are essential. The reason for requiring such a large complex to catalyze a seemingly simple monoglycosyltransfer reaction with moderately sized substrates remains a structural mystery.

After the transfer, PIGL, a single-pass membrane enzyme with the active site facing the cytosol, removes the acetyl group on the glucosamine (Watanabe et al. 1999). The single transmembrane helix (TMH) of PIGL may act as an anchor or confer substrate selectivity by approximating the active site near the membrane.

The glucosylamine-modified PI undergoes an intramembrane flip facilitated by the recently (Wang et al. 2022) identified CLPTM1L (cleft-lip and palate-associated transmembrane protein 1-like) and possibly other nonspecific flippases that are yet to be identified. CLPTM1L is predicted to have eight TMHs with structural similarity to SWEET transporters (Jumper et al. 2021). CLPTM1L is mainly located in the ER in HEK293 cells, but it can also be found in the plasma membrane of some carcinoma and antibody-producing cells. Its overexpression is associated with several cancer types (Wang et al. 2014), including lung (Ni et al. 2012), cervical (Lando et al. 2009), and pancreas cancers (Jia et al. 2014), as well as cisplatin resistance in human ovarian cancer cells (Yamamoto et al. 2001); in fact, CLPTM1L is also known as cisplatin resistance-related protein 9 (CRR9). Monoclonal antibodies targeting CLPTM1L have been shown to re-sensitize drug-resistant cancer cell (Parashar et al. 2021; Puskás et al. 2016). In addition, mutations of the encoding gene are associated with breast (Haiman et al. 2011) and lung cancer (McKay et al. 2008; Rafnar et al. 2009). However, the link between its scramblase activity and roles in carcinogenesis remains to be elucidated.

At the lumen leaflet of the ER membrane, PIGW, a predicted 13-transmembrane protein with no known structural homology to other proteins, catalyzes the

acylation of the inositol moiety at the hydroxyl group on C2 position using acyl coenzyme A as the donor (Chiyonobu et al. 2014). PIGW is a validated drug target for pathogenic fungi, as chemically blocking this step disrupts their GPI-AP synthesis and inhibits their growth (Watanabe et al. 2012). Further, mutations in PIGW that confer resistance to inhibitors have been identified (Mann et al. 2015). Several compounds, including E1210, BIQ, and G884, have been shown to inhibit fungal PIGW but not the human homolog (Yadav and Khan 2018), making PIGW an attractive target for developing antimycotic drugs.

Most, but not all, of the triacylated lipid undergoes a remodeling step that converts the acyl chain to a 1-acyl-2-alkyl form. It has been established that the ether lipid donor during the remodeling step is generated by peroxisomes (Kanzawa et al. 2009), but the responsible remodeling enzyme(s) remain unknown due to the non-stringent requirement of the process and thus the lack of a strong phenotype which impedes genome-wide screening.

Next, four mannose residues extend the glycan core, with three mannoses modified by EtNPs in the order of Man1 (PIGM/PIGX) (Maeda et al. 2001), Man2 (PIGV) (Fabre et al. 2005; Kang et al. 2005), EtNP1 (PIGN) (Hong et al. 1999), Man3 (PIGB) (Takahashi et al. 1996), EtNP3 (PIGF/PIGO) (Hong et al. 2000), EtNP2 (PIGF/PIGG) (Shishioh et al. 2005), and Man4 (PIGZ) (Taron et al. 2004) by the enzymes indicated in the brackets. Dol-P-Man, a dolichol phosphate-activated mannose, is used as the donor for the mannosyltransferases, while phosphatidylethanolamine is the source for the EtNP residues. The mannosyltransferases cross the membrane 11–12 times and share structural similarity with GT-C superfamily proteins (Albuquerque-Wendt et al. 2019). In addition to the transmembrane domain, PIGB and PIGZ contain an $\alpha/\beta/\alpha$ -sandwich-soluble domain (Jumper et al. 2021), which may be necessary to bind with the Man/EtNP residues on the GPI, as similarly proposed for the lipid A glycosyltransferase ArnT (Petrout et al. 2016).

All EtNP transferases are predicted to have 16 TMHs and a large soluble domain (~45 kDa) (Jumper et al. 2021). Like the mannosyltransferases, the EtNP transferases exhibit differences in the requirement of a partner. PIGN (MCD4 in fungus) is the target for YW3548 (Hong et al. 1999), an antimycotic compound originally (Sütterlin et al. 1997) isolated from *Codinea simplex* but can also be produced by *Paecilomyces formosus* (Bilal et al. 2018). Treating *S. cerevisiae* with YW3548 or its analogs induces mutations on MCD4 that lead to drug resistance (Mann et al. 2015). Although the application of YW3548 is currently limited due to its activity toward the human enzyme, these findings demonstrate PIGN as a drug target for YW3548-based modification or de novo design with high-resolution structural information in the future.

The GPI so formed is then attached to preproteins by the GPI transamidase (GPI-T), a heteropentameric complex that is the focus of the chapter and will be discussed in detail later. A series of PGAP (post-GPI attachment to proteins) enzymes then remodel GPI-AP in both the acyl chain and the glycan part.

The inositol acyl is often hydrolyzed by PGAP1 (BST1 in yeast) (Tanaka et al. 2004), a 10-TMH protein with a luminal lipase domain. In yeast and human cells, the deacylation is required for ER-Golgi vesicle transport of folded GPI-APs (Fujita et al. 2011; Rodriguez-Gallardo et al. 2022; Tanaka et al. 2004) and degradation of

misfolded GPI-APs by the ERAD pathway (Fujita et al. 2006b; Vashist et al. 2001). Emerging evidence suggests cross talks between the remodeling by PGAP1 and the calnexin cycle to jointly ensure proper folding of N-glycosylated GPI-APs in the ER (Guo et al. 2020; Liu et al. 2018).

After the deacylation, PGAP5, a 2-TMH membrane protein that has been suggested to be a target for the anticancer cantharidin (Sahu and Tomar 2019), removes EtNP2 from GPI-APs, allowing them to be recognized by p24 proteins and assembled into COPII vesicles destined to the Golgi (Fujita et al. 2009).

In the Golgi, human GPI-APs undergo another remodeling step in the lipid part, where the polyunsaturated *sn*-2 acyl is exchanged with a long saturated chain by the sequential action of the lipase PGAP3 and acyltransferase PGAP2 (Fujita et al. 2006a; Tashima et al. 2006). In *S. cerevisiae*, the deacylation and reacylation reactions are catalyzed by PER1 and GUP1 (Bosson et al. 2006), respectively. Interestingly, GUP1 belongs to the MBOAT (membrane-bound O-acyltransferase) family and shares no sequence homology with the human PGAP2, despite their similar functions. Further, the yeast CWH43, a protein that contains a PGAP2-like N-terminal domain and a C-terminal domain homologous to the human PGAP2-interacting protein (Ghugtyal et al. 2007; Umemura et al. 2007), is responsible for the exchange of the whole diradylglycerol chain with a ceramide moiety.

Depending on specific GPI-APs, tissues, and species, Man1 undergoes further modifications with complex saccharides (Tsai et al. 2012). These modifications present yet another level of regulation on GPI-AP activity. For example, high sialylation has been suggested to correlate with the pathogenicity of prion (Bate et al. 2016) although contradicting evidence has also been reported (Katorcha et al. 2016; Stahl et al. 1992).

Discovery of Transamidase Activity

The attachment of mature GPI to client proteins involves a transamidation reaction, which replaces the peptide amide bond at the ω -site with a hybrid amide between EtNP of GPI and the peptidyl carbonyl. This reaction removes a C-terminal signal peptide (CSP) and adds GPI to the new C-terminus (Fig. 16.1b).

Biochemical characterization of the GPI attachment in the early 1990s led to evidence for the existence of a transamidase. Using a cell-free co-translation system, the Udenfriend group reconstituted the GPI attachment of a truncated placental alkaline phosphatase (termed miniPLAP) (Kodukula et al. 1991) in rough microsomes isolated from several mammalian cells (Kodukula et al. 1992). Later (Maxwell et al. 1995b), it was found that GPI-free mature miniPLAP (having its CSP cleaved but no GPI attached) was also formed, and the level of this byproduct increased from 10% to 50% when microsomes from GPI-deficient mutant cells were used, suggesting the formation of a carbonyl intermediate that is attacked by water. This postulation was supported by results from a follow-up study (Maxwell et al. 1995a) showing that the GPI-free miniPLAP can be produced more efficiently in the presence of other amine-containing chemicals such as hydrazine. Further, the

process does not require energy input. These studies led to the conclusion that an active carbonyl intermediate is formed, which is a hallmark for a transamidation reaction. This concept was supported by the discovery of a cysteine protease family subunit that bears sequence similarities to asparaginyl endopeptidase from jack bean which possesses transamidase activity (Benghezal et al. 1996).

Recently (Ness et al. 2022), the Hendrickson group reported the *in vitro* reconstitution of the transamidation reaction using a purified subcomplex with a synthetic, near-natural proprotein substrate. In this assay, a quencher-fluorophore pair is covalently linked to a full-length proprotein of CD55. Processing of the proprotein at the ω -site alleviates the quenching and increases fluorescence. Consistent with the abovementioned cell-free results, the subcomplex can cleave proprotein substrate without amine in the system, and its activity can be stimulated by hydroxylamine. Moreover, mass-spectrometry results provide definitive evidence that the transamidation reaction between hydroxylamine and the proprotein has taken place.

Discovery of GPI-T Subunits and Early Characterizations

Over the course of approximately a decade, several research groups utilized both forward and reverse genetics to identify the subunits of the GPI transamidase complex (GPI-T) using yeast and mammalian cells as model systems. As a result, GPI-T is now known to contain five subunits: GPAA1 (GAA1), PIGK (GPI8), PIGS (GPI16), PIGT (GPI17), and PIGU (GAB1) (the yeast homologs are indicated in parenthesis).

These studies were aided by several key tools and technologies. In yeast, mutant cells deficient in GPI-AP synthesis often display temperature sensitivity (Elrod-Erickson and Kaiser 1996; Leidich et al. 1995; Umemura et al. 2003, 2007; Yang and Banfield 2020), which is caused by defects in cell wall biogenesis. Screening complementing libraries in mutant cells under restrictive temperatures can identify surviving transformants that express the gene(s) responsible for the rescue. Metabolic labeling in mutant cells with ^3H inositol can reveal GPI precursors that accumulate due to defective gene(s), and these precursors can be analyzed to infer the activity of the defective enzymes. Additionally, the maturation of a GPI-AP marker (GAS1) which changes the glycosylation pattern requires GPI anchoring, making it useful for monitoring GPI-AP synthesis in wild-type and mutant cells (Pittet and Conzelmann 2007).

Unlike yeast cells, lack of GPI-AP synthesis does not noticeably affect the growth of several common mammalian cell lines in Petri dishes. Nevertheless, mutant cells defective in GPI-AP biosynthesis can be counterselected using the pore-forming bacterial toxin aerolysin, which binds to mammalian GPI-APs (Hong et al. 2003). GPI-AP biosynthesis can also be conveniently analyzed using flow cytometry with surface GPI-AP markers (Liu et al. 2021). Finally, the miniPLAP assay mentioned earlier (Maxwell et al. 1995a) has been very useful for characterizing GPI-T activity from mutant and rescued cells.

GPAA1 (GAA1, ~70 kDa) is the first GPI-T subunit to have been identified by the Riezman group (Hamburger et al. 1995) in the 1990s. In their study of a chemically mutagenized yeast strain named *end2*, they discovered that the defect in endocytosis was due to incomplete maturation of the model GPI-AP GAS1, rather than a direct effect on endocytosis. The hallmark of GAS1 maturation is the increase in its molecular weight to 105 kDa by glycosylation in the ER and to 125 kDa by further glycosylation in the Golgi. The ER to Golgi transport requires GPI anchoring (Doering and Schekman 1996), which makes the hallmark useful for screening defects in GPI-AP biosynthesis. When grown at permissive temperatures, *end2* showed impaired GAS1 maturation but unaffected GPI synthesis. Functional rescue using expression libraries identified the responsible gene as *gaal* which stands for *GPI anchor attachment*. Later, genes encoding the human (Hiroi et al. 1998) and murine homologs (Hiroi et al. 2000) were cloned, and their tissue expression patterns were characterized.

Amino acid hydrophobicity analysis suggested that GAA1 is a multi-crossing membrane protein, and its membrane topology was probed by the Menon group (Vainauskas et al. 2002) using antibodies recognizing recombinant epitope tags under selective permeabilization conditions. The results supported a 7-TMH topology with a large glycosylated lumen domain. It was later discovered (Vainauskas and Menon 2005) that the localization of GAA1 was not due to retention by ER retrieval signals but rather a passive process. Using systematic truncation and co-immunoprecipitation, the same group found that a C-terminal hydrophobic segment containing a functionally critical intramembrane proline was responsible for binding with GPI (Vainauskas and Menon 2004). Recently, two bioinformatics studies (Eisenhaber et al. 2014; Su et al. 2020) predicted structural similarity between GPAA1 and M28-type metallo-aminopeptidase and proposed that GPAA1 is the catalytic subunit that forms the amide link between phosphoethanolamine of GPI and the cleaved proprotein.

The second subunit of GPI-T is named PIGK for the human enzyme and GPI8 for the yeast homolog based on the mutant cells from which they were identified: *phosphatidylinositol glycan class K* of human K562 cells (Su et al. 2020) and *glycosylphosphatidylinositol 8* of *S. cerevisiae* cells (Benghezal et al. 1995), respectively. Taking advantage of the fact that yeast cells bearing defective GPI8 and GPI7 (the EtNP2 transferase in GPI synthesis) can grow well at 24 °C but not 37 °C, the Conzelmann group cloned the *GPI8* gene from a complementing insert that restored the temperature tolerance of the double mutant cells (Benghezal et al. 1996). They also cloned the human homolog PIGK based on functional complementation. Around the same time, Yu and colleagues identified the *PIGK* gene that, when over-expressed, restored surface staining of GPI-AP markers in the class K mutant cells (Yu et al. 1997). For reverse genetics, compromising GPI8 function in *T. brucei* by gene disruption or RNA interference resulted in growth defects for the bloodstream form and inability to infect insect cells for the procyclic form (Lillico et al. 2003). Interestingly, the disruption of the *GPI8* gene in *Leishmania mexicana* abolished GPI-APs, but the mutant parasite remained viable and infective (Hilley et al. 2000).

The human and yeast PIGK proteins contain a luminal domain of approximately 33 kDa that shares sequence similarity with cysteine proteases, as well as a single transmembrane domain at the C-terminus which is absent in some homologs, such

as the *T. brucei* GPI8 (Sharma et al. 2000). The similarity between PIGK and cysteine protease was supported by biochemical evidence from several research studies around the same time (Meyer et al. 2000; Ohishi et al. 2000). Thus, mutating Cys199/Cys206 or His157/His164 of the yeast/human GPI8/PIGK to alanine, or adding the thio-reactive iodoacetamide to purified *Leishmania* GPI8, abolished GPI-AP anchoring in cell-based or in cell-free assays. Using photocrosslinking, the Menon group found that GPI8 is linked to proprotein substrate (Vidugiriene et al. 2001), providing further evidence for its role in catalysis and suggesting physical proximity between proprotein substrates and GPI8 during the transamidation reaction. In addition, their results suggested possible additional subunits. Similar conclusions were drawn from another cross-linking study (Spurway et al. 2001) that compares the PIGK-proprotein interaction in various systems that are defective in different stages of GPI-AP synthesis. PIGK is cross-linked with proprotein substrates in the wild-type K562 cells but not the class K mutant cells. The cross-linking is promoted in cells that are either deficient in GPI synthesis or when the proprotein bears a mutated ω -site, probably owing to their prolonged interaction with PIGK. These findings indicate that PIGK is a catalytic subunit.

The identities of other subunits were soon to emerge. In 2001, the Conzelmann group identified GPI16 (PIGT) (Fraering et al. 2001). They overexpressed GST (glutathione S-transferase)-tagged GPI8 in yeast *gpi8-1* cells and purified GPI-T in digitonin, a mild detergent that usually preserves the integrity of membrane protein complexes. In blue-native gel electrophoresis, the complex has an apparent molecular weight of 430–650 kDa. Mass-spectrometry analysis of the trypsin-digested complex identified fragments of GPI16. A similar approach led to the discovery of human PIGT by the Kinoshita group (Ohishi et al. 2001), which additionally identified PIGS, a 75-kDa protein with two transmembrane helices, as the fourth subunit for the human GPI-T. Functional complementation experiments established GPI16 and GPI17 as the yeast homolog of PIGT and PIGS, respectively. Notably, the different output from the two studies was later attributed to the fact that the yeast GPI17 (PIGS) does not stably associate with the other subunits during affinity pull-down (Zhu et al. 2005). By sequence homology analysis, PIGT from *T. brucei* was also identified as an essential subunit of GPI-T (Hong et al. 2006).

The fifth subunit was discovered by the Kinoshita group (Hong et al. 2003) using a mutant CHO strain that is resistant to aerolysin. The mutant cells showed a similar GPI accumulation profile to GPI-T-deficient cells. Microsomes isolated from the mutant cells did not generate the carbonyl intermediate between proprotein and GPI8 in the cell-free assay. In addition, overexpression of the other four subunits did not restore these phenotypes. These results led to the conclusion that there is a fifth subunit of the GPI-T complex which they identified from complementing libraries and named it as PIGU. Interestingly, this subunit was co-purified during their earlier study (Ohishi et al. 2001) that led to the discovery of PIGT and PIGS. However, owing to the high hydrophobicity, PIGU immigrated as smeared bands in SDS-PAGE and was thus overlooked. Adding urea to the SDS-PAGE sharpened the band and allowed confirmation of PIGU by mass-spectrometry. Like other subunits, the yeast homolog can partially rescue GPI anchoring in *PIGU* mutant cells. Finally, using the same GST-affinity purification approach, TTA1 and TTA2 were identified as the

functional homolog of PIGS and PIGU in *T. brucei*, although they do not share appreciable sequence homology with the yeast or the human counterparts (Nagamune et al. 2003).

The assembly of the GPI-T complex has been probed by the dependence of the stability of subunits on each other. PIG-T is believed to play an essential role for complex assembly, as it dictates the stable expression of GAA1 and GPI8 (Ohishi et al. 2001). Further, both the human (Ohishi et al. 2003) and the trypanosoma PIGK (GPI8) (Hong et al. 2006) form an inter-subunit disulfide with their respective PIGT (GPI16). In addition, the soluble domains of GPI8 (PIGK) and GAA1 (Gamage et al. 2017), or together with the soluble domain of GPI17 (PIGT) (Ness et al. 2022), assemble into a subcomplex, which can catalyze transamidation reactions in vitro.

While PIGS is also essential, it is expected to play a less important role in the complex assembly, as the three-membered subcomplex GPAA1/PIGK/PIGT could still form without PIGS in the mammalian cells (Ohishi et al. 2001). This is consistent with the results from yeast GPI-T studies, which show that GPI17 dissociates from GPI8 during affinity purification (Zhu et al. 2005). A more in-depth stability study (Kawaguchi et al. 2019) of the mammalian GPI-T subunits revealed that PIGT, PIGU, and GPAA1 mutually stabilize each other. The stability of PIGS depends on these three subunits, whereas the stability of PIGK depends on all other subunits. These results suggest that a stable subcomplex formed by PIGT/PIGU/GPAA1 is required to assemble PIGS and PIGK into a functional transamidase.

GPI-T has been proposed to exist as a homodimer of the heteropentamer based on several lines of evidence. On blue-native gel, the yeast complex has an apparent molecular weight of 430–650 kDa, which is 2–3 times of its theoretical molecular weight of ~290 kDa (Fraering et al. 2001). This result was recently confirmed by the Wu group (Yi et al. 2017), which reports a band between 480 kDa and 720 kDa. In addition, the in vitro-assembled complex composed of the soluble domains of the yeast GAA1 and GPI8 forms an $\alpha 2\beta 2$ -stoichiometry, as shown by blue-native electrophoresis gel and size exclusion chromatography results (Gamage et al. 2017). Finally, the yeast GPI8 forms a disulfide-linked dimer when expressed and purified from insect cells (Yi et al. 2017).

Characteristics of GPI-T's Proprotein Substrates

A typical protein destined for GPI anchoring has an N-terminal signal peptide that directs it to the ER (Fig. 16.1b). This form, referred to as the preproprotein, is cleaved by signal peptidase to form proprotein. GPI-T then cleaves the proproteins at the ω -site and transfer GPI en bloc to the new C-terminus (Kinoshita 2020).

One of the remarkable features of GPI-T is its ability to recognize proproteins based on a pattern in the CSP region, rather than a consensus sequence (Kinoshita 2020) (Fig. 16.1b). The pattern consists of a small ω -residue to which GPI is latter attached, an $\omega+1$ residue that can be any amino acid except proline (Gerber et al. 1992), a small $\omega+2$ residue (Kodukula et al. 1993), a hydrophilic linker with 8–12 residues, and a hydrophobic segment with 15–20 hydrophobic residues. The

ω -residue can be glycine, alanine, serine, asparagine, aspartate, or, more recently, threonine, leucine, and methionine, albeit at lower frequencies (Masuishi et al. 2016). Some GPI-APs have multiple ω -residues. Interestingly, the hydrophobicity of the CSP must be within a narrow range for efficient anchoring (Galian et al. 2012); it must be higher than the signal peptide for signal peptidases but lower than most transmembrane helices. In addition to CSP, a flexible linker of approximately ten residues proceeds to the ω -site in GPI-APs. These features have enabled the development of various algorithms (Gíslason et al. 2021; Pierleoni et al. 2008; Sauer et al. 2023) to predict GPI-APs and their ω -sites. Notably, the pattern, although appearing to be vague, exhibits species specificity, as swapping CSP between species can impair or abolish GPI anchoring (Morissette et al. 2012). This difference provides a basis for developing species-specific inhibitors to block GPI anchoring in eukaryotic pathogens such as fungi and protozoan parasites.

The Promiscuity of GPI-T

GPI-T is a promiscuous enzyme. It can process a diverse range of proprotein substrates containing the CSP motif with a vague pattern, and even cleave a tetrapeptide substrate which lacks a signal peptide, as demonstrated by the GPI8 from *T. brucei* (Kang et al. 2002). In addition, besides its native substrate GPI, several amine-containing chemicals, such as the abovementioned hydrazine and hydroxylamine, can mediate the transamidation reaction when assayed in vitro (Maxwell et al. 1995a; Ness et al. 2022). Further, GPI-T can cleave proproteins without GPI attachment, thus essentially acting as a protease (using water as a nucleophile) (Kang et al. 2002; Maxwell et al. 1995b). Within GPI, EtNP3 has been long been believed to be the linking group. However, a recent study found EtNP2 to be a preferred linkage for some GPI-APs such as 5'-nucleotidase Ecto and NetrinG2 (Ishida et al. 2022).

Given the broad specificity of GPI-T, its protease activity must be safeguarded to avoid potential harm caused by accidental cleavage. Proteases often rely on strict sequence specificity (such as thrombin) (Chang 1985), acidity (such as lysosomal proteases) (Reinheckel and Tholen 2022), inhibitors (Bobek and Levine 1992), or tightly regulated activation of proenzymes (such as caspases) (Kesavardhana et al. 2020) to control activity. However, GPI-T does not appear to have such characteristics. Its safeguard mechanisms remain to be elucidated.

Structural Insights into GPI-T Assembly and Function

Overall Structure

Recently, two papers describing the experimental structure of GPI-T at 2.53 Å (Xu et al. 2022) and 3.4 Å resolution (Zhang et al. 2022) have been published. The two structures are overall similar, and the 2.53 Å structure is mostly discussed in this chapter.

GPI-T is composed of a heteropentamer with 1:1 subunit stoichiometry with a dimension of approximately $140 \times 150 \times 80$ Å. Its large luminal domain resembles the shape of a canon that sits on the carriage-shaped transmembrane domain (Fig. 16.2a). Except for a few transmembrane helices (TMHs) that protrude out of the membrane, no significant cytoplasmic masses are present. The transmembrane domain contains 24 TMHs, including 8 from GPAA1, 12 from PIGU, 2 from PIGS, and 1 each from PIGK and PIGT (Fig. 16.2b). Except for PIGU, the other four subunits all contribute to the luminal domain. The structure reveals three glycosylation sites, including one each from GPAA1, PIGS, and PIGT. In addition, it captures a calcium ion in the subunit PIGK (Xu et al. 2022).

Subunit Assembly

The assembly mechanisms previously predicted based on stability data (Kawaguchi et al. 2019; Ohishi et al. 2001) are remarkably consistent with those revealed by the cryo-EM structures. As mentioned above, the stability data suggest a central role of PIGT; in the structure, the complex is indeed assembled around PIGT and PIGU (Figs. 16.2a and 16.3a). Similarly, the stability data suggest GPAA1/PIGT/PIGU forms a subcomplex; in the structure, their interactions occur in a comparably larger buried area than with other subunits. Biochemical and stability data suggest a less important role of PIGS for complex assembly; in the structure, PIGS is located at the periphery of the complex. Finally, the stability of PIGK requires the presence of all other subunits; in the structure, GPAA1 and PIGS approach the PIGU and PIGT, creating a large, three-sided cavity in the luminal domain to host PIGK (Fig. 16.2a).

Central to the complex assembly is the PIGT-PIGU platform (Fig. 16.3a) which is held together via several notable features. The 12 TMHs of PIGU arrange into 2 rings. The outer ring consists of traversing TMHs that shield the inner ring which is made of short TMHs that make sharp turns inside the membrane. This creates a void space to confine five amphipathic helices (AHs) with their hydrophobic halves anchoring in the membrane and hydrophilic halves facing the lumen. This structural arrangement suits its function, because AHs are more efficient than TMHs in creating surfaces parallel to the membrane plane for interactions with luminal domains by orienting side chains to the membrane interface. In line with this principle, PIGU interacts with the TMH of PIGT via two long TMHs (TMH 8/12) which would contain more sidechain surfaces than short TMHs for interactions within the membrane plane. Therefore, PIGU is geometrically optimized to interact with PIGT.

Within the central PIGT/PIGU platform, PIGT engages with the luminal domains while PIGU attracts the TMHs of the other subunits. The luminal domain of PIGT contains a curved ten-stranded β -sheet that resembles a rib cage (Fig. 16.3b). Although not directly involved in subunit interaction, this structure acts like a skeleton to support other elements that are involved in subunit binding. The concave side of the rib cage is filled with loops that interact with the linker region in PIGK,

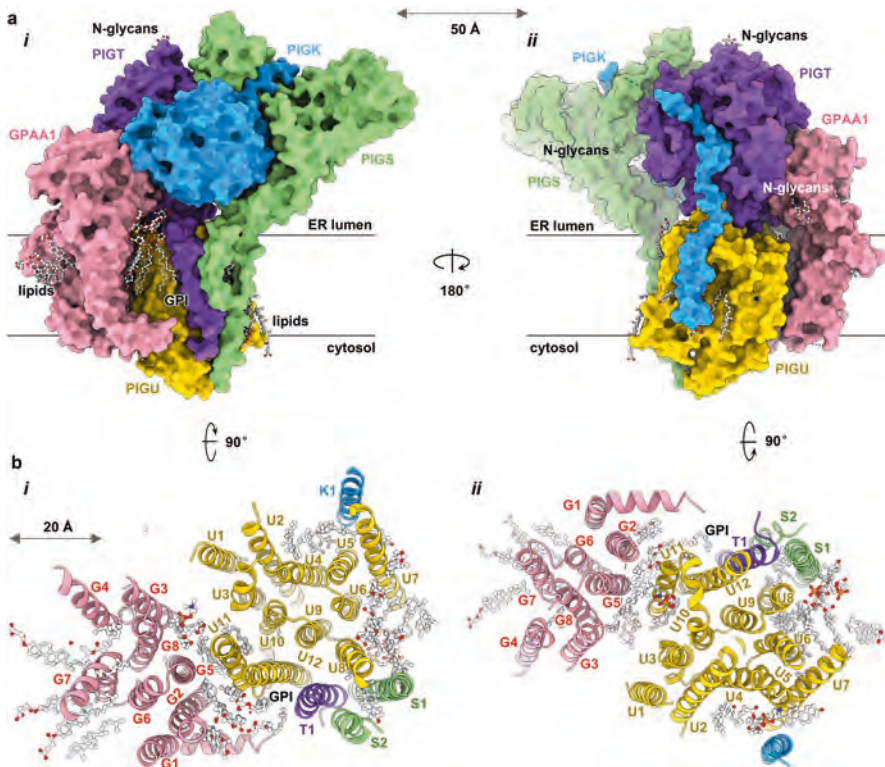


Fig. 16.2 Structure of GPI-T. (a) Side view of the GPI-T complex in surface representation. Lipids, including its substrate GPI, detergents, and glycans are shown as stick-ball representation. Membrane boundaries are marked by lines. *ER* endoplasmic reticulum. (b) Normal views of the transmembrane domain from ER lumen (left) and cytosol (right). Transmembrane helices (TMHs) are labeled with numbers in addition to letters G/K/S/T/U for GPAA1/PIGK/PIGS/PIGT/PIGU. Lipids and detergents are shown as ball and stick representations

and a nine-stranded β -sandwich domain that binds with GPAA1 and PIGT. The convex side is decorated with short helices and β -strands that bind with PIGS and PIGK. At the membrane domain, PIGU interacts with all other four subunits through direct or lipid-mediated hydrophobic interactions (Fig. 16.2b). Thus, GPI-T is held together by a rich network of interactions in both the luminal and transmembrane domains.

Two characteristics of the assembly mechanism for the catalytic subunit PIGK are worth noting. First, PIGK uses four loops to mediate interactions with the other subunits in the three-sided luminal cavity (Fig. 16.3c). Because loops are more flexible, they may act as a buffer to protect the main protease domain from environmental conformational changes, creating a relatively stable environment for the catalytic domain. Second, PIGK is secured to the rest of the complex through multivalent interactions, with the luminal domain embraced in the cavity (Fig. 16.2a), the sole TMH of PIGK planted in the membrane domain and associated with PIGU

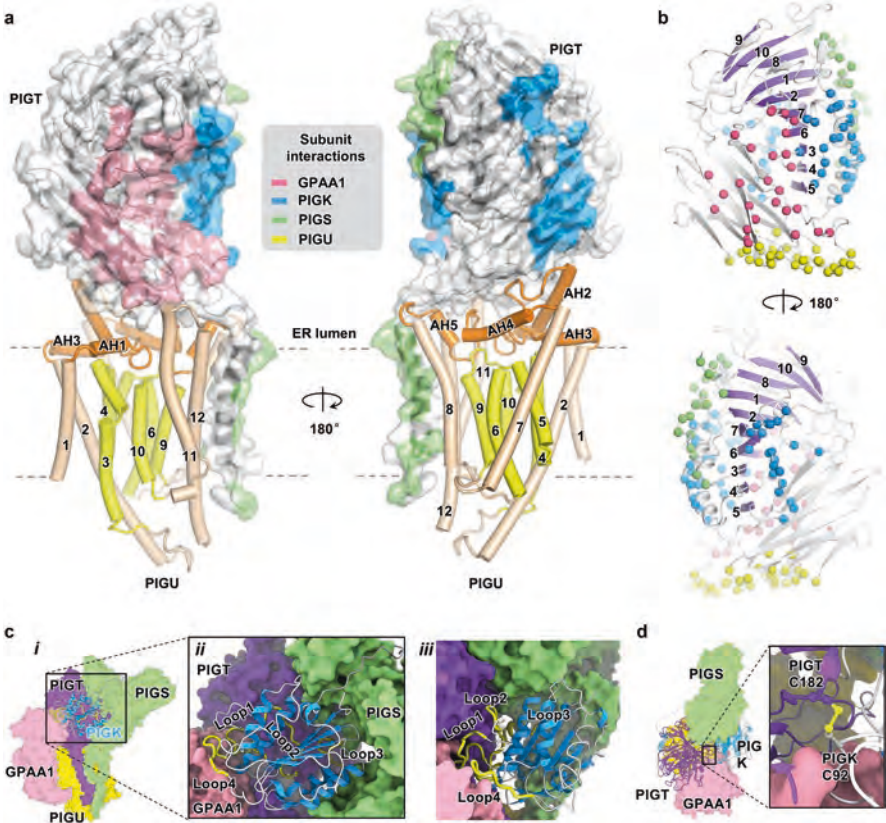


Fig. 16.3 Notable features of GPI-T assembly. (a) The PIGT-PIGU platform. PIGT is shown as cartoon and semitransparent surface representations. Different colors in PIGT indicate regions that interact with other subunits as explained in the gray box. PIGU is shown as cartoon representation with helices as cylinders. The relatively long and short TMHs and the amphipathic helices (AHs) are colored differently. Numbers indicated TMHs 1–10. (b) The luminal domain of PIGT. The central β -sheet is colored purple with numbers indicating the ten strands. Residues that interact with other subunits are shown as $C\alpha$ -spheres with colors explained in the gray box. (c) The four loops that nest the main protease domain of PIGK in the three-sided cavity. The loops are labeled as thick loop (yellow). (i), the overview. (ii) and (iii), expanded view at two different angles. Loop1, residues 89–101; loop2, residues 112–125; loop3, residues 137–151; loop4, residues 237–243. (d) Overview and expanded view of GPI-T and the PIGT-PIGK disulfide viewed from ER lumen

(Fig. 16.2b), and the connecting loop hanging on grooves in PIGS and PIGT (Fig. 16.2a). In addition to non-covalent interactions, Cys92 from PIGK further nails the luminal domain to PIGT by forming an inter-subunit disulfide bond with Cys182 (Fig. 16.3d). Although this disulfide bond is not found in some species such as *S. cerevisiae* (Yi et al. 2017), it is crucial for GPI-T assembly, as mutations that eliminate the cross-linking reduce activity of both the human (Ohishi et al. 2003; Xu et al. 2022) and the *T. brucei* GPI-T complex (Hong et al. 2006).

Active Site Architecture of GPI-T Likely Confers Substrate Selectivity

Capturing substrates or products within the active site of an enzyme is crucial for understanding its substrate recognition and catalytic mechanisms. This was anticipated to be challenging for GPI-T because chemical synthesis of GPI with lipid chains is a difficult task. However, an endogenous GPI was co-purified and trapped in both cryo-EM structures (Xu et al. 2022; Zhang et al. 2022), which revealed an elongated active site extending from the middle of the membrane to the catalytic site of PIGK, measuring over 50 Å (Fig. 16.4a).

The GPI-binding site is of a shared type. At the membrane domain, the lipid chains insert in a cavity between GPAA1 and PIGU/PIGT/PIGS. At the juxtamembrane region, a depression structure formed by PIGT, GPAA1, and PIGK holds the glycan core which has been visualized to the first mannose (Fig. 16.4b). The lipid part mostly interacts with the TMH11/TMH12 of PIGU, while the glycan core binds with residues from GPAA1 and PIGT (Fig. 16.4a). Interestingly, Asp521 and Ser523 from PIGT are more sensitive to mutagenesis than residues from other subunits (Xu et al. 2022), suggesting a critical role of PIGT in GPI binding. This was somewhat unexpected because PIGT has been long believed to be a scaffolding subunit (Kawaguchi et al. 2019; Ohishi et al. 2001). In addition, although GPAA1 and PIGU are involved in GPI binding, the responsible regions differ from those postulated from previous characterizations (GPAA1 TMH8 and PIGU TMH9) (Hong et al. 2003; Vainauskas and Menon 2004) (Fig. 16.1b, c).

Although how GPI-T recognizes its proprotein substrate remains to be investigated, the overall positioning and orientation of proprotein can now be speculated.

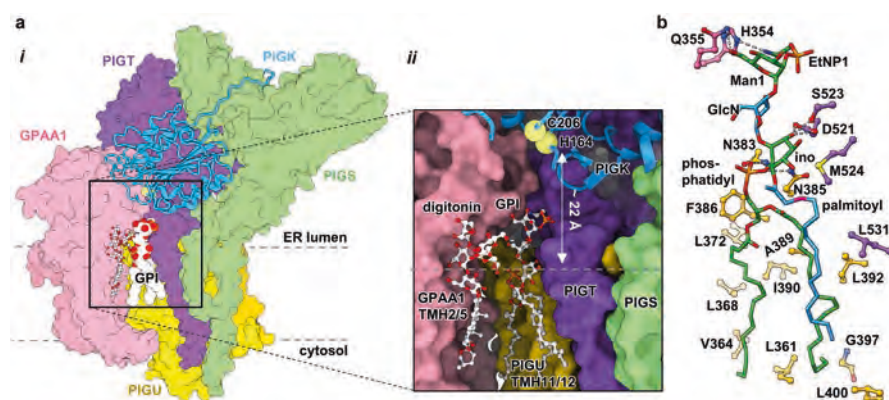


Fig. 16.4 The composite GPI-binding site. (a) The GPI-binding site near the catalytic dyad (yellow sphere). In the overview (i), GPI is shown as sphere representation. In the expanded view (ii), GPI is shown as ball and stick representation alongside a digitonin molecule. Subunits are shown as surface representations except for PIGK which is in cartoon representation. (b) Detailed interactions between GPI (alternatingly colored with green and blue) and GPI-T residues (colored according to subunit coloring in a). Dash lines indicate hydrogen bonding

The assembly mechanism fixes the catalytic subunit PIGK in a “mid-air” position in relation to the ER membrane, leaving an approximate 22 Å space between the catalytic dyad and the membrane interface. For orientation, the ω -residue must be near the catalytic dyad, and the hydrophobic region in the CSP is expected to insert into the membrane, probably at the same cavity that holds GPI. This arrangement would place the 8–12 amino acid spacer (Fig. 16.1b) in the 22 Å space. Thus, GPI-T may use the topological matching between the active site architecture and the hydrophobic/hydrophilic pattern to confer substrate specificity (Fig. 16.4a). This hypothesis remains to be investigated by future structural studies.

Structure Feature of Individual Subunits

The Catalytic Subunit PIGK

The PIGK structure is globular and contains a mostly parallel, six-stranded β -sheet sandwiched by connecting α -helices. This fold is shared by other Cys proteases such as caspases (Earnshaw et al. 1999) and legumains (Dall and Brandstetter 2016) despite their low sequence similarity (Xu et al. 2022). In addition, the catalytic dyad (Cys206-His168) and oxyanion-hole residues (Gly165, Gly166) are positioned similarly in PIGK and legumains (Fig. 16.5a). Further, the S1 pocket which, by analogy to caspases, would hold the ω -site during the cleavage, is composed of Arg44, His45, Asp231, and Glu187 (Dall and Brandstetter 2016). (Based on the Schechter-Berge nomenclature (Schechter and Berger 1967), the residue with the scissile bond is named P1, and residues N- and C-terminal to P1 are respectively named as P2, P3, and P1', P2', etc. Accordingly, pockets that hold such residues are named as S1/S1', etc.) PIGK contains three identical residues (Arg60, His61, Asp247) and one conserved residue (Asp204) in the structurally equivalent positions (Fig. 16.5b). Mutating these residues impairs activity (Xu et al. 2022), suggesting that PIGK uses a similar mechanism to host and cleave proprotein substrates.

The zwitterionic nature of the legumain S1 pocket has been postulated to confer strict substrate selectivity with an asparagine at the P1 position (Dall and Brandstetter 2013). Thus, aspartate, although having similar volume as asparagine, cannot serve as a P1 residue under neutral conditions because of charge repulsion with the two acidic residues in the S1 pocket (Fig. 16.5a). Under low pH conditions, however, legumains can cleave an aspartate-P1 substrate because the repulsive interactions are reduced by protonation of both the P1 and S1 residues. Although having a similar S1 pocket composition with legumains, PIGK readily uses aspartate as an ω -residue (P1) despite being located in the ER lumen, which has a neutral pH. In addition, caspases, which cleave substrates with strict specificity with an aspartate at P1, use a basic S1 pocket that is made of a glutamine and two arginine residues (Earnshaw et al. 1999). This raises the possibility that PIGK uses a different composition or organization of the S1 pocket to accommodate an aspartate ω -residue.

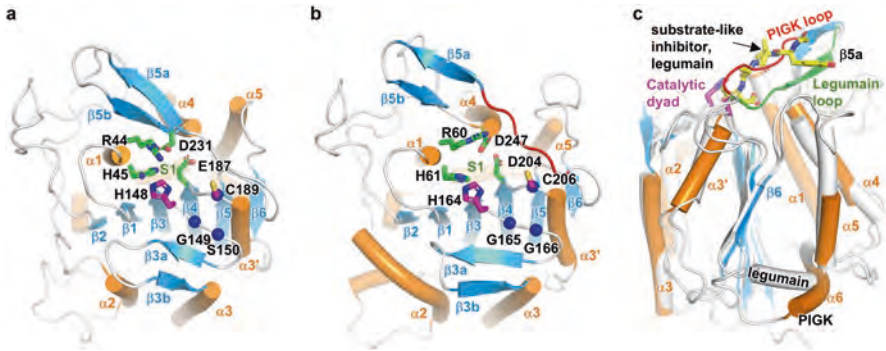


Fig. 16.5 Structure comparison between PIGK and legumain. **(a)** Structure of the catalytic domain of legumain (PDB ID 4AWA). **(b)** Structure of the luminal domain of PIGK. **(c)** Alignment of the structures of legumain and PIGK. Protein structures are shown as cartoon representations color-coded by secondary structure elements (α -helix, orange; β -sheet, blue; loop, white). The S1 pocket residues (green) and catalytic dyad (magenta) are shown as sticks, and the oxyanion hole residues are shown as C α -sphere (blue). In **(b)** and **(c)**, the loop where PIGK differs from legumains is colored red. A covalent inhibitor (yellow stick) of legumain is shown in **(c)** to illustrate its clashes with the PIGK loop (red) but not the legumain loop (green)

Although PIGK is overall similar to legumains, a loop near the catalytic dyad displays noticeable differences that may have functional significance. Specifically, the loop that makes part of the subsite is “elevated” in PIGK, preventing the entrance of the substrate superposed from legumains (Fig. 16.5c). It is possible that this arrangement still works for PIGK since the substrate selectivity mostly relies on the residues C-terminal to the ω -site (the clashing occurs at the N-terminal part of the peptidyl inhibitor) or that PIGK in the cryo-EM structures may exist in an inactive state. These hypotheses remain to be tested by structure determination of the enzyme-substrate complex.

GPAA1 Contains a Protease-Like Domain and a Lipid-Binding Pocket

The shape of GPAA1 resembles that of a portico, with roughly equal mass in its soluble and the transmembrane regions (Fig. 16.6a). The soluble domain sits directly “atop” of the transmembrane domain, suggesting that GPAA1 may be functional alone with possible moonlight activities.

Previous bioinformatics studies (Eisenhaber et al. 2014; Su et al. 2020) have proposed GPAA1 as a catalytic subunit. In their model, PIGK cleaves proprotein substrates at the ω -site, and GPAA1 attaches GPI to the carboxyl of the ω -residue through the ethanolamine moiety. This proposal, mostly based on the fact that the soluble domain has a similar fold to M28, a type of zinc protease (Eisenhaber et al. 2014), suggests that GPAA1 catalyzes the reverse reaction of a protease, i.e., the GPI attachment step that generates, rather than breaks, an amide bond.

Agreeing with the bioinformatics predictions (Eisenhaber et al. 2014), GPAA1 indeed assumes a Zn protease fold, characterizing a central eight-stranded β -sheet sandwiched by α -helices (Fig. 16.6b, c). However, detailed analysis of side chains, the positioning of which are generally not reliable in predicted models, indicates GPAA1 does not have protease/ligase functions. Specifically, the metal-binding motif in the Zn protease that contains two histidine residues and four acidic side-chains is not present in GPAA1 (Fig. 16.6c). Mutagenesis experiments also show that altering nearby histidine and acidic residues (Fig. 16.6c) does not affect GPI-T activity, as evidenced by wild-type level of surface staining of GPI-AP markers (Xu et al. 2022).

In contrast to a previously reported 7-TMH model (Vainauskas et al. 2002), the membrane domain of GPAA1 contains eight TMHs and four amphipathic helices (AHs). The eight TMHs are arranged into two helix bundles (TMH 1/2/5/6 and TMH 3/4/7/8) (Fig. 16.6a), which separate from each other at the lumen leaflet to create a funnel-shaped pocket that is enclosed by AH2 and AH4. Notably, TMH 7 breaks into roughly two equal parts (TMH 7a and 7b) at Pro553, which facilitates the formation of the funnel by tilting TMH 7b away from TMH 8 and sealing the funnel with TMH 7a in the middle of the membrane (Fig. 16.6a). The higher-resolution structure (Xu et al. 2022) reveals a phospholipid (modeled as a phosphatidylethanolamine) in this pocket, which appears to have high affinity for membrane lipids as it attracts four ordered cholesterol hemisuccinate molecules at its periphery (Fig. 16.6a).

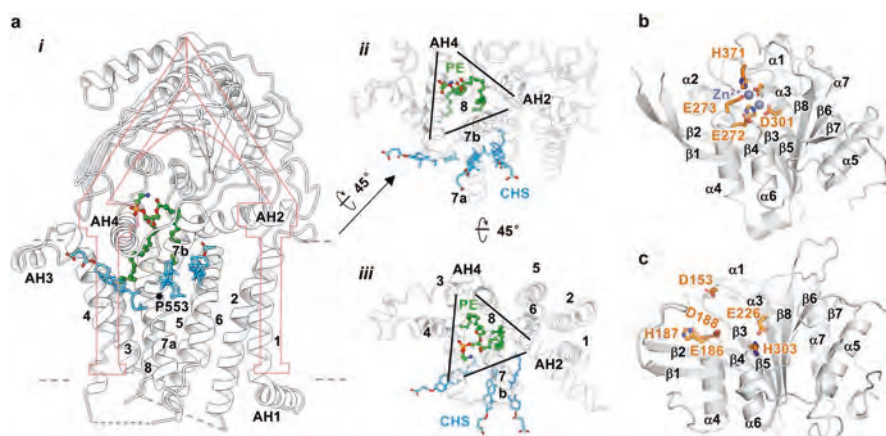


Fig. 16.6 Structure of GPAA1. (a) The overall structure of GPAA1. Red sketches indicate the resemblance between the GPAA1 structure and a portico (i). Transmembrane and amphipathic helices (AHs) are labeled. A phosphatidylethanolamine (PE, green) and three cholesterol hemisuccinate molecules (CHS, blue) are shown as stick representation. The lipid pocket and the arrangement of the transmembrane helices are shown at two different angles in (ii) and (iii). (b) The structure of a Zn protease (PDB ID 2EK8). (c) The structure of the luminal domain of GPAA1. Residues and the Zn²⁺ ions are labeled appropriately

Although the functional significance of the lipid-hosting pocket is currently unclear, biochemical and structural data may provide some clues for future investigations. Truncation analysis has identified a region covering TMH8 and AH4 (residues 564–621) as important for GPI binding (Vainauskas and Menon 2004). Although the cryo-EM structure models a phospholipid with two acyl chains (Fig. 16.6a), the pocket may be capable of hosting a GPI with three acyl chains with some flexibility from TMH 7b, thus acting as a reservoir for free GPI or a station to host newly modified GPI-AP before channeling to PGAP enzyme(s) for remodeling.

Previously (Hamburger et al. 1995), it was found that the overexpression of the yeast GAA1 suppresses the defect in the maturation of GAS1 mutants with alterations at or near the ω -site. Because the maturation process depends on GPI anchoring, this result may be interpreted as GPAA1 being involved in substrate binding, possibly through its protease-like domain. In addition, as reasoned above, GAA1 can probably exist as a standalone protein. The excess amount of non-complexed GAA1 and hence the lipid-binding pocket may reserve free GPI which is scarce in yeast membranes. This function may compensate the lower processivity of GAS1 mutants by providing more GPI substrates. Alternatively, free GAA1 may facilitate maturation of GAS1 by forming weak complexes with downstream machineries responsible for glycosylation quality control or vesicle transport of GPI-APs. Intriguingly, GPI-T was co-purified with a RING finger protein (RNF121) that is likely an ubiquitin E3 ligase (Zhang et al. 2022). Although the resolution of the cryo-EM map was not high enough to reveal atomic details (Zhang et al. 2022), the extra density corresponding to RNF121 was associated with GPAA1, implying that GPAA1 participates in the endoplasmic reticulum-associated degradation.

Structural Feature of Individual Subunits: PIGU Adopts a GT-C Fold

PIGU is the most hydrophobic subunit of GPI-T, with the majority of its residues buried in the membrane. Its overall structure bears a striking resemblance to that of the glycosyltransferases of the C (GT-C) superfamily (Bohl et al. 2021), which use polyisoprenol-linked donors for glycosylation. The GT-C member glucosyltransferase Alg6 (Bloch et al. 2020) is used here for illustration. The 12 TMHs of PIGU align to TMH 1–10 and TMH 13–14 of Alg6, while the 5 AHs are also aligned well (Fig. 16.7). However, AH4 from PIGU is longer than that from Alg6 (Fig. 16.7). AH4 sits in a membrane-facing cavity formed by two tilting V-shaped TMHs (TMH 7/8) on either side and the short TMH 6 at the bottom. The cavity in PIGU is filled with three CHS molecules, whereas in Alg6, it is occupied by the dolichol phosphate substrate (Fig. 16.7). When aligned, AH4 obstructs the hydrophilic head group of the superimposed substrate from accessing the active site (Fig. 16.7). This suggests that PIGU may have evolved from the same ancestor as GT-C enzymes but lost the catalytic elements.

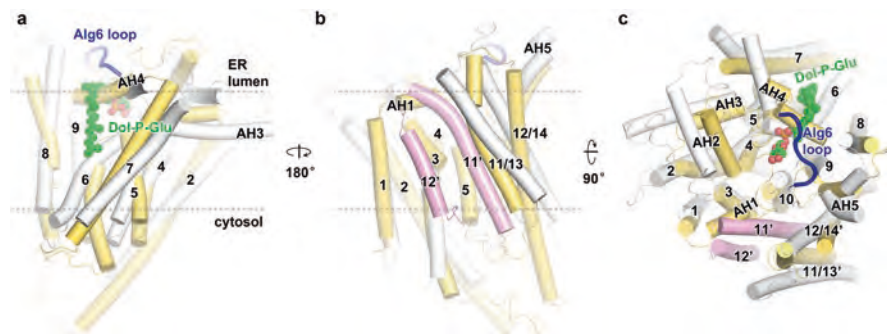


Fig. 16.7 Structure comparison between PIGU and Alg6. Superposed structures of PIGU (orange) and Alg6 (gray and pink) are shown at three angles (*i–iii*). Glucosyl-dolichol-phosphate (Dol-P-Glu), a lipid substrate of Alg6, is shown as small-sphere representation (green). The substrate-gating loop in Alg6 is colored blue. Transmembrane helices (TMHs) 1–10 which are common to PIGU and Alg6 are numbered with a single number. The extra two TMHs of Alg6 (TMH11 and 12) are numbered with a prime. TMH11 and 12 of PIGU align with TMH13 and 14 of Alg6. AH amphipathic helix

Structural Feature of Individual Subunits: PIGS and PIGT

PIGS contains two TMHs that anchor it to the ER membrane, and a large shamrock-shaped luminal domain. This shape is well suited for binding with PIGK. Thus, one groove embraces the protease domain of PIGK with shape complementarity, while the other groove accommodates the linker that connects the protease domain and the TMH of PIGK (Fig. 16.2a). The structural features of PIGT are described earlier (Fig. 16.3b).

Structural Rationalization of Disease Mutations

Reflecting GPI-T's essential role in GPI-AP biosynthesis, genetic mutations that cause reduced surface staining of GPI-APs and neurodevelopmental and congenital disorders have been reported for all subunits. Specifically, 14 point mutations have been identified in GPAA1 (M50K, S51L, S54N, M55T, W176S, L190P, L191P, H306R, A316V, N323S, L350R, A389P, L520R, W611R) (Castle et al. 2021; Li et al. 2020; Nguyen et al. 2017), 7 in PIGK (S53F, L86P, D88N, Y106S, M246K, C275R) (Nguyen et al. 2020), 2 in PIGS (L34P, E308G) (Nguyen et al. 2018), 10 in PIGT (T183P, E184K, E237Q, G360V, G366R/W, R448W, R507Q/W, V528M) (Ben Ayed et al. 2023; Jezela-Stanek et al. 2020b; Kohashi et al. 2018; Kvarnung et al. 2013; Nakashima et al. 2014; Pagnamenta et al. 2017; Skauli et al. 2016; Yang

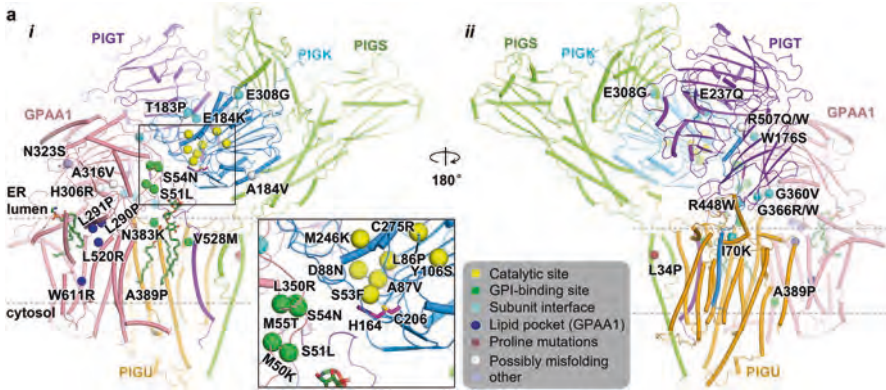


Fig. 16.8 Structural interpretation of disease mutations associated with GPI-T. Point mutations reported to date are mapped to the GPI-T structure viewed at two different angles (*i*, *ii*). The C α of the mutated residues are color-coded as explained in the dark-gray box. Membrane boundaries are marked with dashed lines. GPI and the phosphatidylethanolamine from the GPAA1 lipid-binding pocket are shown as stick representations (green)

et al. 2018), and 2 in PIGU (I70K, N383K) (Knaus et al. 2019). The GPI-T structures provide a framework to map and understand the mechanism of the disease mutations. Out of the 35 identified mutations, 15 are concentrated near the catalytic dyad and the GPI-binding site (Fig. 16.8). Such mutation may impair GPI-T activity by distorting the active site or weakening substrate binding. Nine of the remaining 20 mutations occur at the inter-subunit interface, including PIGT T183P/E184K/G360V/G366R/G366W/R488W, PIGS E308G, GPAA1 W176S, and PIGU I70K. These mutations are expected to affect GPI-T assembly, and therefore the stability and activity of the complex. Four proline mutations, including GPAA1 A389P/L290P/L291P and PIGS L34P, are expected to distort the helix structure of TMH X, AH2, and TMH 1, respectively. Such mainchain distortion may lead to misfolding and degradation of the affected subunits or instability of the complex. Of note, L290P, L291P, together with L520R and W611R, occur near the lipid-binding pocket of GPAA1. Finally, some of the mutations, such as PIGT W176S and PIGK Y106S, occur in the interior of the globular subdomains. These mutations may promote misfolding by clashing or creating a mismatching microenvironment with surrounding residues (Xu et al. 2022).

Conclusions and Outlook

Enzymes, transporters, and regulators of the GPI-AP biosynthesis pathway are of great physiological and pharmaceutical importance. Since the discovery of the first subunit in the 1990s, extensive cell biology, biochemistry, and bioinformatics studies of GPI-T have provided rich information on its substrate specificity, subunit characteristics, and catalytic mechanism. The recent GPI-T structures provide

accurate assembly mechanism of the complex, clarify the catalytic subunit, identify the GPI-binding site, and rationalize the disease mutations.

Biochemical and structural studies of GPI-AP biosynthesis enzymes have been challenging; in fact, GPI-T is the first enzyme of the pathway to have the structure experimentally determined. This is largely due to the facts that almost all (~20) enzymes in the pathway are transmembrane proteins, GPI lipids are not commercially available, and the chemical synthesis of GPI is extremely challenging owing to the complicated glycan core and the lipid chains. Fortunately, an endogenous GPI molecule was co-purified with GPI-T in the cryo-EM samples, which provides promising possibility for future biochemical and structural studies. Thus, engineered cells capable of producing desired GPI intermediates may offer an attractive solution for obtaining substrates and products of enzymes of the pathway.

Despite the new insights from GPI-T structures, several important questions remain unanswered. For example, the structural basis for proprotein recognition, the catalytic mechanism, and how GPI-T achieves both broad substrate specificity and catalytic fidelity are not fully understood. Further, the integration of seemingly weak CSP consensus features into a strong selectivity filter remains an intriguing topic in structural biology. In addition, while the structures explain many previous discoveries, some aspects still require further investigation. For example, an intra-membrane proline (Pro609) of the GPAA1 TMH 8 has been reported to be functionally critical for GPI-T activity (Vainauskas and Menon 2004). Proline is known to disrupt helices, but it does not seem to be the case of the TMH8 which is a continuous helix in the structures (Fig. 16.6a). Other possibilities may be related to its critical role in forming the lipid-binding pocket, although the functional significance of the pocket is yet to be identified. Intriguingly, two disease mutations of GPAA1 occur on AH2 (Fig. 16.8) which is part of the lipid-binding pocket (Fig. 16.6a). Furthermore, the role of the protease-like domain in GPAA1 remains unclear. A recent study reports a disease mutation (N323S) near the pseudo-active site, which is facing the bulk solvents (Fig. 16.8), making it difficult to interpret based on the current model. Finally, new information about this complex keeps emerging. For example, EtNP2 has been found to be the preferred bridging moiety for some of the GPI-APs (Ishida et al. 2022). How proprotein substrate induces different orientation of GPI at the active site would require careful structural investigations.

GPI-AP biogenesis is overall conserved among eukaryotes. In many cases, homologs are interchangeable between remote species such as yeast and humans (Hamburger et al. 1995; Hong et al. 2003; Ohishi et al. 2001). However, species specificity exists. For example, the human GPI-T can only process *T. brucei* proproteins with poor efficiency (Moran and Caras 1994), and a fourth mannose on GPI is required for the yeast GPI-T, but not for the human homolog (Taron et al. 2004). Furthermore, the EtNP1 capping on GPI is important for binding with the human GPI-T (Vainauskas and Menon 2006), but GPI from *T. brucei* does not contain this moiety (Hong and Kinoshita 2009). These differences open opportunities for structure-based design of *T. brucei*-specific inhibitors as anti-parasite drugs. In addition to the current structures, this practice would require high-resolution structures of GPI-T from the pathogens.

Acknowledgements I thank the National Natural Science Foundation of China (82151215) and the Strategic Priority Research Program of CAS (XDB37020204) and Science and Technology Commission of Shanghai Municipality (22ZR1468300) for financial support. Structural figures were generated using ChimeraX or PyMol.

References

- Albuquerque-Wendt A, Hütte HJ, Buettner FFR, Routier FH, Bakker H (2019) Membrane topological model of glycosyltransferases of the GT-C superfamily. *Int J Mol Sci* 20
- Bate C, Nolan W, Williams A (2016) Sialic acid on the glycosylphosphatidylinositol anchor regulates PrP-mediated cell signaling and prion formation. *J Biol Chem* 291:160–170
- Becker Kojić ZA, Terness P (2002) A novel human erythrocyte glycosylphosphatidylinositol (GPI)-anchored glycoprotein ACA: ISOLATION, PURIFICATION, PRIMARY STRUCTURE DETERMINATION, AND MOLECULAR PARAMETERS OF ITS LIPID STRUCTURE*. *J Biol Chem* 277:40472–40478
- Beihammer G, Maresch D, Altmann F, Strasser R (2020) Glycosylphosphatidylinositol-anchor synthesis in plants: a glycobiology perspective. *Front Plant Sci* 11:611188
- Ben Ayed I, Jallouli O, Murakami Y, Souissi A, Mallouli S, Bouzid A, Kamoun F, Elloumi I, Frikha F, Tlili A et al (2023) Case report: Functional analysis of the p.Arg507Trp variant of the PIGT gene supporting the moderate epilepsy phenotype of mutations in the C-terminal region. *Front Neurol* 14
- Benachour A, Sipos G, Flury I, Reggiori F, Canivenc-Gansel E, Vionnet C, Conzelmann A, Benghezal M (1999) Deletion of GPI7, a yeast gene required for addition of a side chain to the glycosylphosphatidylinositol (GPI) core structure, affects GPI protein transport, remodeling, and cell wall integrity*. *J Biol Chem* 274:15251–15261
- Benghezal M, Benachour A, Rusconi S, Aebi M, Conzelmann A (1996) Yeast Gpi8p is essential for GPI anchor attachment onto proteins. *EMBO J* 15:6575–6583
- Benghezal M, Lipke PN, Conzelmann A (1995) Identification of six complementation classes involved in the biosynthesis of glycosylphosphatidylinositol anchors in *Saccharomyces cerevisiae*. *J Cell Biol* 130:1333–1344
- Bilal S, Ali L, Khan AL, Shahzad R, Asaf S, Imran M, Kang S-M, Kim S-K, Lee I-J (2018) Endophytic fungus *Paecilomyces formosus* LHL10 produces sester-terpenoid YW3548 and cyclic peptide that inhibit urease and α -glucosidase enzyme activities. *Arch Microbiol* 200:1493–1502
- Bloch JS, Pesciullesi G, Boilevin J, Nosol K, Irobalieva RN, Darbre T, Aebi M, Kossiakoff AA, Raymond J-L, Locher KP (2020) Structure and mechanism of the ER-based glucosyltransferase ALG6. *Nature* 579:443–447
- Bobek LA, Levine MJ (1992) Cystatins — Inhibitors of cysteine proteinases. *Crit Rev Oral Biol Med* 3:307–332
- Bohl H, Bai L, Li H (2021) Recent progress in structural studies on the GT-C superfamily of protein glycosyltransferases. *Subcell Biochem* 96:259–271
- Bosson R, Jaquenoud M, Conzelmann A (2006) GUP1 of *Saccharomyces cerevisiae* encodes an O-acyltransferase involved in remodeling of the GPI anchor. *Mol Biol Cell* 17:2636–2645
- Brown DA, Rose JK (1992) Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical cell surface. *Cell* 68:533–544
- Caiolfa VR, Zamai M, Malengo G, Andolfo A, Madsen CD, Sutin J, Digman MA, Gratton E, Blasi F, Sidenius N (2007) Monomer dimer dynamics and distribution of GPI-anchored uPAR are determined by cell surface protein assemblies. *J Cell Biol* 179:1067–1082
- Cash SJ, McGue BP, Reynolds TS, Crist ER (2020) PIGA related disorder as a range of phenotypes rather than two distinct subtypes. *Brain Dev* 42:205–210

- Castle AMR, Salian S, Bassan H, Sofrin-Drucker E, Cusmai R, Herman KC, Heron D, Keren B, Johnstone DL, Mears W et al (2021) Expanding the phenotypic spectrum of GPI anchoring deficiency due to biallelic variants in GPAA1. *Neurol Genet* 7:e631
- Chang JY (1985) Thrombin specificity. Requirement for apolar amino acids adjacent to the thrombin cleavage site of polypeptide substrate. *Eur J Biochem* 151:217–224
- Chiyonobu T, Inoue N, Morimoto M, Kinoshita T, Murakami Y (2014) Glycosylphosphatidylinositol (GPI) anchor deficiency caused by mutations in PIGW is associated with West syndrome and hyperphosphatasia with mental retardation syndrome. *J Med Genet* 51:203–207
- Conzelmann A, Puoti A, Lester RL, Desponds C (1992) Two different types of lipid moieties are present in glycosylphosphoinositol-anchored membrane proteins of *Saccharomyces cerevisiae*. *EMBO J* 11:457–466
- Dall E, Brandstetter H (2013) Mechanistic and structural studies on legumain explain its zymogenicity, distinct activation pathways, and regulation. *Proc Natl Acad Sci USA* 110:10940–10945
- Dall E, Brandstetter H (2016) Structure and function of legumain in health and disease. *Biochimie* 122:126–150
- Doering TL, Schekman R (1996) GPI anchor attachment is required for Gas1p transport from the endoplasmic reticulum in COP II vesicles. *EMBO J* 15:182–191
- Duval R, Nicolas G, Willemetz A, Murakami Y, Mikdar M, Vignaud C, Megahed H, Cartron J-P, Masson C, Wehbi S et al (2021) Inherited glycosylphosphatidylinositol defects cause the rare Emm-negative blood phenotype and developmental disorders. *Blood* 137:3660–3669
- Earnshaw WC, Martins LM, Kaufmann SH (1999) Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annu Rev Biochem* 68:383–424
- Eisenhaber B, Eisenhaber S, Kwang TY, Grüber G, Eisenhaber F (2014) Transamidase subunit GAA1/GPAA1 is a M28 family metallo-peptide-synthetase that catalyzes the peptide bond formation between the substrate protein's omega-site and the GPI lipid anchor's phosphoethanolamine. *Cell Cycle (Georgetown, Tex)* 13:1912–1917
- Elrod-Erickson MJ, Kaiser CA (1996) Genes that control the fidelity of endoplasmic reticulum to Golgi transport identified as suppressors of vesicle budding mutations. *Mol Biol Cell* 7:1043–1058
- Fabre AL, Orlean P, Taron CH (2005) *Saccharomyces cerevisiae* Ybr004c and its human homologue are required for addition of the second mannose during glycosylphosphatidylinositol precursor assembly. *FEBS J* 272:1160–1168
- Fenollar À, Ros-Lucas A, Pfa Alberione M, Martínez-Peinado N, Ramírez M, Ángel Rosales-Motos M, Lee Y, Alonso-Padilla J, Izquierdo L (2022) Compounds targeting GPI biosynthesis or N-glycosylation are active against *Plasmodium falciparum*. *Comput Struct Biotechnol J* 20:850–863
- Ferguson MA, Homans SW, Dwek RA, Rademacher TW (1988) Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. *Science* 239:753–759
- Fraering P, Imhof I, Meyer U, Strub JM, van Dorsselaer A, Vionnet C, Conzelmann A (2001) The GPI transamidase complex of *Saccharomyces cerevisiae* contains Gaa1p, Gpi8p, and Gpi16p. *Mol Biol Cell* 12:3295–3306
- Fu Y, Estoppey D, Roggo S, Pistorius D, Fuchs F, Studer C, Ibrahim AS, Aust T, Grandjean F, Mihalic M et al (2020) Jawsamycin exhibits in vivo antifungal properties by inhibiting Spt14/Gpi3-mediated biosynthesis of glycosylphosphatidylinositol. *Nat Commun* 11:3387
- Fujihara Y, Ikawa M (2016) GPI-AP release in cellular, developmental, and reproductive biology. *J Lipid Res* 57:538–545
- Fujita M, Maeda Y, Ra M, Yamaguchi Y, Taguchi R, Kinoshita T (2009) GPI glycan remodeling by PGAP5 regulates transport of GPI-anchored proteins from the ER to the Golgi. *Cell* 139:352–365
- Fujita M, Umemura M, Yoko-o T, Jigami Y (2006a) PER1 is required for GPI-phospholipase A2 activity and involved in lipid remodeling of GPI-anchored proteins. *Mol Biol Cell* 17:5253–5264

- Fujita M, Watanabe R, Jaensch N, Romanova-Michaelides M, Satoh T, Kato M, Riezman H, Yamaguchi Y, Maeda Y, Kinoshita T (2011) Sorting of GPI-anchored proteins into ER exit sites by p24 proteins is dependent on remodeled GPI. *J Cell Biol* 194:61–75
- Fujita M, Yoko-o T, Jigami Y (2006b) Inositol deacylation by Bst1p is required for the quality control of glycosylphosphatidylinositol-anchored proteins. *Mol Biol Cell* 17:834–850
- Galian C, Björkholm P, Bulleid N, von Heijne G (2012) Efficient glycosylphosphatidylinositol (GPI) modification of membrane proteins requires a C-terminal anchoring signal of marginal hydrophobicity. *J Biol Chem* 287:16399–16409
- Game DG, Hendrickson TL (2013) GPI transamidase and GPI anchored proteins: oncogenes and biomarkers for cancer. *Crit Rev Biochem Mol Biol* 48:446–464
- Game DG, Varma Y, Meitzler JL, Morissette R, Ness TJ, Hendrickson TL (2017) The soluble domains of Gpi8 and Gaa1, two subunits of glycosylphosphatidylinositol transamidase (GPI-T), assemble into a complex. *Arch Biochem Biophys* 633:58–67
- Gebremariam T, Alkhazraji S, Alqarihi A, Jeon HH, Gu Y, Kapoor M, Shaw KJ, Ibrahim AS (2019) APX001 is effective in the treatment of murine invasive pulmonary aspergillosis. *Antimicrob Agents Chemother* 63
- Gerber LD, Kodukula K, Udenfriend S (1992) Phosphatidylinositol glycan (PI-G) anchored membrane proteins. Amino acid requirements adjacent to the site of cleavage and PI-G attachment in the COOH-terminal signal peptide. *J Biol Chem* 267:12168–12173
- Ghugtyal V, Vionnet C, Roubaty C, Conzelmann A (2007) CWH43 is required for the introduction of ceramides into GPI anchors in *Saccharomyces cerevisiae*. *Mol Microbiol* 65:1493–1502
- Gíslason MH, Nielsen H, Almagro Armenteros JJ, Johansen AR (2021) Prediction of GPI-anchored proteins with pointer neural networks. *Curr Res Biotechnol* 3:6–13
- Guo X-Y, Liu Y-S, Gao X-D, Kinoshita T, Fujita M (2020) Calnexin mediates the maturation of GPI-anchors through ER retention. *J Biol Chem* 295:16393–16410
- Güther ML, de Almeida ML, Yoshida N, Ferguson MA (1992) Structural studies on the glycosylphosphatidylinositol membrane anchor of *Trypanosoma cruzi* 1G7-antigen. The structure of the glycan core. *J Biol Chem* 267:6820–6828
- Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, Wang X, Ademuyiwa F, Ahmed S, Ambrosone CB et al (2011) A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat Genet* 43:1210–1214
- Hamburger D, Egerton M, Riezman H (1995) Yeast Gaa1p is required for attachment of a completed GPI anchor onto proteins. *J Cell Biol* 129:629–639
- Hilley JD, Zawadzki JL, McConville MJ, Coombs GH, Mottram JC (2000) *Leishmania mexicana* mutants lacking glycosylphosphatidylinositol (GPI):protein transamidase provide insights into the biosynthesis and functions of GPI-anchored proteins. *Mol Biol Cell* 11:1183–1195
- Hiroi Y, Chen R, Sawa H, Hosoda T, Kudoh S, Kobayashi Y, Aburatani H, Nagashima K, Nagai R, Yazaki Y et al (2000) Cloning of murine glycosyl phosphatidylinositol anchor attachment protein, GPAA1. *Am J Physiol Cell Physiol* 279:C205–C212
- Hiroi Y, Komuro I, Chen R, Hosoda T, Mizuno T, Kudoh S, Georgescu SP, Medof ME, Yazaki Y (1998) Molecular cloning of human homolog of yeast GAA1 which is required for attachment of glycosylphosphatidylinositols to proteins. *FEBS Lett* 421:252–258
- Holtz AM, Harrington AW, McNamara ER, Kielian A, Soul JS, Martinez-Ojeda M, Levy PT (2020) Expanding the phenotypic spectrum of Mabry Syndrome with novel PIGO gene variants associated with hyperphosphatasia, intractable epilepsy, and complex gastrointestinal and urogenital malformations. *Eur J Med Genet* 63:103802
- Homans SW, Ferguson MA, Dwek RA, Rademacher TW, Anand R, Williams AF (1988) Complete structure of the glycosyl phosphatidylinositol membrane anchor of rat brain Thy-1 glycoprotein. *Nature* 333:269–272
- Hong Y, Kinoshita T (2009) Trypanosome glycosylphosphatidylinositol biosynthesis. *Korean J Parasitol* 47:197–204
- Hong Y, Maeda Y, Watanabe R, Inoue N, Ohishi K, Kinoshita T (2000) Requirement of PIG-F and PIG-O for transferring phosphoethanolamine to the third mannose in glycosylphosphatidylinositol. *J Biol Chem* 275:20911–20919

- Hong Y, Maeda Y, Watanabe R, Ohishi K, Mishkind M, Riezman H, Kinoshita T (1999) Pig-n, a mammalian homologue of yeast Mcd4p, is involved in transferring phosphoethanolamine to the first mannose of the glycosylphosphatidylinositol. *J Biol Chem* 274:35099–35106
- Hong Y, Nagamune K, Ohishi K, Morita YS, Ashida H, Maeda Y, Kinoshita T (2006) TbGPI16 is an essential component of GPI transamidase in *Trypanosoma brucei*. *FEBS Lett* 580:603–606
- Hong Y, Ohishi K, Kang JY, Tanaka S, Inoue N, Nishimura J-I, Maeda Y, Kinoshita T (2003) Human PIG-U and yeast Cdc91p are the fifth subunit of GPI transamidase that attaches GPI-anchors to proteins. *Mol Biol Cell* 14:1780–1789
- Huang X, Liu Y, Ni T, Li L, Yan L, An M, Zhang D, Jiang Y (2020) 11g, a potent antifungal candidate, enhances *Candida albicans* immunogenicity by unmasking β -glucan in fungal cell wall. *Front Microbiol* 11:1324
- Ishida M, Maki Y, Ninomiya A, Takada Y, Campeau P, Kinoshita T, Murakami Y (2022) Ethanolamine-phosphate on the second mannose is a preferential bridge for some GPI-anchored proteins. *EMBO Rep* 23:e54352
- Jezela-Stanek A, Mierzevska H, Szczepanik E (2020a) Vertical nystagmus as a feature of PIGN-related glycosylphosphatidylinositol biosynthesis defects. *Clin Neurol Neurosurg* 196:106033
- Jezela-Stanek A, Szczepanik E, Mierzevska H, Rydzanicz M, Rutkowska K, Knaus A, Śmigiel R, Stepniak I, Markiewicz MG, Boniel S et al (2020b) Evidence of the milder phenotypic spectrum of c.1582G>A PIGT variant: delineation based on seven novel Polish patients. *Clin Genet* 98:468–476
- Jia J, Bosley AD, Thompson A, Hoskins JW, Cheuk A, Collins I, Parikh H, Xiao Z, Ylaya K, Dzyadyk M et al (2014) CLPTM1L promotes growth and enhances aneuploidy in pancreatic cancer cells. *Cancer Res* 74:2785–2795
- Johnstone DL, Nguyen TTM, Zamboni J, Kernohan KD, St-Denis A, Baratang NV, Hartley T, Geraghty MT, Richer J, Majewski J et al (2020) Early infantile epileptic encephalopathy due to biallelic pathogenic variants in PIGQ: Report of seven new subjects and review of the literature. *J Inher Metab Dis* 43:1321–1332
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A et al (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583–589
- Kang JY, Hong Y, Ashida H, Shishioh N, Murakami Y, Morita YS, Maeda Y, Kinoshita T (2005) PIG-V involved in transferring the second mannose in glycosylphosphatidylinositol. *J Biol Chem* 280:9489–9497
- Kang X, Szallies A, Rawer M, Echner H, Duzenko M (2002) GPI anchor transamidase of *Trypanosoma brucei*: in vitro assay of the recombinant protein and VSG anchor exchange. *J Cell Sci* 115:2529–2539
- Kanzawa N, Maeda Y, Ogiso H, Murakami Y, Taguchi R, Kinoshita T (2009) Peroxisome dependency of alkyl-containing GPI-anchor biosynthesis in the endoplasmic reticulum. *Proc Natl Acad Sci USA* 106:17711–17716
- Katorcha E, Srivastava S, Klimova N, Baskakov IV (2016) Sialylation of glycosylphosphatidylinositol (GPI) anchors of mammalian prions is regulated in a host-, tissue-, and cell-specific manner*. *J Biol Chem* 291:17009–17019
- Kawaguchi K, Sato T, Kondo S, Yamamoto-Hino M, Goto S (2019) Stability of the transamidase complex catalyzing GPI anchoring of proteins. *Biochem Biophys Res Commun* 512:584–590
- Kesavardhana S, Malireddi RKS, Kanneganti T-D (2020) Caspases in cell death, inflammation, and pyroptosis. *Annu Rev Immunol* 38:567–595
- Kinoshita T (2014) Biosynthesis and deficiencies of glycosylphosphatidylinositol. *Proc Jpn Acad Ser B* 90:130–143
- Kinoshita T (2020) Biosynthesis and biology of mammalian GPI-anchored proteins. *Open Biol* 10:190290
- Kinoshita T, Fujita M (2016) Thematic Review Series: Glycosylphosphatidylinositol (GPI) anchors: biochemistry and cell biology biosynthesis of GPI-anchored proteins: special emphasis on GPI lipid remodeling. *J Lipid Res* 57:6–24

- Kinoshita T, Maeda Y, Fujita M (2013) Transport of glycosylphosphatidylinositol-anchored proteins from the endoplasmic reticulum. *Biochim Biophys Acta, Mol Cell Res* 1833:2473–2478
- Knaus A, Kortüm F, Kleefstra T, Stray-Pedersen A, Đukić D, Murakami Y, Gerstner T, van Bokhoven H, Iqbal Z, Horn D et al (2019) Mutations in *PIGU* impair the function of the GPI transamidase complex, causing severe intellectual disability, epilepsy, and brain anomalies. *Am J Hum Genet* 105:395–402
- Kodukula K, Cines D, Amthauer R, Gerber L, Udenfriend S (1992) Biosynthesis of phosphatidylinositol-glycan (PI-G)-anchored membrane proteins in cell-free systems: cleavage of the nascent protein and addition of the PI-G moiety depend on the size of the COOH-terminal signal peptide. *Proc Natl Acad Sci U S A* 89:1350–1353
- Kodukula K, Gerber LD, Amthauer R, Brink L, Udenfriend S (1993) Biosynthesis of glycosylphosphatidylinositol (GPI)-anchored membrane proteins in intact cells: specific amino acid requirements adjacent to the site of cleavage and GPI attachment. *J Cell Biol* 120:657–664
- Kodukula K, Micanovic R, Gerber L, Tamburrini M, Brink L, Udenfriend S (1991) Biosynthesis of phosphatidylinositol glycan-anchored membrane proteins. Design of a simple protein substrate to characterize the enzyme that cleaves the COOH-terminal signal peptide*. *J Biol Chem* 266:4464–4470
- Kohashi K, Ishiyama A, Yuasa S, Tanaka T, Miya K, Adachi Y, Sato N, Saitsu H, Ohba C, Matsumoto N et al (2018) Epileptic apnea in a patient with inherited glycosylphosphatidylinositol anchor deficiency and *PIGT* mutations. *Brain Dev* 40:53–57
- Kvarnung M, Nilsson D, Lindstrand A, Korenke GC, Chiang SC, Blennow E, Bergmann M, Stöbberg T, Mäkitie O, Anderlid BM et al (2013) A novel intellectual disability syndrome caused by GPI anchor deficiency due to homozygous mutations in *PIGT*. *J Med Genet* 50:521–528
- Lando M, Holden M, Bergersen LC, Svendsrud DH, Stokke T, Sundfør K, Glad IK, Kristensen GB, Lyng H (2009) Gene dosage, expression, and ontology analysis identifies driver genes in the carcinogenesis and chemoradioresistance of cervical cancer. *PLoS Genet* 5:e1000719
- Lebreton S, Paladino S, Liu D, Nitti M, von Blume J, Pinton P, Zurzolo C (2021) Calcium levels in the Golgi complex regulate clustering and apical sorting of GPI-APs in polarized epithelial cells. *Proc Natl Acad Sci U S A* 118:e2014709118
- Leidich SD, Kostova Z, Latek RR, Costello LC, Drapp DA, Gray W, Fassler JS, Orlean P (1995) Temperature-sensitive yeast GPI anchoring mutants *gpi2* and *gpi3* are defective in the synthesis of N-acetylglucosaminyl phosphatidylinositol. CLONING OF THE *GPI2* GENE*. *J Biol Chem* 270:13029–13035
- Li Y, Yang L, Yang J, Shi J, Chai P, Ge S, Wang Y, Fan X, Jia R (2020) A novel variant in *GPAAL1*, encoding a GPI transamidase complex protein, causes inherited vascular anomalies with various phenotypes. *Hum Genet* 139:1499–1511
- Lillico S, Field MC, Blundell P, Coombs GH, Mottram JC (2003) Essential roles for GPI-anchored proteins in African trypanosomes revealed using mutants deficient in *GPI8*. *Mol Biol Cell* 14:1182–1194
- Liu SS, Liu YS, Guo XY, Murakami Y, Yang G, Gao XD, Kinoshita T, Fujita M (2021) A knock-out cell library of GPI biosynthetic genes for functional studies of GPI-anchored proteins. *Commun Biol* 4:777
- Liu W, Zou Z, Huang X, Shen H, He LJ, Chen SM, Li LP, Yan L, Zhang SQ, Zhang JD et al (2016) *Bst1* is required for *Candida albicans* infecting host via facilitating cell wall anchorage of Glycosylphosphatidyl inositol anchored proteins. *Sci Rep* 6:34854
- Liu Y-S, Fujita M (2020) Mammalian GPI-anchor modifications and the enzymes involved. *Biochem Soc Trans* 48:1129–1138
- Liu YS, Guo XY, Hirata T, Rong Y, Motooka D, Kitajima T, Murakami Y, Gao XD, Nakamura S, Kinoshita T et al (2018) N-Glycan-dependent protein folding and endoplasmic reticulum retention regulate GPI-anchor processing. *J Cell Biol* 217:585–599
- Loong L, Tardivo A, Knaus A, Hashim M, Pagnamenta AT, Alt K, Böhrer-Rabel H, Caro-Llopis A, Cole T, Distelmaier F et al (2023) Biallelic variants in *PIGN* cause Fryns syndrome, multiple

- congenital anomalies-hypotonia-seizures syndrome, and neurologic phenotypes: a genotype-phenotype correlation study. *Genet Med* 25:37–48
- Maeda Y, Tashima Y, Houjou T, Fujita M, Yoko-o T, Jigami Y, Taguchi R, Kinoshita T (2007) Fatty acid remodeling of GPI-anchored proteins is required for their raft association. *Mol Biol Cell* 18:1497–1506
- Maeda Y, Watanabe R, Harris CL, Hong Y, Ohishi K, Kinoshita K, Kinoshita T (2001) PIG-M transfers the first mannose to glycosylphosphatidylinositol on the luminal side of the ER. *EMBO J* 20:250–261
- Mann PA, McLellan CA, Koseoglu S, Si Q, Kuzmin E, Flattery A, Harris G, Sher X, Murgolo N, Wang H et al (2015) Chemical genomics-based antifungal drug discovery: targeting glycosylphosphatidylinositol (GPI) precursor biosynthesis. *ACS Infect Dis* 1:59–72
- Masuishi Y, Kimura Y, Arakawa N, Hirano H (2016) Identification of glycosylphosphatidylinositol-anchored proteins and ω -sites using TiO₂-based affinity purification followed by hydrogen fluoride treatment. *J Proteome* 139:77–83
- Maxwell SE, Ramalingam S, Gerber LD, Brink L, Udenfriend S (1995a) An active carbonyl formed during glycosylphosphatidylinositol addition to a protein is evidence of catalysis by a transamidase. *J Biol Chem* 270:19576–19582
- Maxwell SE, Ramalingam S, Gerber LD, Udenfriend S (1995b) Cleavage without anchor addition accompanies the processing of a nascent protein to its glycosylphosphatidylinositol-anchored form. *Proc Natl Acad Sci U S A* 92:1550–1554
- McConville MJ, Ferguson MA (1993) The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. *Biochem J* 294(Pt 2):305–324
- McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J et al (2008) Lung cancer susceptibility locus at 5p15.33. *Nat Genet* 40:1404–1406
- McKean DM, Niswander L (2012) Defects in GPI biosynthesis perturb Cripto signaling during forebrain development in two new mouse models of holoprosencephaly. *Biology open* 1:874–883
- Meyer U, Benghezal M, Imhof I, Conzelmann A (2000) Active site determination of Gpi8p, a caspase-related enzyme required for glycosylphosphatidylinositol anchor addition to proteins. *Biochemistry* 39:3461–3471
- Moran P, Caras IW (1994) Requirements for glycosylphosphatidylinositol attachment are similar but not identical in mammalian cells and parasitic protozoa. *J Cell Biol* 125:333–343
- Morissette R, Varma Y, Hendrickson TL (2012) Defining the boundaries of species specificity for the *Saccharomyces cerevisiae* glycosylphosphatidylinositol transamidase using a quantitative in vivo assay. *Biosci Rep* 32:577–586
- Murakami Y, Siripanyaphinyo U, Hong Y, Tashima Y, Maeda Y, Kinoshita T (2005) The initial enzyme for glycosylphosphatidylinositol biosynthesis requires PIG-Y, a seventh component. *Mol Biol Cell* 16:5236–5246
- Nagamune K, Ohishi K, Ashida H, Hong Y, Hino J, Kangawa K, Inoue N, Maeda Y, Kinoshita T (2003) GPI transamidase of *Trypanosoma brucei* has two previously uncharacterized (trypanosomatid transamidase 1 and 2) and three common subunits. *Proc Natl Acad Sci U S A* 100:10682–10687
- Naik RS, Branch OH, Woods AS, Vijaykumar M, Perkins DJ, Nahlen BL, Lal AA, Cotter RJ, Costello CE, Ockenhouse CF et al (2000) Glycosylphosphatidylinositol anchors of *Plasmodium falciparum*: molecular characterization and naturally elicited antibody response that may provide immunity to malaria pathogenesis. *J Exp Med* 192:1563–1576
- Nakashima M, Kashii H, Murakami Y, Kato M, Tsurusaki Y, Miyake N, Kubota M, Kinoshita T, Saito H, Matsumoto N (2014) Novel compound heterozygous PIGT mutations caused multiple congenital anomalies-hypotonia-seizures syndrome 3. *Neurogenetics* 15:193–200

- Ness TJ, Gamage DG, Ekanayaka SA, Hendrickson TL (2022) A soluble, minimalistic glycosylphosphatidylinositol transamidase (GPI-T) retains transamidation activity. *Biochemistry* 61:1273–1285
- Nguyen TTM, Murakami Y, Mobilio S, Niceta M, Zampino G, Philippe C, Moutton S, Zaki MS, James KN, Musaev D et al (2020) Bi-allelic variants in the GPI transamidase subunit PIGK cause a neurodevelopmental syndrome with hypotonia, cerebellar atrophy, and epilepsy. *Am J Hum Genet* 106:484–495
- Nguyen TTM, Murakami Y, Sheridan E, Ehresmann S, Rousseau J, St-Denis A, Chai G, Ajeawung NF, Fairbrother L, Reimschisel T et al (2017) Mutations in GPAA1, encoding a GPI transamidase complex protein, cause developmental delay, epilepsy, cerebellar atrophy, and osteopenia. *Am J Hum Genet* 101:856–865
- Nguyen TTM, Murakami Y, Wigby KM, Baratang NV, Rousseau J, St-Denis A, Rosenfeld JA, Laniewski SC, Jones J, Iglesias AD et al (2018) Mutations in PIGS, encoding a GPI transamidase, cause a neurological syndrome ranging from fetal akinesia to epileptic encephalopathy. *Am J Hum Genet* 103:602–611
- Ni Z, Tao K, Chen G, Chen Q, Tang J, Luo X, Yin P, Tang J, Wang X (2012) CLPTM1L is overexpressed in lung cancer and associated with apoptosis. *PLoS One* 7:e52598
- Ohishi K, Inoue N, Kinoshita T (2001) PIG-S and PIG-T, essential for GPI anchor attachment to proteins, form a complex with GAA1 and GPI8. *EMBO J* 20:4088–4098
- Ohishi K, Inoue N, Maeda Y, Takeda J, Riezman H, Kinoshita T (2000) Gaa1p and gpi8p are components of a glycosylphosphatidylinositol (GPI) transamidase that mediates attachment of GPI to proteins. *Mol Biol Cell* 11:1523–1533
- Ohishi K, Nagamune K, Maeda Y, Kinoshita T (2003) Two subunits of glycosylphosphatidylinositol transamidase, GPI8 and PIG-T, form a functionally important intermolecular disulfide bridge. *J Biol Chem* 278:13959–13967
- Orlean P, Menon AK (2007) Thematic review series: Lipid posttranslational modifications. GPI anchoring of protein in yeast and mammalian cells, or: how we learned to stop worrying and love glycopospholipids. *J Lipid Res* 48:993–1011
- Pagnamenta AT, Murakami Y, Taylor JM, Anzilotti C, Howard MF, Miller V, Johnson DS, Tadros S, Mansour S, Temple IK et al (2017) Analysis of exome data for 4293 trios suggests GPI-anchor biogenesis defects are a rare cause of developmental disorders. *Eur J Hum Genet* 25:669–679
- Paladino S, Lebreton S, Lelek M, Riccio P, De Nicola S, Zimmer C, Zurzolo C (2017) GPI-anchored proteins are confined in subdiffraction clusters at the apical surface of polarized epithelial cells. *Biochem J* 474:4075–4090
- Parashar D, Geethadevi A, McAllister D, Ebben J, Peterson FC, Jensen DR, Bishop E, Pradeep S, Volkman BF, Dwinell MB et al (2021) Targeted biologic inhibition of both tumor cell-intrinsic and intercellular CLPTM1L/CRR9-mediated chemotherapeutic drug resistance. *NPJ Precis Oncol* 5:16
- Peron A, Iascone M, Salvatici E, Cavarani B, Marchetti D, Corno S, Vignoli A (2020) PIGW-related glycosylphosphatidylinositol deficiency: description of a new patient and review of the literature. *Am J Med Genet A* 182:1477–1482
- Petrou VI, Herrera CM, Schultz KM, Clarke OB, Vendome J, Tomasek D, Banerjee S, Rajashankar KR, Belcher Dufrisne M, Kloss B et al (2016) Structures of aminoarabinose transferase ArnT suggest a molecular basis for lipid A glycosylation. *Science* 351:608–612
- Pierleoni A, Martelli PL, Casadio R (2008) PredGPI: a GPI-anchor predictor. *BMC Bioinform* 9:392
- Pittet M, Conzelmann A (2007) Biosynthesis and function of GPI proteins in the yeast *Saccharomyces cerevisiae*. *Biochim Biophys Acta* 1771:405–420
- Puskás LG, Mán I, Szebeni G, Tiszlavicz L, Tsai S, James MA (2016) Novel anti-CRR9/CLPTM1L antibodies with antitumorigenic activity inhibit cell surface accumulation, PI3K interaction, and survival signaling. *Mol Cancer Ther* 15:985–997
- Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, Jakobsdottir M, Helgadóttir H, Thorlacius S, Aben KKH et al (2009) Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet* 41:221–227

- Reinheckel T, Tholen M (2022) Low-level lysosomal membrane permeabilization for limited release and sublethal functions of cathepsin proteases in the cytosol and nucleus. *FEBS Open Bio* 12:694–707
- Richard M, Ibatu-Ombetta S, Dromer F, Bordon-Pallier F, Jouault T, Gaillardin C (2002) Complete glycosylphosphatidylinositol anchors are required in *Candida albicans* for full morphogenesis, virulence and resistance to macrophages. *Mol Microbiol* 44:841–853
- Richier P, Arpagaus M, Toutant JP (1992) Glycolipid-anchored acetylcholinesterases from rabbit lymphocytes and erythrocytes differ in their sensitivity to phosphatidylinositol-specific phospholipase C. *Biochim Biophys Acta* 1112:83–88
- Rodriguez-Gallardo S, Sabido-Bozo S, Ikeda A, Araki M, Okazaki K, Nakano M, Aguilera-Romero A, Cortes-Gomez A, Lopez S, Waga M et al (2022) Quality-controlled ceramide-based GPI-anchored protein sorting into selective ER exit sites. *Cell Rep* 39:110768
- Sahu PK, Tomar RS (2019) The natural anticancer agent cantharidin alters GPI-anchored protein sorting by targeting Cdc1-mediated remodeling in endoplasmic reticulum. *J Biol Chem* 294:3837–3852
- Salian S, Benkerroum H, Nguyen TTM, Nampoothiri S, Kinoshita T, Félix TM, Stewart F, Sisodiya SM, Murakami Y, Campeau PM (2021) PIGF deficiency causes a phenotype overlapping with DOORS syndrome. *Hum Genet* 140:879–884
- Salian S, Guo X-Y, Murakami Y, Kinoshita T, Kaur P, Shukla A, Girisha KM, Fujita M, Campeau PM (2022) C18orf32 loss-of-function is associated with a neurodevelopmental disorder with hypotonia and contractures. *Hum Genet* 141:1423–1429
- Sauer LM, Canovas R, Roche D, Shams-Eldin H, Ravel P, Colinge J, Schwarz RT, Ben Mamoun C, Rivals E, Cornillot E (2023) FT-GPI, a highly sensitive and accurate predictor of GPI-anchored proteins, reveals the composition and evolution of the GPI proteome in *Plasmodium* species. *Malar J* 22:27
- Schechter I, Berger A (1967) On the size of the active site in proteases. I. Papain *Biochem Biophys Res Commun* 27:157–162
- Schneider P, Ferguson MA, McConville MJ, Mehlert A, Homans SW, Bordier C (1990) Structure of the glycosyl-phosphatidylinositol membrane anchor of the *Leishmania major* promastigote surface protease. *J Biol Chem* 265:16955–16964
- Seong J, Wang Y, Kinoshita T, Maeda Y (2013) Implications of lipid moiety in oligomerization and immunoreactivities of GPI-anchored proteins. *J Lipid Res* 54:1077–1091
- Sezgin E, Levental I, Mayor S, Eggeling C (2017) The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nat Rev Mol Cell Biol* 18:361–374
- Sharma DK, Hilley JD, Bangs JD, Coombs GH, Mottram JC, Menon AK (2000) Soluble GPI8 restores glycosylphosphatidylinositol anchoring in a trypanosome cell-free system depleted of luminal endoplasmic reticulum proteins. *Biochem J* 351(Pt 3):717–722
- Sharma P, Varma R, Sarasij RC, Ira, Gousset K, Krishnamoorthy G, Rao M, Mayor S (2004) Nanoscale organization of multiple GPI-anchored proteins in living cell membranes. *Cell* 116:577–589
- Shishioh N, Hong Y, Ohishi K, Ashida H, Maeda Y, Kinoshita T (2005) GPI7 is the second partner of PIG-F and involved in modification of glycosylphosphatidylinositol*. *J Biol Chem* 280:9728–9734
- Skauli N, Wallace S, Chiang SC, Barøy T, Holmgren A, Stray-Pedersen A, Bryceson YT, Strømme P, Frengen E, Misceo D (2016) Novel PIGT variant in two brothers: expansion of the multiple congenital anomalies-hypotonia seizures syndrome 3 phenotype. *Genes (Basel)* 7
- Smith TK, Crossman A, Brimacombe JS, Ferguson MA (2004) Chemical validation of GPI biosynthesis as a drug target against African sleeping sickness. *EMBO J* 23:4701–4708
- Spurway TD, Dalley JA, High S, Bulleid NJ (2001) Early events in glycosylphosphatidylinositol anchor addition. Substrate proteins associate with the transamidase subunit gpi8p. *J Biol Chem* 276:15975–15982
- Stahl N, Baldwin M, Hecker R, Pan K-M, Burlingame A, Prusiner S (1992) Glycosylphosphatidylinositol anchors of the scrapie and cellular prion proteins contain sialic acid. *Biochemistry* 31:5043–5053

- Su CT, Sinha S, Eisenhaber B, Eisenhaber F (2020) Structural modelling of the luminal domain of human GPAA1, the metallo-peptide synthetase subunit of the transamidase complex, reveals zinc-binding mode and two flaps surrounding the active site. *Biol Direct* 15:14
- Sütterlin C, Horvath A, Gerold P, Schwarz RT, Wang Y, Dreyfuss M, Riezman H (1997) Identification of a species-specific inhibitor of glycosylphosphatidylinositol synthesis. *EMBO J* 16:6374–6383
- Suzuki KG, Kasai RS, Hirose KM, Nemoto YL, Ishibashi M, Miwa Y, Fujiwara TK, Kusumi A (2012) Transient GPI-anchored protein homodimers are units for raft organization and function. *Nat Chem Biol* 8:774–783
- Takahashi M, Inoue N, Ohishi K, Maeda Y, Nakamura N, Endo Y, Fujita T, Takeda J, Kinoshita T (1996) PIG-B, a membrane protein of the endoplasmic reticulum with a large luminal domain, is involved in transferring the third mannose of the GPI anchor. *EMBO J* 15:4254–4261
- Tanaka S, Maeda Y, Tashima Y, Kinoshita T (2004) Inositol deacylation of glycosylphosphatidylinositol-anchored proteins is mediated by mammalian PGAP1 and yeast Bst1p*. *J Biol Chem* 279:14256–14263
- Taron BW, Colussi PA, Wiedman JM, Orlean P, Taron CH (2004) Human Smp3p adds a fourth mannose to yeast and human glycosylphosphatidylinositol precursors in vivo. *J Biol Chem* 279:36083–36092
- Tashima Y, Taguchi R, Murata C, Ashida H, Kinoshita T, Maeda Y (2006) PGAP2 is essential for correct processing and stable expression of GPI-anchored proteins. *Mol Biol Cell* 17:1410–1420
- Thompson MD, Knaus AA, Barshop BA, Caliebe A, Muhle H, Nguyen TTM, Baratang NV, Kinoshita T, Percy ME, Campeau PM et al (2020) A post glycosylphosphatidylinositol (GPI) attachment to proteins, type 2 (PGAP2) variant identified in Mabry syndrome index cases: Molecular genetics of the prototypical inherited GPI disorder. *Eur J Med Genet* 63:103822
- Thompson MD, Li X, Spencer-Manzon M, Andrade DM, Murakami Y, Kinoshita T, Carpenter TO (2023) Excluding digenic inheritance of PGAP2 and PGAP3 variants in mabry syndrome (OMIM 239300) patient: phenotypic spectrum associated with PGAP2 gene variants in hyperphosphatasia with mental retardation syndrome-3 (HPMRS3). *Genes* 14(2):359
- Tremblay-Laganière C, Maroofian R, Nguyen TTM, Karimiani EG, Kirmani S, Akbar F, Ibrahim S, Afroz B, Doosti M, Ashrafzadeh F et al (2021) PIGG variant pathogenicity assessment reveals characteristic features within 19 families. *Genet Med* 23:1873–1881
- Tsai YH, Liu X, Seeberger PH (2012) Chemical biology of glycosylphosphatidylinositol anchors. *Angew Chem (Int ed in English)* 124:11438–11456
- Umemura M, Fujita M, Yoko-o T, Fukamizu A, Jigami Y (2007) *Saccharomyces cerevisiae* CWH43 is involved in the remodeling of the lipid moiety of GPI anchors to ceramides. *Mol Biol Cell* 18:4304–4316
- Umemura M, Okamoto M, Nakayama K, Sagane K, Tsukahara K, Hata K, Jigami Y (2003) GWT1 gene is required for inositol acylation of glycosylphosphatidylinositol anchors in yeast. *J Biol Chem* 278:23639–23647
- Vainauskas S, Maeda Y, Kurniawan H, Kinoshita T, Menon AK (2002) Structural requirements for the recruitment of Gaa1 into a functional glycosylphosphatidylinositol transamidase complex. *J Biol Chem* 277:30535–30542
- Vainauskas S, Menon AK (2004) A conserved proline in the last transmembrane segment of Gaa1 is required for glycosylphosphatidylinositol (GPI) recognition by GPI transamidase. *J Biol Chem* 279:6540–6545
- Vainauskas S, Menon AK (2005) Endoplasmic reticulum localization of Gaa1 and PIG-T, subunits of the glycosylphosphatidylinositol transamidase complex. *J Biol Chem* 280:16402–16409
- Vainauskas S, Menon AK (2006) Ethanolamine phosphate linked to the first mannose residue of glycosylphosphatidylinositol (GPI) lipids is a major feature of the GPI structure that is recognized by human GPI transamidase*. *J Biol Chem* 281:38358–38364
- Vashist S, Kim W, Belden WJ, Spear ED, Barlowe C, Ng DTW (2001) Distinct retrieval and retention mechanisms are required for the quality control of endoplasmic reticulum protein folding. *J Cell Biol* 155:355–368

- Vetro A, Pisano T, Chiaro S, Procopio E, Guerra A, Parrini E, Mei D, Virdò S, Mangone G, Azzari C et al (2020) Early infantile epileptic-dyskinetic encephalopathy due to biallelic PIGP mutations. *Neurol Genet* 6:e387
- Vidugiriene J, Vainauskas S, Johnson AE, Menon AK (2001) Endoplasmic reticulum proteins involved in glycosylphosphatidylinositol-anchor attachment: photocrosslinking studies in a cell-free system. *Eur J Biochem* 268:2290–2300
- Viriyakosol S, Kapoor M, Okamoto S, Covell J, Soltow QA, Trzoss M, Shaw KJ, Fierer J (2019) APX001 and other Gwt1 inhibitor prodrugs are effective in experimental *Coccidioides immitis* pneumonia. *Antimicrob Agents Chemother* 63
- Walter EI, Roberts WL, Rosenberry TL, Ratnoff WD, Medof ME (1990) Structural basis for variations in the sensitivity of human decay accelerating factor to phosphatidylinositol-specific phospholipase C cleavage. *J Immunol* 144:1030–1036
- Wang Y, Menon AK, Maki Y, Liu YS, Iwasaki Y, Fujita M, Guerrero PA, Silva DV, Seeberger PH, Murakami Y et al (2022) Genome-wide CRISPR screen reveals CLPTM1L as a lipid scramblase required for efficient glycosylphosphatidylinositol biosynthesis. *Proc Natl Acad Sci U S A* 119:e2115083119
- Wang Z, Zhu B, Zhang M, Parikh H, Jia J, Chung CC, Sampson JN, Hoskins JW, Hutchinson A, Burdette L et al (2014) Imputation and subset-based association analysis across different cancer types identifies multiple independent risk loci in the TERT-CLPTM1L region on chromosome 5p15.33. *Hum Mol Genet* 23:6616–6633
- Watanabe NA, Miyazaki M, Horii T, Sagane K, Tsukahara K, Hata K (2012) E1210, a new broad-spectrum antifungal, suppresses *Candida albicans* hyphal growth through inhibition of glycosylphosphatidylinositol biosynthesis. *Antimicrob Agents Chemother* 56:960–971
- Watanabe R, Ohishi K, Maeda Y, Nakamura N, Kinoshita T (1999) Mammalian PIG-L and its yeast homologue Gpi12p are N-acetylglucosaminylphosphatidylinositol de-N-acetylases essential in glycosylphosphatidylinositol biosynthesis. *Biochem J* 339(Pt 1):185–192
- Wiederhold NP, Najvar LK, Shaw KJ, Jaramillo R, Patterson H, Olivo M, Catano G, Patterson TF (2019) Efficacy of delayed therapy with fosmanogepix (APX001) in a murine model of *Candida auris* invasive candidiasis. *Antimicrob Agents Chemother* 63
- Wu G, Guo Z, Chatterjee A, Huang X, Rubin E, Wu F, Mambo E, Chang X, Osada M, Sook Kim M et al (2006) Overexpression of glycosylphosphatidylinositol (GPI) transamidase subunits phosphatidylinositol glycan class T and/or GPI anchor attachment 1 induces tumorigenesis and contributes to invasion in human breast cancer. *Cancer Res* 66:9829–9836
- Xu Y, Jia G, Li T, Zhou Z, Luo Y, Chao Y, Bao J, Su Z, Qu Q, Li D (2022) Molecular insights into biogenesis of glycosylphosphatidylinositol anchor proteins. *Nat Commun* 13:2617
- Yadav B, Bhatnagar S, Ahmad MF, Jain P, Pratyusha VA, Kumar P, Komath SS (2014) First step of glycosylphosphatidylinositol (GPI) biosynthesis cross-talks with ergosterol biosynthesis and Ras signaling in *Candida albicans*. *J Biol Chem* 289:3365–3382
- Yadav U, Khan MA (2018) Targeting the GPI biosynthetic pathway. *Pathogens Glob Health* 112:115–122
- Yamamoto K, Okamoto A, Isonishi S, Ochiai K, Ohtake Y (2001) A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. *Biochem Biophys Res Commun* 280:1148–1154
- Yang G, Banfield DK (2020) Cdc1p is a Golgi-localized glycosylphosphatidylinositol-anchored protein remodelase. *Mol Biol Cell* 31:2883–2891
- Yang L, Peng J, Yin XM, Pang N, Chen C, Wu TH, Zou XM, Yin F (2018) Homozygous PIGT mutation lead to multiple congenital anomalies-hypotonia seizures syndrome 3. *Front Genet* 9:153
- Yi L, Bozkurt G, Li Q, Lo S, Menon AK, Wu H (2017) Disulfide bond formation and N-glycosylation modulate protein-protein interactions in GPI-transamidase (GPIT). *Sci Rep* 8:45912
- Yu J, Nagarajan S, Knez JJ, Udenfriend S, Chen R, Medof ME (1997) The affected gene underlying the class K glycosylphosphatidylinositol (GPI) surface protein defect codes for the GPI transamidase. *Proc Natl Acad Sci U S A* 94:12580–12585

- Zanni G, D'Abrusco F, Nicita F, Cascioli S, Tosi M, Corrente F, Serpieri V, Ciccone R, Motta M, Vasco G et al (2022) PIGQ-related glycosylphosphatidylinositol deficiency associated with non-progressive congenital ataxia. *Cerebellum* 21:525–530
- Zhang H, Su J, Li B, Gao Y, Liu M, He L, Xu H, Dong Y, Zhang XC, Zhao Y (2022) Structure of human glycosylphosphatidylinositol transamidase. *Nat Struct Mol Biol* 29:203–209
- Zhang X-X, Ni B, Li Q, Hu L-P, Jiang S-H, Li R-K, Tian G-A, Zhu L-L, Li J, Zhang X-L et al (2019) GPAA1 promotes gastric cancer progression via upregulation of GPI-anchored protein and enhancement of ERBB signalling pathway. *J Exp Clin Cancer Res* 38:214
- Zhu Y, Fraering P, Vionnet C, Conzelmann A (2005) Gpi17p does not stably interact with other subunits of glycosylphosphatidylinositol transamidase in *Saccharomyces cerevisiae*. *Biochim Biophys Acta* 1735:79–88
- Zurzolo C, van't Hof W, van Meer G, Rodriguez-Boulan E (1994) VIP21/caveolin, glycosphingolipid clusters and the sorting of glycosylphosphatidylinositol-anchored proteins in epithelial cells. *EMBO J* 13:42–53

Chapter 17

Maturation and Assembly of mTOR Complexes by the HSP90-R2TP-TTT Chaperone System: Molecular Insights and Mechanisms



Andrés López-Perrote, Marina Serna, and Oscar Llorca

Abstract The mechanistic target of rapamycin (mTOR) is a master regulator of cell growth and metabolism, integrating environmental signals to regulate anabolic and catabolic processes, regulating lipid synthesis, growth factor-induced cell proliferation, cell survival, and migration. These activities are performed as part of two distinct complexes, mTORC1 and mTORC2, each with specific roles. mTORC1 and mTORC2 are elaborated dimeric structures formed by the interaction of mTOR with specific partners. mTOR functions only as part of these large complexes, but their assembly and activation require a dedicated and sophisticated chaperone system. mTOR folding and assembly are temporarily separated with the TELO2-TTI1-TTI2 (TTT) complex assisting the cotranslational folding of mTOR into a native conformation. Matured mTOR is then transferred to the R2TP complex for assembly of active mTORC1 and mTORC2 complexes. R2TP works in concert with the HSP90 chaperone to promote the incorporation of additional subunits to mTOR and dimerization. This review summarizes our current knowledge on how the HSP90-R2TP-TTT chaperone system facilitates the maturation and assembly of active mTORC1 and mTORC2 complexes, discussing interactions, structures, and mechanisms.

Keywords mTORC1 · mTORC2 · Chaperones · TELO2 · TTT · RUVBL1 · RUVBL2 · R2TP · HSP90

A. López-Perrote (✉) · M. Serna · O. Llorca (✉)
Spanish National Cancer Research Centre (CNIO), Structural Biology Programme, Melchor
Fernández Almagro 3, Madrid, Spain
e-mail: aperrote@cnio.es; mserna@cnio.es; ollorca@cnio.es

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2024

J. R. Harris, J. Marles-Wright (eds.), *Macromolecular Protein Complexes V*,
Subcellular Biochemistry 104, https://doi.org/10.1007/978-3-031-58843-3_17

mTOR Assembles Two Large Multisubunit Complexes, mTORC1 and mTORC2

The mechanistic (formerly referred to as mammalian) target of rapamycin (mTOR) is an evolutionarily conserved serine/threonine kinase, and a member of the family of phosphatidylinositol 3-kinase-related kinases (PIKKs) (Keith and Schreiber 1995). In addition to mTOR, the PIKK family includes other five members: ataxia-telangiectasia mutated (ATM), ataxia- and Rad3-related (ATR) and DNA-dependent protein kinase catalytic subunit (DNA-PKcs) that are involved in DNA damage signaling and DNA repair, suppressor of morphogenesis in genitalia 1 (SMG1) that activates nonsense-mediated mRNA decay (NMD), and transformation/transcription domain-associated protein (TRRAP), a transcriptional regulator (Lempiainen and Halazonetis 2009). PIKKs share a related structural organization including the conserved C-terminal kinase domain (KIN) that is preceded by an array of helical repeats, mostly HEAT (Huntington Elongation factor 3 A subunit of PP2A and TOR1) and TPR (tetratricopeptide) repeats (Fig. 17.1). mTOR signaling assumes a crucial role in sensing and responding to cellular environment cues, serving as a central regulator of cellular growth, proliferation, metabolism, autophagy, and translational control (reviewed in Saxton and Sabatini 2017). Consequently, dysregulation of mTOR signaling profoundly impacts cellular homeostasis, and has been implicated in various pathological conditions, including cancer development and progression, inflammation, diabetes, and neurodegeneration (reviewed in Laplante and Sabatini 2012).

The mTOR kinase domain is structurally stabilized by the interaction with the mammalian lethal with SEC13 protein 8 (mLST8) (Yang et al. 2013), and both proteins constitute the core of two functionally distinct multiprotein complexes named mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2) (Zoncu et al. 2011) (Fig. 17.2). mTORC1, a rapamycin-sensitive complex (Loewith et al. 2002), tightly couples the cellular energy status and nutrient levels to protein synthesis and



Fig. 17.1 The PIKK family. Schematic representation depicting the domains of the PIKK members mTOR, ATR, ATM, SMG1, TRRAP, and DNA-PKcs aligned based on conserved regions. HEAT (Huntington Elongation factor 3 A subunit of PP2A and TOR1), FAT (FRAP, ATM, and TRRAP), and KIN (Kinase) domains are indicated. The FRB (FKBP12 rapamycin binding) and FATC (FAT C-terminal) domains are colored yellow and red, respectively. Insertion region in SMG1 is indicated in gray. The amino acids of each protein are indicated

autophagy, thereby promoting cell growth and proliferation (Hara et al. 1998). mTORC2 only responds partially after long-term treatment with rapamycin (Sarbasov et al. 2006), and its downstream effectors play fundamental roles in regulating cell metabolism, in particular, lipid synthesis, growth factor-induced cell proliferation, cell survival, migration, and insulin signaling.

Both signaling complexes are assembled as dimers containing two copies of mTOR and of the accessory subunits, and only dimeric complexes can be activated to phosphorylate multiple downstream effectors. Signals regulating mTORC1 include amino acids, which activate Rag GTPases to promote mTORC1 localization and activation at the lysosomes (Kim et al. 2008), but also mitogens (Roux et al. 2004), oxygen levels (Brugarolas et al. 2004), and energy status (Gwinn et al. 2008) (Fig. 17.3). In addition, glucose and glutamine regulate mTORC1 through a

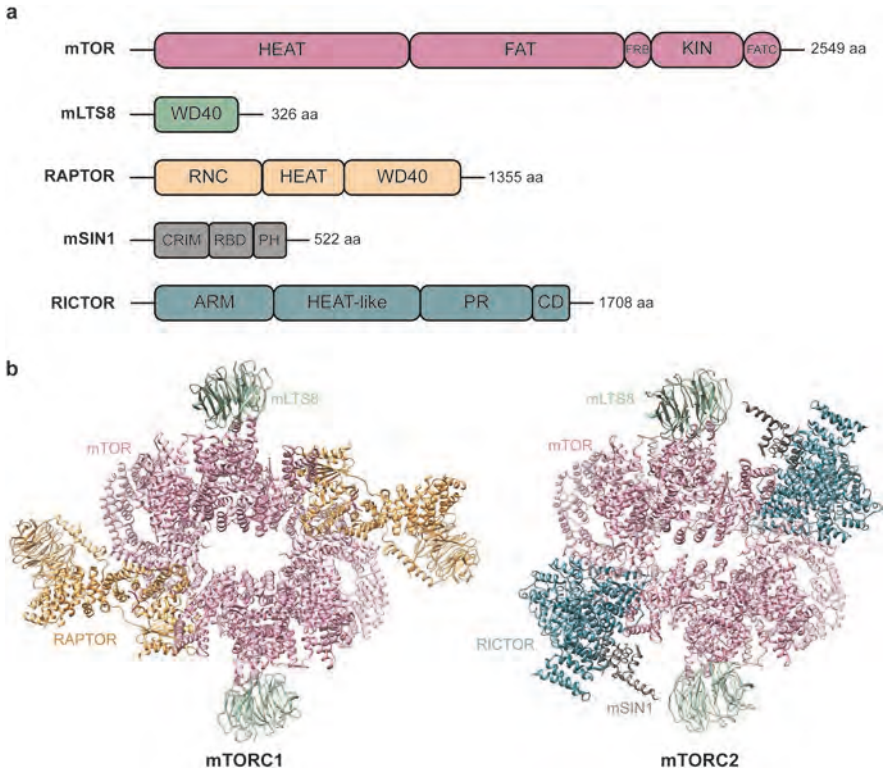


Fig. 17.2 Architecture of mTOR complexes. (a) Domain organization of mTOR complex subunits. W-D 40 repeats (WD40), RAPTOR N-terminal conserved (RNC), conserved region in the middle (CRIM), pleckstrin homology (PH), Armadillo repeat domain (ARM), phosphorylation site region (PR), and C-terminal domain (CD) motifs are indicated. (b) Atomic structure of mTORC1 (left, PDB-6BCX) and mTORC2 (right, PDB-5ZCS). Color codes are as indicated in (a): mTOR in pink, mLTS8 in green, mTORC1-specific subunit RAPTOR in orange, and mTORC2 components RICTOR and mSIN1 in blue and gray, respectively. Both complexes are shown with the same orientation based on mTOR subunit position

mechanism that is independent of the tuberous sclerosis complex (TSC) and the AMP-activated protein kinase (AMPK) signaling (Kim et al. 2013). Deprivation of amino acids in cells promotes mTOR cytoplasmatic localization, while it is located at the lysosome in the presence of amino acids (Sancak et al. 2008). Details of the main pathways regulating mTORC1 and mTORC2 and their downstream effectors have been reviewed extensively elsewhere (Laplane and Sabatini 2012; Kim and Guan 2019; Liu and Sabatini 2020; Napolitano et al. 2022).

Regulatory-associated protein of mTOR (RAPTOR) and rapamycin-insensitive companion of mTOR (RICTOR) are the subunits defining mTORC1 and mTORC2, respectively. mTOR and its obligated accessory protein mLST8 interact with RAPTOR which regulates mTOR catalytic activity and substrate specificity to form mTORC1. This 452 kDa protein complex possesses an intricately arranged multi-subunit architecture elucidated through a combination of cryo-electron microscopy (cryoEM), X-ray crystallography, and functionally assays (Aylett et al. 2016) (Fig. 17.2). mTORC2 is assembled by the interaction of mTOR and mLST8 with RICTOR, and mammalian stress-activated MAP kinase-interacting protein 1 (mSIN1), but in addition, it can contain other constituents that are cell type-specific. Several structural studies of mTORC2 have revealed the way RICTOR and mSIN1 subunits interact in mTORC2 (Fig. 17.2) and the structural basis for the biochemical regulation of the complex (Chen et al. 2018; Scaiola et al. 2020; Yu et al. 2022).

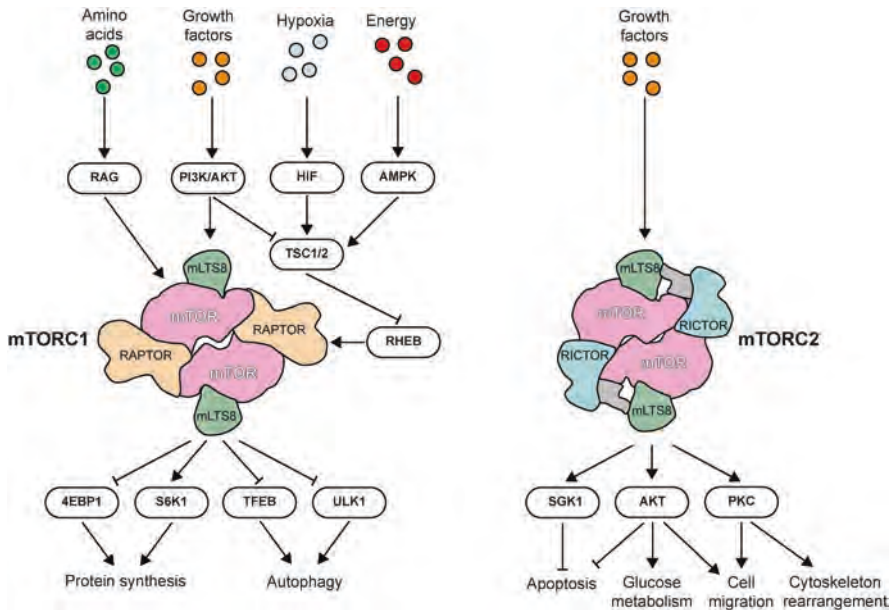


Fig. 17.3 Signaling pathways of mTOR complexes. Simplified scheme of the signals regulating the mTORC1 (left) and mTORC2 (right) complexes signaling cascades. mTORC1 integrates stimuli from oxygen levels, growth factors, amino acids, and energy to activate anabolism and repress catabolism. mTORC2 responds to growth factors to regulate cell metabolism and survival

The Maturation and Assembly of mTOR Complexes Requires the Assistance of Chaperones

The complex structures of mTORC1 and mTORC2 require a sophisticated maturation and assembly pathway. Prior to their assembly, each of the protein constituents must undergo proper folding and in some cases protection from misfolding and aggregation while waiting for the interactions with other components of the mTORC1 and mTORC2 complexes. All this is achieved, thanks to several molecular chaperones and cochaperones. Chaperones are a diverse group of proteins responsible for assisting in protein folding, stability, and quality control within cells, whereas cochaperones are non-client binding proteins that help chaperones, for example, by regulating their ATPase activity.

The chaperonin-containing TCP-1 (CCT) is the eukaryotic homologue of the bacterial heat shock protein 60 (HSP60) chaperones, a family also known as chaperonins. Genetic disruption of the ATPase activity of CCT has been associated with phenotypes similar to those observed in yeast TOR signaling loss (Amit et al. 2010). Moreover, overexpression of CCT in yeast alleviates phenotypes associated with temperature-sensitive mutations of TOR and LST8 (Schmidt et al. 1996; Kabir et al. 2005). Also, human interactome studies reveal interactions of CCT with mLST8 and Raptor but not with mTOR (Huttlin et al. 2017). These and other results have prompted several authors to propose that CCT plays a role in formation of mTORC1 complexes by facilitating the folding of the β -propeller domains of both mTORC1 components, RAPTOR and mLST8 (Cuellar et al. 2019). The maturation of mTORC2, on its part, has been linked to the heat shock protein 70 (HSP70) chaperone which interacts with RICTOR, suggesting that this chaperone could promote proper assembly of these complexes both constitutively and under heat shock conditions (Martin et al. 2008). In addition, recent studies suggest that CCT and members of the HSP family of proteins may have also regulatory functions beyond mere protein folding (Tang et al. 2021; Wang et al. 2022; Horejsi et al. 2010; Takai et al. 2010).

Therefore, CCT and HSP70 which are chaperones that assist in the folding and maturation of many proteins also participate in the folding of several of the proteins that integrate the mTORC1 and mTORC2 complexes. However, the assembly of mTORC1 and mTORC2 also requires an additional set of chaperones and cochaperones that are specific for the mTOR complexes of those of other members of the PIKK family, and whose structure, function, and regulation are significantly more complex than other chaperones. These are the complex formed by the proteins Telomere length regulation protein TEL2 homologue (TELO2), TELO2-interacting protein 1 (TTI1) and TELO2-interacting protein 2 (TTI2), and the HSP90 chaperone working in concert with the R2TP complex (Takai et al. 2010; Kim et al. 2013), and which are described in detail in the following sections.

Discovery of the Telo2-TTI1-TTI2 Complex and Its Function in mTOR Maturation

TELO2 is a protein first described as a factor involved in telomere length maintenance in budding yeast (Lustig and Petes 1986), but subsequent functional studies determined that the mammalian homologue Telo2 was instead implicated in the expression and stability of mTOR (Kano and Yanagida 2007; Takai et al. 2007). Cells deficient in Telo2 show reduced levels of mTOR protein without alteration of the mRNA levels, being mTOR synthesized at normal rates in the absence of Telo2; however, the half-life of the kinase is drastically reduced (Takai et al. 2007). These observations suggested that Telo2 was involved in the stability of mTOR, and several experiments have demonstrated that Telo2 forms a complex with mTOR. Endogenous human mTOR is immunoprecipitated by Telo2 from HeLa and MEFs cells, and the interaction between Telo2 and mTOR seems direct since Telo2 pulls down a fragment of mTOR containing a portion of the HEAT-repeat domains (amino acids 400–806) (Takai et al. 2007).

TELO2 associates preferentially with newly synthesized mTOR. Recovery of newly synthesized mTOR from Telo2-deficient MEFs cells is reduced in immunoprecipitation (IP) experiments using either RAPTOR or RICTOR subunits as bait, suggesting that Telo2 affects the incorporation of mTOR to the mTORC1 and mTORC2 complexes (Takai et al. 2007; Kaizuka et al. 2010). A general role of Telo2 in the maturation of PIKK was further supported by the observation that Telo2 preferentially bound newly synthesized not only mTOR but also ATM, ATR, and DNA-PKcs (Takai et al. 2010).

TELO2 is part of a trimeric complex together with TTI1 and TTI2, which were first found as Telo2 interactors in a proteomic analysis in fission yeast (Hayashi et al. 2007; Shevchenko et al. 2008). In mammals, Telo2 was found to physically associate with TTI1 and TTI2 in genome-wide RNAi screen experiments aimed to investigate IR resistance genes, forming the Telo2-TTI1-TTI2 (TTT) complex (Hurov et al. 2010). In addition to a role in DNA damage response (DDR) in mammals due to the affection of ATM and ATR levels, the TTT complex was found to maintain the abundance of mTOR and the rest of kinases of the PIKK family. TTI1 depletion in U2OS cells causes defects in signaling downstream from mTORC2. TTI1 associates with Telo2 and mTOR in immunoprecipitation experiments, and TTI2 associates with both TTI1 and Telo2 (Hurov et al. 2010; Kaizuka et al. 2010). Depletion for TTI1 reduces the amounts of mTOR-coprecipitated RAPTOR and RICTOR, suggesting the disassembly of both mTORC1 and mTORC2 complexes (Kaizuka et al. 2010). TTI1, TTI2, and Telo2 cofractionate in a large molecular weight complex on size exclusion chromatography of 293 T cells expressing inducible TTI2, further corroborating that they function together as part of a larger multimeric complex (Hurov et al. 2010).

TTT is also part of a complex interacting network with other complexes. The first evidence for this came from IP-mass spectrometry experiments from HeLa S3 cells using each of the TTT subunits as bait, in which Telo2 coprecipitated with TTI1

and TTI2, but also with the multifaceted ATPases RuvB-like 1 (RUVBL1) and RuvB-like 2 (RUVBL2), which contribute to the expression of PIKK by affecting its mRNA levels (Izumi et al. 2010). TELO2 also coprecipitates the heat shock proteins HSP90, HSP70, and HSP40. On other hand, TTI2 coprecipitates the HSP90 chaperone, RNA polymerase II-associated protein 3 (RPAP3), and PIH1 domain-containing protein 1 (PIH1D1). Both RPAP3 and PIH1D1 proteins are well-known interactors with RUVBL1 and RUVBL2, and will be further discussed later (Jeronimo et al. 2007; Boulon et al. 2008; Gano and Simon 2010).

Altogether, the association of TTT with mTOR and molecular chaperones suggested that TTT is necessary for the maturation and assembly of mTOR complexes, and this seemed to occur as part of a pathway involving HSP90 and its cochaperone partners RUVBL1, RUVBL2, RPAP3, and PIH1D1. In fact, inhibition of HSP90 diminished the association of TELO2 to mTOR and other PIKKs in HeLa cells (Horejsi et al. 2010), without affecting protein levels (Takai et al. 2007). It has also been shown that Tra1 and Tra2 in *S. pombe*, homologues of the PIKK family member TRRAP, require HSP90 and TTT for their incorporation into the SAGA and NuA4 complexes (Elias-Villalobos et al. 2019).

TELO2-TT1-TTI2 Assists mTOR Cotranslationally

Recent work by Toullec and colleagues has clarified the main role of TTT during the maturation of mTOR complexes using *S. pombe* as model organism (Toullec et al. 2021). The TTT complex assists in the folding and maturation of mTOR and all other PIKKs rather than a direct function in the assembly of mTORC1 and mTORC2. This occurs cotranslationally, thereby temporarily separating mTOR folding from its assembly into a larger complex.

By using tandem affinity purification of endogenous TELO2, TTI1, and TTI2 coupled to quantitative mass spectrometry (TAP-MASS), they revealed the formation of a stable complex of TELO2, TTI1, and TTI2, which interacts with mTOR and other PIKK. Interestingly, the interaction of the TTT components with RUVBL1 and RUVBL2 is weak or transient, and no interaction with HSP90 is detected in their experimental approach. Both steady-state levels and stability of the *S. pombe* mTOR homologue Tor1 (together with TRAPP homologues Tra1 and Tra2) are decreased in TTI2-conditional knockout (CKO) strains (Toullec et al. 2021).

Tra1 mRNA levels are enriched in RNA immunoprecipitation (RIP) experiments of TTT subunits, an effect abolished by treatment with ribosome-dissociating agents, which supported a model where TTT binds the PIKK-encoding mRNA and recognizes nascent PIKK polypeptides (Toullec et al. 2021). TTI2 was proposed to be the subunit of the TTT complex that binds to the nascent polypeptides since the interaction of TTT with Tra1 decreases upon TTI2 depletion but not in the absence of TELO2 (Toullec et al. 2021). A structure-function study using mutants of yeast Tra1 and Tor1 revealed that their FAT and KINASE domains (Fig. 17.1) are essential and sufficient for binding to TTI2, whereas the PIKK HEAT-repeat domain

seemed dispensable. Furthermore, the FAT and KINASE domains interact with TTI2 cotranslationally in RIP experiments using these truncated PIKK mutants.

Altogether, Toullec and colleagues proposed that nascent polypeptide chains of mTOR and other PIKKs are recognized by the TTI2 subunit of the TTT complex once the FAT and KINASE domains located at the C-terminus are translated (Toullec et al. 2021). This would explain how TTT can assist in the folding of all the members of the PIKK family since the structural unit formed by the FAT and KINASE domains is highly conserved among the PIKKs whereas major differences are located in the extension and structure of the HEAT repeat region at the N-terminus (Imseng et al. 2018) (Fig. 17.1). TTT would stabilize the nascent PIKKs in a non-native state during translation of the FAT and KINASE domains until the last residues are synthesized by the ribosome, defining the TTT as a bona fide cotranslational chaperone of PIKKs.

The functions of TTT seem to be restricted to the folding and stability of mTOR (and the PIKKs) but not to other subunits of mTORC1 and mTORC2. Subunits Mip1 (RAPTOR homologue in *S. pombe*) and Ste20 (RICTOR homologue) were not found to interact with TTT cotranslationally (Toullec et al. 2021). In agreement, TTT associates with PIKKs but not with their partners in mammals (Takai et al. 2007, 2010). All this argues for a model where the TTT complex serves to fold and stabilize mTOR during translation, prior to the transfer to other chaperone complexes dedicated to the promotion of interaction with additional partners that comprise mTORC1 and mTORC2.

Structure of the TELO2-TT1-TTI2 Complex

Besides biochemical evidence of cellular associations and direct interactions of the TTT subunits with mTOR, structural information of some of these interactions and of the TTT complex have contributed to our understanding of the molecular mechanism involved in the maturation of this kinase by TTT. The three constituents of the TTT complex are composed mostly by helical repeats folded into α -solenoid structures (Fig. 17.4a), and its architecture has been recently determined by a combination of cryoEM and interaction mapping experiments (Pal et al. 2021; Kim et al. 2022) (Fig. 17.4b). The TTI1 subunit is the largest, forming a curved helical domain placed at the center of the complex and interacting with the N-terminal domain (NTD) of TELO2 through its central region (middle domain, MD), and with TTI2 through its C-terminal segment (CTD) (Kim et al. 2022). TTI1 interacts with the FAT domain and the HEAT repeats of ATM in yeast-two-hybrid experiments (Kim et al. 2022), whereas TTI2 was found to bind with PIKK kinases during cotranslational folding (Toullec et al. 2021), suggesting that the extended helical structure formed by TTI1 and TTI2 could be involved in the association with hydrophobic regions of the PIKKs exposed during its translation and folding.

The atomic structure of Tel2, the yeast homologue of TELO2, shows a protein composed of two helical solenoids, NTD and CTD, connected by a flexible linker.

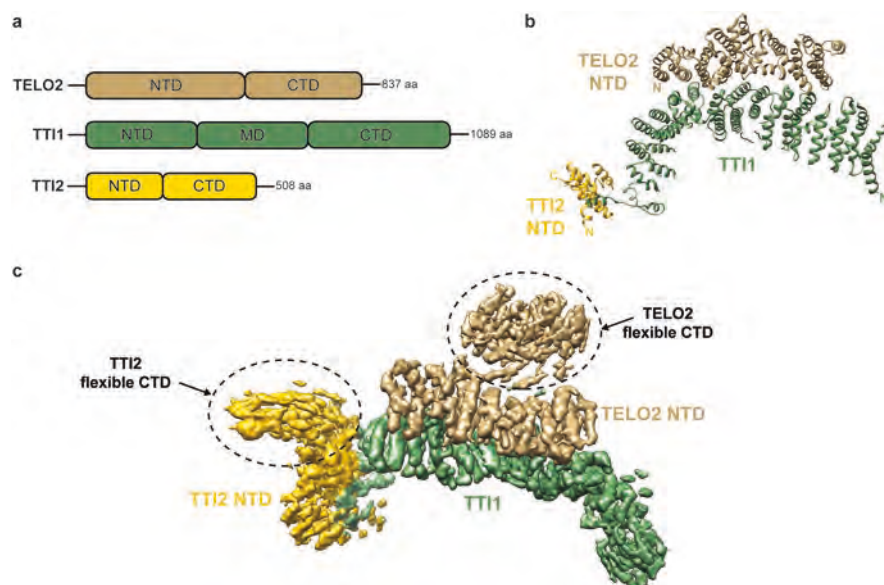


Fig. 17.4 Structure of the TTT complex. **(a)** TELO2 (brown), TTI1 (green) and TTI2 (yellow) representation showing amino acids and domain organization (NTD, N-terminal domain; MD, middle domain; CTD, C-terminal domain). **(b)** Atomic structure of the TTT complex (PDB-7F4U). TTI1 interacts with the NTD of TELO2 through its MD helical repeats, and with the NTD of TTI2 through the CTD. **(c)** CryoEM structure of the TTT complex (EMD-21454) used for the atomic model building in **(b)** showing additional flexible regions corresponding to TELO2 CTD and TTI1 CTD. The model is tilted 60° respect to the view shown in **(b)**

In the cryoEM structures of the TTT complex, it is the larger NTD of TELO2 which interacts with TTI1, whereas the CTD remains flexibly attached, hindering the determination of its atomic structure in the complex (Pal et al. 2021; Kim et al. 2022). However, some cryoEM maps revealed density compatible with this C-terminal domain in several positions in the vicinity of TELO2 NTD (Kim et al. 2022) (Fig. 17.4c). The linker between the two domains of TELO2 harbors a consensus sequence containing two serine residues (Ser487 and Ser491) that are constitutively phosphorylated by CK2 *in vivo* (Takai et al. 2010; Horejsi et al. 2010). These phosphosites serve as an anchoring motif for the N-terminal protein interacting with Hsp90 1 (PIH1) domain of PIH1D1, one of the components of R2TP. Crystal structures revealed that PIH-N forms a complex with the TELO2 phosphopeptide and that the conserved DpSDD motif in TELO2 is crucial for binding (Horejsi et al. 2014; Pal et al. 2014).

The R2TP Complex and HSP90 Cooperate to Assemble mTOR Complexes

Beyond the folding of mTOR and its partners, the assembly of functional mTORC1 and mTORC2 complexes needs the interaction of mTOR with its partners and dimerization. This requires a dedicated chaperone machinery integrated by the R2TP complex and HSP90 (Takai et al. 2007; Zhao and Houry 2007). Whereas TTT seems to work specifically for the maturation of the PIKK kinases (Toullec et al. 2021), the R2TP complex cooperates with HSP90 in the maturation not only of the PIKKs but also of several other large complexes such as spliceosomal U4 and U5 snRNPs, snoRNP complexes of the L7Ae family such as telomerase, RNA polymerase II, and DNA repair-sensor MRN complex (Lynham and Houry 2022). It is believed that for each family of clients, specific factors and adaptors exist to regulate R2TP and to contribute to transfer each client to R2TP (Vaughan 2014; Rodriguez and Llorca 2020). For instance, the URI prefoldin-like complex participates in the assembly of RNA polymerase II and other RNA polymerases in cooperation with HSP90 and R2TP, but not in other processes such the assembly of mTOR complexes (Boulon et al. 2010).

Some of the early evidence connecting the activity of R2TP with the maturation of PIKKs comes from studies showing that RUVBL1 coprecipitates all the six members of the PIKK family in IP experiments, including mTOR (Izumi et al. 2010). Subsequently it was found that treatment of HeLa cells with the HSP90 inhibitor 17-AAG promotes a reduction in the amount of mTOR protein and a reduction in mTOR-mediated phosphorylation of downstream targets (Izumi et al. 2012). Such an effect correlates with the physical association of HSP90, RUVBL1, RUVBL2, and TELO2 in cells. HSP90 copurifies with RUVBL1, while TELO2 coprecipitates with HSP90, together with the TTI1 and TTI2 subunits. Moreover, TELO2 and both TTI1 and TTI2 copurify with RUVBL1-RUVBL2 and with HSP90. In addition, the RPAP3 and PIH1D1 subunits from the R2TP complex, as well as RNA polymerase II subunit B5 (RBP5) and unconventional prefoldin RPB5-interactor 1 (URI1) from the URI1 prefoldin-like complex, coimmunoprecipitate with RUVBL1 and with TELO2, TTI1, and TTI2 (Izumi et al. 2012). Such an intricate network of protein associations suggests the formation of complexes containing HSP90, RUVBL1-RUVBL2, TTT, RPAP3, PIH1D1, and URI1. Consistently with a regulatory effect on mTOR, knockout of RUVBL1, RUVBL2, or TELO2 decreased the amount of the PIKK, while the knockout of the associated proteins TTI1, TTI2, PIH1D1, and URI1 only affected the PIKK signaling pathway (Izumi et al. 2012).

The R2TP complex is formed by the interaction of four proteins, the RUVBL1 and RUVBL2 ATPases which form a heterohexameric ring (Fig. 17.5a, b), RPAP3, and PIH1D1 (Fig. 17.5c). This RUVBL1-RUVBL2-RPAP3-PIH1D1 complex was named R2TP after the yeast homologue proteins Rvb2-Rvb2-Tah1-Pih1 (Zhao et al. 2005). RUVBL1 and RUVBL2 proteins are AAA+ ATPases first described as homologues of bacterial RuvB (Kanemaki et al. 1997) that are essential for cell

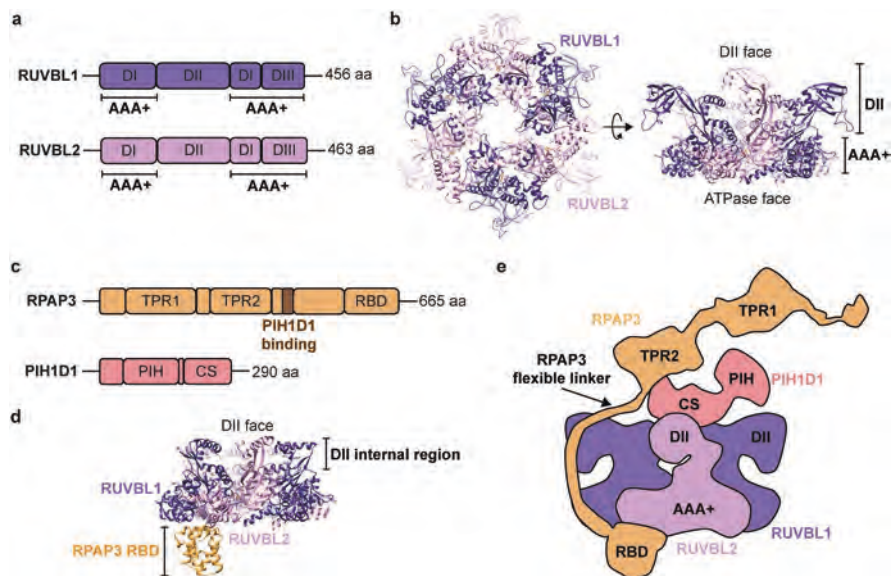


Fig. 17.5 RUVBL1-RUVBL2 and R2TP architecture. **(a)** Sequence scheme of RUVBL1 (purple) and RUVBL2 (pink) ATPases indicating representative domains and amino acids. **(b)** Atomic structure of the RUVBL1-RUVBL2 heterohexamer (generated from PDB-2C9O and PDB 6H7X) indicating the catalytic AAA+ core and the DII domain surface. Two views, top and side, are shown using color codes as in **(a)**. The ATPase and the DII faces of the hexameric ring are indicated. **(c)** RPAP3 (orange) and PIH1D1 (red) domains scheme and amino acids. **(d)** Atomic structure of the human R2TP complex (PDB-6FO1) showing the AAA+ core of RUVBL1-RUVBL2 bound to one RBD domain of RPAP3. Note that the OB-fold domains of RUVBL1 and RUVBL2 subunits are not solved in this atomic model due to its high flexibility. The DII face of the RUVBL1-RUVBL2 is indicated. **(e)** Full-length R2TP cartoon representation. RPAP3 subunit anchors to the AAA+ side of the RUVBL1-RUVBL2 hexamer through its specialized RBD domain, flexible connected to the TPR domains which are oriented to the DII interface. PIH1D1 is engaged by the interaction with a DII domain and a region of RPAP3 near the TPR2

viability, regulating a plethora of cellular process including chromatin architecture dynamics, cell division, replication, transcription, or genomic integrity maintenance, among others (reviewed in Dauden et al. 2021).

The R2TP subunit RPAP3 contains two TPR domains that can recognize the MEEVD motif of HSP90, supporting that R2TP functions in concert with HSP90 (Pal et al. 2014; Morgan et al. 2015). HSP90 is a conserved molecular chaperone that binds client proteins at later stages of their folding and promotes their stability, maturation, and assembly into larger complexes, and its activity usually requires accessory proteins or cochaperones (Young et al. 2004; Biebl and Buchner 2019). RPAP3 also contains a RUVBL2-binding domain (RBD) at the C-terminus that serves to engage RUVBL2 (Martino et al. 2018). PIH1D1 is composed by the N-terminal PIH1 domain, mentioned above when describing that this domain recognizes the conserved DpSDD motif in TELO2, and a CHORD-containing protein

and SGT1 domain (CS) domain that interacts with RPAP3 (Horejsi et al. 2010; Pal et al. 2014) and also with the DII domains of RUVBL2 (Munoz-Hernandez et al. 2019) (Fig. 17.5c, d).

Architecture of the R2TP Complex

R2TP is considered one of the most sophisticated HSP90 cochaperones yet described, and its structure is organized around a heterohexameric ring of the RUVBL1 and RUVBL2 AAA+ ATPases, which serves as a scaffold for the rest of subunits. Domains 1 (DI) and 3 (DIII) of RUVBL1 and RUVBL2 form a ring of alternating subunits and harboring the catalytic activity. Domain 2 (DII) appears between the AAA+ domains in the sequence and is organized as a protruding oligonucleotide/oligosaccharide-binding (OB)-fold in the structure (Fig. 17.5a, b). Thus, six DII domains, three from each type of subunit, project from the hexamer and define two faces of the ring, the ATPase-face and the DII face, which seems to have different functions (Fig. 17.5b). DII domains are involved in protein-protein interactions, being the DII face of the ring where most R2TP partners interact (Matias et al. 2006; Gorynia et al. 2011; Lopez-Perrote et al. 2012; Ewens et al. 2016; Silva-Martin et al. 2016; Silva et al. 2018). In fact, in all other complexes containing RUVBL1-RUVBL2 but which are unrelated to R2TP proteins and partners are found to occupy the DII face of the ring (Ayala et al. 2018; Lopez-Perrote et al. 2020; Serna et al. 2022).

The architecture of the complete human R2TP complex was first defined by cryoEM (Martino et al. 2018). The structure revealed that RUVBL1 and RUVBL2 form the core of the R2TP complex where the rest of components attach (Fig. 17.5d, e). RPAP3 interacts with RUVBL2 through its C-terminal alpha helical RBD domain, being the only RPAP3 domain required for RUVBL1-RUVBL2 binding. Interestingly, this interaction happens at the ATPase face of the RUVBL1-RUVBL2 ring (Fig. 17.5d), which has been the only protein/domain yet described that can recognize that face of the ring. Further structural characterization of this specialized RBD domain on RPAP3 was provided by NMR experiments (Maurizy et al. 2018). This domain is present in other proteins, including an alternative isoform of RPAP3 which lacks the PIH1D1-binding motif (RPAP3 isoform 2), sperm-associated antigen 1 (SPAG1), and coiled-coil domain-containing protein 103 (CCDC103), and this has been used to describe a list of R2TP-like complexes that might participate in dynein arm assembly (Maurizy et al. 2018; Rodriguez and Llorca 2020). Cross-linking mass spectrometry and cryoEM showed that RPAP3 then runs outside the edge of the RUVBL1-RUVBL2 toward the DII face of the ring (Martino et al. 2018) (Fig. 17.5e). The PIH1D1 subunit engages with the DII domains of RUVBL1-RUVBL2 and, by interacting with RPAP3, brings the RPAP3 N-terminal end to the DII face of RUVBL1-RUVBL2. Such architecture flexibly positions the HSP90-binding domains TPR in the DII face of the complex. As mentioned, PIH1D1 contains a PIH1 domain which recognizes the phosphorylated version of consensus

motif in several clients including TELO2 (Horejsi et al. 2014; Pal et al. 2014). The structure of the CS domain in complex with a fragment of RPAP3 comprising the most C-terminal TPR domain showed an exclusive insertion region in RPAP3 isoform 1 (here referred to as RPAP3) essential for PIH1D1 binding, and established a PIH1D1-binding motif (Henri et al. 2018). On the other hand, this PIH1D1 CS domain also interacts with the DII domain of RUVBL2, and this interaction regulates the catalytic activity of the ATPases by promoting a conformational rearrangement of an N-terminal segment of the RUVBL2 subunits that acts as a gatekeeper to regulate the accessibility to the nucleotide binding site (Munoz-Hernandez et al. 2019). The PIH1 domain of PIH1D1 would be located at the center of the DII face where it can interact with TELO2 through the consensus sequence (Fig. 17.5e).

All these interactions engage the TPR domains of RPAP3 at the DII face RUVBL1-RUVBL2 ring; however, cryoEM of human R2TP revealed that this region is very flexible, which is likely due to the long linkers connecting different domains in RPAP3. It has been proposed that this flexibility could be functionally relevant to facilitate recruitment of HSP90 and the clients with the sufficient conformational flexibility to manipulate the several members of the PIKK family with related but quite different structures (Martino et al. 2018).

The functionally relevant stoichiometry of the R2TP complex in cells remains to be determined. In vitro, each RUVBL2 subunit can interact with one RPAP3 molecule, and therefore each R2TP complex can contain up to three copies of RPAP3 and three copies of PIH1D1 (Martino et al. 2018). However, cryoEM suggests that even with three copies of RPAP3-PIH1D1, only one PIH1D1 molecule engages with the DII domain at a time (Munoz-Hernandez et al. 2019). This could be due to steric limitations as only one PIH1D1 could fit at the center of the DII face of RUVBL1-RUVBL2. This could imply that maybe only one copy of RPAP3-PIH1D1 is actually needed in the complex. Alternatively, it is also possible that several RPAP3-PIH1D1 heterodimers, which would be flexibly tethered to the RUVBL1-RUVBL2 ring, are needed, but only one actually engages with the DII domains at each time to regulate ATP hydrolysis. Actually, the idea that R2TP could engage two mTOR molecules by two HSP90 dimers held by two RPAP3 molecules make sense if mTOR dimerization must happen within one R2TP complex.

TTT Interacts with RUVBL1-RUVBL2 and Transfers mTOR to R2TP

The binding of TTT to mTOR and other PIKKs has been characterized (Kanoh and Yanagida 2007; Takai et al. 2007, 2010; Kaizuka et al. 2010). In vitro interaction experiments using the yeast R2TP, TTT complex, and an mTORC1-mLST8 complex from *Kluyveromyces marxianus* have shown that the TTI1-TTI2 module of the complex binds mTOR and RUVBL1-RUVBL2 (Pal et al. 2021). Using purified human proteins, R2TP coprecipitates mTOR only when TTT is present supporting

that the TTT complex serves to bridge R2TP and mTOR (Pal et al. 2021). These results reinforce a model where TTI1-TTI2 and not TELO2 binds mTOR and RUVBL1-RUVBL2 to help transfer the kinase from the TTT to R2TP, whereas TELO2 could serve mostly a regulatory function.

After cotranslational folding of mTOR in complex with TTT, the kinase must be transferred to R2TP for further processing and mTORC1/mTORC2 assembly. The predominant model proposes that this occurs by the formation of a complex between mTOR-loaded TTT and the R2TP, and suggests that one of the functions of TTT would be to serve as an adaptor that can bring PIKKs specifically to R2TP. The interaction between PIH1D1 and TELO2 was thought to be the most relevant determinant for the formation of a complex between TTT and R2TP. The CK2-dependent TELO2 phosphosites were not essential for the association of TTT with mTOR or any of the PIKKs with TTT, but they were essential for mTOR in cells. A significant reduction in mTOR levels occurs in conditional TELO2-KO MEFs that is rescued by expression of TELO2 WT but not by TELO2 versions mutated in the phosphosite residues (Horejsi et al. 2010). Therefore, PIH1D1 would serve in part to bring TTT and mTOR to the R2TP complex. The cryoEM map of R2TP in complex with TTT shows that the C-terminal domain of TELO2 and the flexible linker containing the consensus sequence occupies a region in the vicinity of PIH1D1 compatible with the formation of this complex in the DII face of the RUVBL1-RUVBL2 ring (Pal et al. 2021).

However, recent structural studies of the human R2TP-TTT complex have challenged this model after revealing that TTT also forms a complex with RUVBL1-RUVBL2 in the absence of RPAP3-PIH1D1 (Fig. 17.6a), thereby suggesting that mTOR could be transferred from TTT to R2TP using an alternative or complementary mechanism to the recognition of a TELO2 phosphosite by PIH1D1 (Pal et al. 2021). In the structure of RUVBL1-RUVBL2-TTT, the TTT complex is recognized by the OB-fold motifs of the DII domains from two consecutive RUVBL1 and RUVBL2 subunits, and this allows for up to three TTT complexes bound per RUVBL1-RUVBL2 ring under saturating conditions *in vitro* (Fig. 17.6b). In agreement, truncation of the DII domains in RUVBL1 and RUVBL2 abolishes the interaction with the TTT complex (Pal et al. 2021). All these results suggest that the need for phosphorylation of TELO2 for mTORC1 and mTORC2 assembly could be maybe unrelated with a function of recruitment of TTT to R2TP. The interaction of PIH1D1 with the phosphosites of TELO2 could contribute as an additional or alternative anchoring point to the machinery or as part of the regulation of some of the transient steps of a complex and dynamic process (Horejsi et al. 2010, 2014; Pal et al. 2014). An intricate regulation is suggested by the structure of the R2TP-TTT complex that shows that the interaction of TTT with the DII domains of RUVBL1-RUVBL2 is mutually exclusive with the interaction of PIH1D1 with the DII domains (Pal et al. 2021). Also, the interaction of TTT with RUVBL1-RUVBL2 downregulates the ATPase activity of RUVBL1-RUVBL2 (Pal et al. 2021). Although TELO2 does not participate in the interaction of TTT with RUVBL1-RUVBL2, the architecture of the complex places the phosphosite-containing linker between the NTD

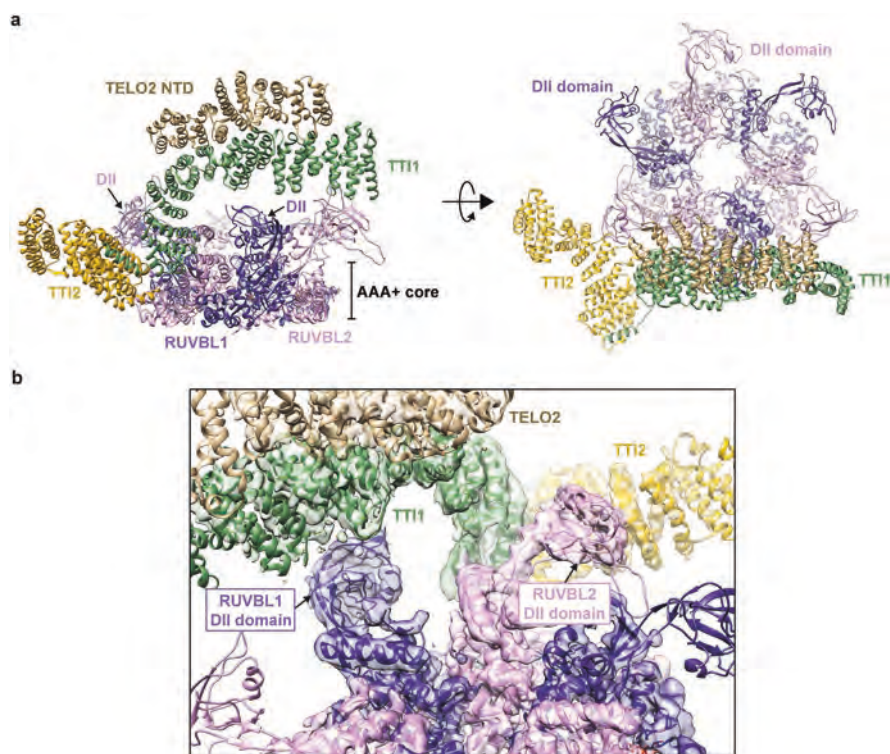


Fig. 17.6 The RUVBL1-RUVBL2-TTT complex. (a) Two views (side and top) of the atomic structure of RUVBL1-RUVBL2 hexamer bound to the TTT complex (PDB-7OLE). (b) A close-up view of the interaction surface between RUVBL1-RUVBL2 and TTT. The RUVBL1-RUVBL2 engages the TTT complex by interacting with TTI1 through the DII domain of one RUVBL1 subunit and with TTI2 through the DII domain of an adjacent RUVBL2 subunit. Atomic structure is shown fitted in the cryoEM map (EMD-12979). Color codes are the same as in Figs. 17.4 and 17.5

and CTD of TELO2 within the DII face of the hexameric ring in a region accessible for binding to PIH1D1 (Pal et al. 2021).

A Model of mTORC1 and mTORC2 Assembly by R2TP-TTT

All the research described in this review, as well as many other works by several authors, have elucidated some of the mechanistic insights of this elaborated pathway of chaperones required for the folding, maturation, and activation of mTOR complexes. Based on all the available evidence, we propose the following model which integrates some of the most accepted views in the field (Fig. 17.7).

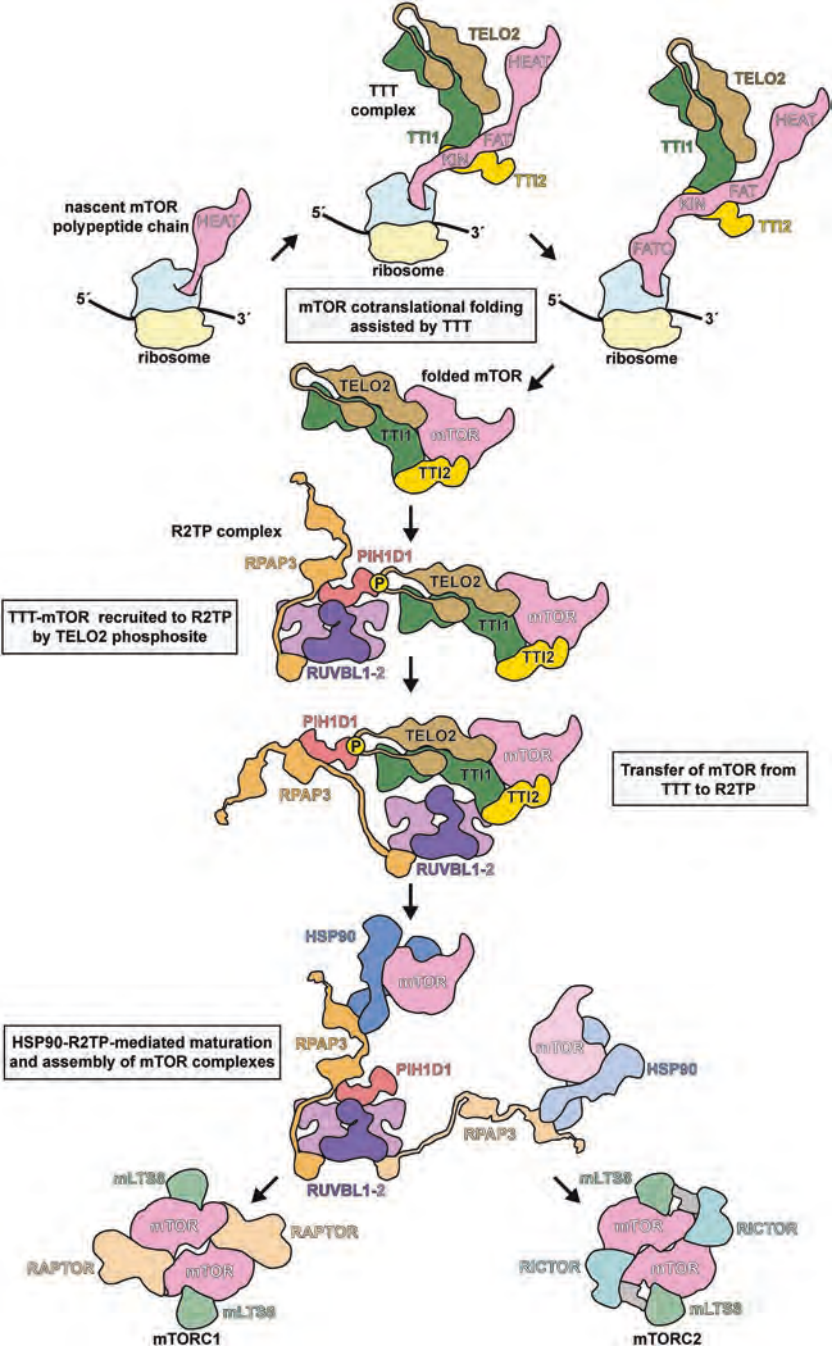


Fig. 17.7 Integrative model of mTOR complex maturation. During the translation of the mTOR protein, the TTT complex assists the process by the recognition of the FAT and KINASE (KIN)

The TTT complex assists the cotranslational folding of mTOR by forming a direct complex. The TTI2 subunit interacts with the mTOR FAT and KINASE domains, which protects this module from misfolding until protein translation is completed (Toullec et al. 2021). The FAT and KINASE domains are well conserved among the PIKKs, rationalizing why TTT assists the folding of every member of the PIKK family (Toullec et al. 2021).

Once mTOR is fully translated, the HSP90 chaperone facilitates the assembly of mTORC1 and mTORC2 complexes working in concert with the R2TP complex. This involves the delivery of mTOR from TTT to R2TP which presumably happens through the formation of an R2TP-TTT complex. The mTOR-loaded TTT complex could interact with the R2TP cochaperone through two mechanisms. On one hand, the PIH1D1 subunit of R2TP recognizes a CK2-phosphosite in TELO2, which provides an anchoring point between the two complexes (Horejsi et al. 2010, 2014; Pal et al. 2014). In addition, the TTI1 and TTI2 subunits interact with the DII domains of two adjacent subunits of RUVBL1-RUVBL2 heterohexameric ring, with an orientation that exposes the phosphosite-containing region of TELO2 to the central section of the DII face of RUVBL1-RUVBL2 accessible for PIH1D1 binding (Pal et al. 2021). Given that TTT can interact directly with RUVBL1-RUVBL2, it is possible that the interaction between PIH1D1 and the CK2-phosphosite in TELO2 could serve other regulatory functions in addition or instead of serving as a way to recruit TTT to the R2TP.

Interestingly RUVBL1-RUVBL2 provides a flexible anchoring for several of the proteins that bind to the heterohexameric ring, and this could be functionally relevant to facilitate the multiple interactions that need to occur in order to assemble the mTOR complexes. RPAP3 is flexibly tethered to RUVBL1-RUVBL2, and this will affect HSP90 when bound to RPAP3 TPR domains (Martino et al. 2018). The DII domains, where other proteins such as PIH1D1 and the TTT complex interact, are also flexible, and they change their conformation upon interaction with proteins.

The RPAP3-PIH1D1 heterodimer regulates nucleotide exchange in RUVBL1-RUVBL2 (Munoz-Hernandez et al. 2019), and presumably regulates its ATPase activity as described for the yeast R2TP counterpart (Rivera-Calzada et al. 2017). The TTT complex downregulates RUVBL1-RUVBL2 ATPase activity by stabilizing a conformation that possibly hampers the nucleotide exchange required for ATP hydrolysis. RPAP3-PIH1D1 and TTT might both regulate the ATPase activity in the R2TP complex during the assembly of mTOR complexes, which must include

Fig. 17.7 (continued) domains mainly through the TTI2 subunit, stabilizing the nascent polypeptide chain and preventing aggregation or misfolding. Once the most C-terminal domain FACT is translated, the folded mTOR protein is released from the ribosome. Presumably, the mTOR-TTT complex is recruited to the HSP90-R2TP machinery by the recognition of a phosphosite in TELO2 by the PIH1D1 subunit, allowing for the further transfer of mTOR from the TTT complex to R2TP in a process that might be controlled by the modulation of the ATPase activity of RUVBL1-RUVBL2. Recruitment of HSP90 by the TPR domains of RPAP3 favors the interaction and processing of mTOR by the chaperone. mTOR dimerization and recruitment of accessory subunits through a less-well understood mechanism must occur to assemble the active mTORC1 and mTORC2 complexes

several steps such loading of mTOR and the partners, dimerization, and the release of the fully assembled complexes (Pal et al. 2021). The details of all these steps are still not understood. Each R2TP complex might manipulate one mTOR molecule and its binding to other partners, but it is also possible that each R2TP, which can comprise more than one RPAP3-HSP90 complex, could also deal simultaneously with two mTOR molecules to promote dimerization (Fig. 17.7).

The R2TP-TTT Chaperones Modulate mTORC1

There is accumulating evidence that the regulation of the R2TP-TTT chaperone system can serve cells to modulate mTORC1 activity. For example, it has been proposed that Wacky, the *Drosophila* homologue of the human protein WW domain-containing adaptor protein with coiled-coil (WAC), regulates the interaction between TTT and R2TP and this, in turn, regulates mTORC1 activity. Wacky interacts with mTORC1 and components of the TTT-R2TP cochaperone system in an energy-dependent manner, and its depletion results in impaired mTOR signaling (David-Morrison et al. 2016). While the precise mechanism remains to be elucidated, the authors interpreted their results as suggesting that Wacky would regulate the interaction between TTT and R2TP, and this would serve to regulate the assembly and activity of mTORC1 complexes (David-Morrison et al. 2016).

Interestingly, a chemical library screening identified piperlongumine (PL) to specifically promote a cytotoxic effect in cancer cells with high mTORC1 activity, and this molecule functioned by disrupting the interaction between the TTT complex and RUVBL1-RUVBL2 (Shin et al. 2020). PL interacts with RUVBL1-RUVBL2 in IP experiments of cells treated with biotin-labeled PL. Moreover, PL reduces the levels of TTI1 and RUVBL2 coprecipitated by TELO2, indicating that this compound affects the interaction between RUVBL1-RUVBL2 and the TTT complex (Shin et al. 2020). Knockdown of RUVBL1-RUVBL2 induces death of cells with mTORC1 hyperactivation, together with a downregulation of the levels of counterpart proteins TELO2, TTI1, TTI2, mTOR, and other PIKK. Similarly, knockdown of TELO2 promotes cell death. Overexpression of RUVBL2 and components of the TTT complex in cancer cells correlated with high activity of mTORC1, highlighting a compensatory mechanism in which the hyperactivation of mTORC1 in oncogenic processes requires an increment of the regulatory complex RUVBL1-RUVBL2-TTT (Shin et al. 2020).

We hypothesize that the regulation of the R2TP-TTT chaperones system either by cells or by drugs can serve to modulate the activity of mTORC1. Reinforcing this hypothesis, several reports have linked the regulation of the R2TP-TTT chaperones system in response to deprivation of specific amino acids with the downmodulation of TORC1. The interaction of TTT with RUVBL1-RUVBL2 is disrupted after the simultaneous deprivation of glucose and glutamine, and it was proposed that this reduced the amount of mTORC1 assembly and also inhibited the location of mTORC1 at the lysosome that is required for activation (Kim et al. 2013). The TTT

complex could regulate mTORC1 also by other mechanisms since it has been described that the MAPK-interacting kinase (MNK) contributes to mTORC1 signaling by a pathway that requires TELO2 (Brown and Gromeier 2017).

Conclusions and Outlook

The maturation and assembly of mTORC1 and mTORC2 is a complex cellular pathway involving an elaborated network of chaperones, the TTT and R2TP complexes, working together with HSP90 and regulated by ATP hydrolysis and various other proteins. We are just starting to unravel the mechanistic details of this complex. Major questions remained to be addressed, such as how mTOR and its partners are loaded to R2TP, how R2TP manipulates all these clients to promote the formation of a fully assembled dimeric mTORC1/mTORC2 working in concert with HSP90, how the assembled complex is released from the chaperone and cochaperone, and what is the role of ATP hydrolysis in all these steps. In addition, we are just starting to study how the cells make use of this pathway to regulate mTORC1 activity in response to some nutrient conditions. Further understanding will require the contributions of biochemistry, molecular cell biology, and structural biology, especially the use of cryoEM to determine the structure of these large complexes.

The R2TP and TTT chaperones are important players in cancer. For instance, studies on hepatocellular carcinoma (HCC) patients have identified a downregulation of the neural calcium sensor protein hippocalcin-like 1 (HPCAL1) (Zhang et al. 2016). HPCAL1-deficient tumors show a reprogramming of lipid metabolism which is related to hyperactivation of mTOR (Chen et al. 2022). Cellular experiments showed that HPCAL1 interacts with RUVBL1 in cellular experiments and mice liver tumors, and inhibits the assembly of the TTT complex, thus inhibiting mTOR activation. This, together with the evidence that components of the R2TP complex can be targeted to disrupt cancer cells (Assimon et al. 2019; Nano et al. 2020; Ju et al. 2020; Kakiyama et al. 2020), opens the way to target these chaperones with potential benefits in cancer.

References

- Amit M, Weisberg SJ, Nadler-Holly M, McCormack EA, Feldmesser E, Kaganovich D, Willison KR, Horovitz A (2010) Equivalent mutations in the eight subunits of the chaperonin CCT produce dramatically different cellular and gene expression phenotypes. *J Mol Biol* 401(3):532–543. <https://doi.org/10.1016/j.jmb.2010.06.037>
- Assimon VA, Tang Y, Vargas JD, Lee GJ, Wu ZY, Lou K, Yao B, Menon MK, Pios A, Perez KC, Madriaga A, Buchowiecki PK, Rolfe M, Shawver L, Jiao X, Le Moigne R, Zhou HJ, Anderson DJ (2019) CB-6644 is a selective inhibitor of the RUVBL1/2 complex with anticancer activity. *ACS Chem Biol* 14(2):236–244. <https://doi.org/10.1021/acscchembio.8b00904>

- Ayala R, Willhoft O, Aramayo RJ, Wilkinson M, McCormack EA, Ocloo L, Wigley DB, Zhang X (2018) Structure and regulation of the human INO80-nucleosome complex. *Nature* 556(7701):391–395. <https://doi.org/10.1038/s41586-018-0021-6>
- Aylett CH, Sauer E, Imseng S, Boehringer D, Hall MN, Ban N, Maier T (2016) Architecture of human mTOR complex 1. *Science* 351(6268):48–52. <https://doi.org/10.1126/science.aaa3870>
- Biebl MM, Buchner J (2019) Structure, function, and regulation of the Hsp90 machinery. *Cold Spring Harb Perspect Biol* 11(9). <https://doi.org/10.1101/cshperspect.a034017>
- Boulon S, Marmier-Gourrier N, Pradet-Balade B, Wurth L, Verheggen C, Jady BE, Rothe B, Pescia C, Robert MC, Kiss T, Bardoni B, Krol A, Branlant C, Allmang C, Bertrand E, Charpentier B (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol* 180(3):579–595. <https://doi.org/10.1083/jcb.200708110>
- Boulon S, Pradet-Balade B, Verheggen C, Molle D, Boireau S, Georgieva M, Azzag K, Robert MC, Ahmad Y, Neel H, Lamond AI, Bertrand E (2010) HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol Cell* 39(6):912–924. <https://doi.org/10.1016/j.molcel.2010.08.023>
- Brown MC, Gromeier M (2017) MNK controls mTORC1:substrate association through regulation of TELO2 binding with mTORC1. *Cell Rep* 18(6):1444–1457. <https://doi.org/10.1016/j.celrep.2017.01.023>
- Brugarolas J, Lei K, Hurley RL, Manning BD, Reiling JH, Hafen E, Witters LA, Ellisen LW, Kaelin WG Jr (2004) Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. *Genes Dev* 18(23):2893–2904. <https://doi.org/10.1101/gad.1256804>
- Chen T, Yuan Z, Lei Z, Duan J, Xue J, Lu T, Yan G, Zhang L, Liu Y, Li Q, Zhang Y (2022) Hippocalcin-Like 1 blunts liver lipid metabolism to suppress tumorigenesis via directly targeting RUVBL1-mTOR signaling. *Theranostics* 12(17):7450–7464. <https://doi.org/10.7150/thno.75936>
- Chen X, Liu M, Tian Y, Li J, Qi Y, Zhao D, Wu Z, Huang M, Wong CCL, Wang HW, Wang J, Yang H, Xu Y (2018) Cryo-EM structure of human mTOR complex 2. *Cell Res* 28(5):518–528. <https://doi.org/10.1038/s41422-018-0029-3>
- Cuellar J, Ludlam WG, Tensmeyer NC, Aoba T, Dhavale M, Santiago C, Bueno-Carrasco MT, Mann MJ, Plimpton RL, Makaju A, Franklin S, Willardson BM, Valpuesta JM (2019) Structural and functional analysis of the role of the chaperonin CCT in mTOR complex assembly. *Nat Commun* 10(1):2865. <https://doi.org/10.1038/s41467-019-10781-1>
- Dauden MI, Lopez-Perrote A, Llorca O (2021) RUVBL1-RUVBL2 AAA-ATPase: a versatile scaffold for multiple complexes and functions. *Curr Opin Struct Biol* 67:78–85. <https://doi.org/10.1016/j.sbi.2020.08.010>
- David-Morrison G, Xu Z, Rui YN, Chang WL, Jaiswal M, Yamamoto S, Xiong B, Zhang K, Sandoval H, Duraine L, Zuo Z, Zhang S, Bellen HJ (2016) WAC regulates mTOR activity by acting as an adaptor for the TTT and pontin/reptin complexes. *Dev Cell* 36(2):139–151. <https://doi.org/10.1016/j.devcel.2015.12.019>
- Elias-Villalobos A, Toullec D, Faux C, Seveno M, Helmlinger D (2019) Chaperone-mediated ordered assembly of the SAGA and NuA4 transcription co-activator complexes in yeast. *Nat Commun* 10(1):5237. <https://doi.org/10.1038/s41467-019-13243-w>
- Ewens CA, Su M, Zhao L, Nano N, Houry WA, Southworth DR (2016) Architecture and nucleotide-dependent conformational changes of the Rvb1-Rvb2 AAA+ complex revealed by cryoelectron microscopy. *Structure* 24(5):657–666. <https://doi.org/10.1016/j.str.2016.03.018>
- Gano JJ, Simon JA (2010) A proteomic investigation of ligand-dependent HSP90 complexes reveals CHORDC1 as a novel ADP-dependent HSP90-interacting protein. *Mol Cell Proteomics* 9(2):255–270. <https://doi.org/10.1074/mcp.M900261-MCP200>
- Gorynia S, Bandejas TM, Pinho FG, McVey CE, Vornrhein C, Round A, Svergun DI, Donner P, Matias PM, Carrondo MA (2011) Structural and functional insights into a dodecameric molecular machine – the RuvBL1/RuvBL2 complex. *J Struct Biol* 176(3):279–291. <https://doi.org/10.1016/j.jsb.2011.09.001>

- Gwinn DM, Shackelford DB, Egan DF, Mihaylova MM, Mery A, Vasquez DS, Turk BE, Shaw RJ (2008) AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol Cell* 30(2):214–226. <https://doi.org/10.1016/j.molcel.2008.03.003>
- Hara K, Yonezawa K, Weng QP, Kozlowski MT, Belham C, Avruch J (1998) Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism. *J Biol Chem* 273(23):14484–14494. <https://doi.org/10.1074/jbc.273.23.14484>
- Hayashi T, Hatanaka M, Nagao K, Nakaseko Y, Kanoh J, Kokubu A, Ebe M, Yanagida M (2007) Rapamycin sensitivity of the *Schizosaccharomyces pombe* tor2 mutant and organization of two highly phosphorylated TOR complexes by specific and common subunits. *Genes Cells* 12(12):1357–1370. <https://doi.org/10.1111/j.1365-2443.2007.01141.x>
- Henri J, Chagot ME, Bourguet M, Abel Y, Terral G, Maurizy C, Aigueperse C, Georgescauld F, Vandermoere F, Saint-Fort R, Behm-Ansmant I, Charpentier B, Pradet-Balade B, Verheggen C, Bertrand E, Meyer P, Cianferani S, Manival X, Quinternet M (2018) Deep structural analysis of RPAP3 and PIH1D1, two components of the HSP90 co-chaperone R2TP complex. *Structure* 26(9):1196–1209. e1198. <https://doi.org/10.1016/j.str.2018.06.002>
- Horejsi Z, Stach L, Flower TG, Joshi D, Flynn H, Skehel JM, O'Reilly NJ, Ogrodowicz RW, Smerdon SJ, Boulton SJ (2014) Phosphorylation-dependent PIH1D1 interactions define substrate specificity of the R2TP cochaperone complex. *Cell Rep* 7(1):19–26. <https://doi.org/10.1016/j.celrep.2014.03.013>
- Horejsi Z, Takai H, Adelman CA, Collis SJ, Flynn H, Maslen S, Skehel JM, de Lange T, Boulton SJ (2010) CK2 phospho-dependent binding of R2TP complex to TEL2 is essential for mTOR and SMG1 stability. *Mol Cell* 39(6):839–850. <https://doi.org/10.1016/j.molcel.2010.08.037>
- Hurov KE, Cotta-Ramusino C, Elledge SJ (2010) A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. *Genes Dev* 24(17):1939–1950. <https://doi.org/10.1101/gad.1934210>
- Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, Colby G, Gebreab F, Gygi MP, Parzen H, Szpyt J, Tam S, Zarraga G, Pontano-Vaites L, Swarup S, White AE, Schweppe DK, Rad R, Erickson BK, Obar RA, Guruharsha KG, Li K, Artavanis-Tsakonas S, Gygi SP, Harper JW (2017) Architecture of the human interactome defines protein communities and disease networks. *Nature* 545(7655):505–509. <https://doi.org/10.1038/nature22366>
- Imseang S, Aylett CH, Maier T (2018) Architecture and activation of phosphatidylinositol 3-kinase related kinases. *Curr Opin Struct Biol* 49:177–189. <https://doi.org/10.1016/j.sbi.2018.03.010>
- Izumi N, Yamashita A, Hirano H, Ohno S (2012) Heat shock protein 90 regulates phosphatidylinositol 3-kinase-related protein kinase family proteins together with the RUVBL1/2 and Tel2-containing co-factor complex. *Cancer Sci* 103(1):50–57. <https://doi.org/10.1111/j.1349-7006.2011.02112.x>
- Izumi N, Yamashita A, Iwamatsu A, Kurata R, Nakamura H, Saari B, Hirano H, Anderson P, Ohno S (2010) AAA+ proteins RUVBL1 and RUVBL2 coordinate PIKK activity and function in nonsense-mediated mRNA decay. *Sci Signal* 3(116):ra27. <https://doi.org/10.1126/scisignal.2000468>
- Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J, Chabot B, Poirier GG, Hughes TR, Blanchette M, Price DH, Coulombe B (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 27(2):262–274. <https://doi.org/10.1016/j.molcel.2007.06.027>
- Ju D, Zhang W, Yan J, Zhao H, Li W, Wang J, Liao M, Xu Z, Wang Z, Zhou G, Mei L, Hou N, Ying S, Cai T, Chen S, Xie X, Lai L, Tang C, Park N, Takahashi JS, Huang N, Qi X, Zhang EE (2020) Chemical perturbations reveal that RUVBL2 regulates the circadian phase in mammals. *Sci Transl Med* 12(542). <https://doi.org/10.1126/scitranslmed.aba0769>
- Kabir MA, Kaminska J, Segel GB, Bethlendy G, Lin P, Della Seta F, Blegen C, Swiderek KM, Zoladek T, Arndt KT, Sherman F (2005) Physiological effects of unassembled chaperonin Cct subunits in the yeast *Saccharomyces cerevisiae*. *Yeast* 22(3):219–239. <https://doi.org/10.1002/yea.1210>

- Kaizuka T, Hara T, Oshiro N, Kikkawa U, Yonezawa K, Takehana K, Iemura S, Natsume T, Mizushima N (2010) Tti1 and Tel2 are critical factors in mammalian target of rapamycin complex assembly. *J Biol Chem* 285(26):20109–20116. <https://doi.org/10.1074/jbc.M110.121699>
- Kakihara Y, Kiguchi T, Ohazama A, Saeki M (2020) R2TP/PAQosome as a promising chemotherapeutic target in cancer. *Jpn Dent Sci Rev* 56(1):38–42. <https://doi.org/10.1016/j.jdsr.2019.08.001>
- Kanemaki M, Makino Y, Yoshida T, Kishimoto T, Koga A, Yamamoto K, Yamamoto M, Moncollin V, Egly JM, Muramatsu M, Tamura T (1997) Molecular cloning of a rat 49-kDa TBP-interacting protein (TIP49) that is highly homologous to the bacterial RuvB. *Biochem Biophys Res Commun* 235(1):64–68. <https://doi.org/10.1006/bbrc.1997.6729>
- Kanoh J, Yanagida M (2007) Tel2: a common partner of PIK-related kinases and a link between DNA checkpoint and nutritional response? *Genes Cells* 12(12):1301–1304. <https://doi.org/10.1111/j.1365-2443.2007.01142.x>
- Keith CT, Schreiber SL (1995) PIK-related kinases: DNA repair, recombination, and cell cycle checkpoints. *Science* 270(5233):50–51. <https://doi.org/10.1126/science.270.5233.50>
- Kim E, Goraksha-Hicks P, Li L, Neufeld TP, Guan KL (2008) Regulation of TORC1 by Rag GTPases in nutrient response. *Nat Cell Biol* 10(8):935–945. <https://doi.org/10.1038/ncb1753>
- Kim J, Guan KL (2019) mTOR as a central hub of nutrient signalling and cell growth. *Nat Cell Biol* 21(1):63–71. <https://doi.org/10.1038/s41556-018-0205-1>
- Kim SG, Hoffman GR, Poulgiannis G, Buel GR, Jang YJ, Lee KW, Kim BY, Erikson RL, Cantley LC, Choo AY, Blenis J (2013) Metabolic stress controls mTORC1 lysosomal localization and dimerization by regulating the TTT-RUVBL1/2 complex. *Mol Cell* 49(1):172–185. <https://doi.org/10.1016/j.molcel.2012.10.003>
- Kim Y, Park J, Joo SY, Kim BG, Jo A, Lee H, Cho Y (2022) Structure of the human TELO2-TTI1-TTI2 complex. *J Mol Biol* 434(2):167370. <https://doi.org/10.1016/j.jmb.2021.167370>
- Laplanche M, Sabatini DM (2012) mTOR signaling in growth control and disease. *Cell* 149(2):274–293. <https://doi.org/10.1016/j.cell.2012.03.017>
- Lempiäinen H, Halazonetis TD (2009) Emerging common themes in regulation of PIKKs and PI3Ks. *EMBO J* 28(20):3067–3073. <https://doi.org/10.1038/emboj.2009.281>
- Liu GY, Sabatini DM (2020) mTOR at the nexus of nutrition, growth, ageing and disease. *Nat Rev Mol Cell Biol* 21(4):183–203. <https://doi.org/10.1038/s41580-019-0199-y>
- Loewith R, Jacinto E, Wullschlegel S, Lorberg A, Crespo JL, Bonenfant D, Oppliger W, Jenoe P, Hall MN (2002) Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Mol Cell* 10(3):457–468. [https://doi.org/10.1016/s1097-2765\(02\)00636-6](https://doi.org/10.1016/s1097-2765(02)00636-6)
- Lopez-Perrote A, Hug N, Gonzalez-Corpas A, Rodriguez CF, Serna M, Garcia-Martin C, Boskovic J, Fernandez-Leiro R, Caceres JF, Llorca O (2020) Regulation of RUVBL1-RUVBL2 AAA-ATPases by the nonsense-mediated mRNA decay factor DHX34, as evidenced by Cryo-EM. *Elife* 9. <https://doi.org/10.7554/eLife.63042>
- Lopez-Perrote A, Munoz-Hernandez H, Gil D, Llorca O (2012) Conformational transitions regulate the exposure of a DNA-binding domain in the RuvBL1-RuvBL2 complex. *Nucleic Acids Res* 40(21):11086–11099. <https://doi.org/10.1093/nar/gks871>
- Lustig AJ, Petes TD (1986) Identification of yeast mutants with altered telomere structure. *Proc Natl Acad Sci U S A* 83(5):1398–1402. <https://doi.org/10.1073/pnas.83.5.1398>
- Lynham J, Houry WA (2022) The role of Hsp90-R2TP in macromolecular complex assembly and stabilization. *Biomolecules* 12(8). <https://doi.org/10.3390/biom12081045>
- Martin J, Masri J, Bernath A, Nishimura RN, Gera J (2008) Hsp70 associates with Rictor and is required for mTORC2 formation and activity. *Biochem Biophys Res Commun* 372(4):578–583. <https://doi.org/10.1016/j.bbrc.2008.05.086>
- Martino F, Pal M, Munoz-Hernandez H, Rodriguez CF, Nunez-Ramirez R, Gil-Carton D, Degliesposti G, Skehel JM, Roe SM, Prodromou C, Pearl LH, Llorca O (2018) RPAP3 provides a flexible scaffold for coupling HSP90 to the human R2TP co-chaperone complex. *Nat Commun* 9(1):1501. <https://doi.org/10.1038/s41467-018-03942-1>

- Matias PM, Gorynia S, Donner P, Carrondo MA (2006) Crystal structure of the human AAA+ protein RuvBL1. *J Biol Chem* 281(50):38918–38929. <https://doi.org/10.1074/jbc.M605625200>
- Maurizy C, Quinet M, Abel Y, Verheggen C, Santo PE, Bourguet M, ACF P, Bragantini B, Chagot ME, Robert MC, Abeza C, Fabre P, Fort P, Vandermoere F, PMF S, Rain JC, Charpentier B, Cianferani S, Bandejas TM, Pradet-Balade B, Manival X, Bertrand E (2018) The RPAP3-Cterminal domain identifies R2TP-like quaternary chaperones. *Nat Commun* 9(1):2093. <https://doi.org/10.1038/s41467-018-04431-1>
- Morgan RM, Pal M, Roe SM, Pearl LH, Prodromou C (2015) Tah1 helix-swap dimerization prevents mixed Hsp90 co-chaperone complexes. *Acta Crystallogr D Biol Crystallogr* 71(Pt 5):1197–1206. <https://doi.org/10.1107/S1399004715004551>
- Munoz-Hernandez H, Pal M, Rodriguez CF, Fernandez-Leiro R, Prodromou C, Pearl LH, Llorca O (2019) Structural mechanism for regulation of the AAA-ATPases RUVBL1-RUVBL2 in the R2TP co-chaperone revealed by cryo-EM. *Sci Adv* 5(5):eaaw1616. <https://doi.org/10.1126/sciadv.aaw1616>
- Nano N, Ugwu F, Seraphim TV, Li T, Azer G, Isaac M, Prakesch M, Barbosa LRS, Ramos CHI, Datti A, Houry WA (2020) Sorafenib as an inhibitor of RUVBL2. *Biomolecules* 10(4). <https://doi.org/10.3390/biom10040605>
- Napolitano G, Di Malta C, Ballabio A (2022) Non-canonical mTORC1 signaling at the lysosome. *Trends Cell Biol* 32(11):920–931. <https://doi.org/10.1016/j.tcb.2022.04.012>
- Pal M, Morgan M, Phelps SE, Roe SM, Parry-Morris S, Downs JA, Polier S, Pearl LH, Prodromou C (2014) Structural basis for phosphorylation-dependent recruitment of Tel2 to Hsp90 by Pih1. *Structure* 22(6):805–818. <https://doi.org/10.1016/j.str.2014.04.001>
- Pal M, Munoz-Hernandez H, Bjorklund D, Zhou L, Degliesposti G, Skehel JM, Hesketh EL, Thompson RF, Pearl LH, Llorca O, Prodromou C (2021) Structure of the TEO2-TTI1-TTI2 complex and its function in TOR recruitment to the R2TP chaperone. *Cell Rep* 36(1):109317. <https://doi.org/10.1016/j.celrep.2021.109317>
- Rivera-Calzada A, Pal M, Munoz-Hernandez H, Luque-Ortega JR, Gil-Carton D, Degliesposti G, Skehel JM, Prodromou C, Pearl LH, Llorca O (2017) The structure of the R2TP complex defines a platform for recruiting diverse client proteins to the HSP90 molecular chaperone system. *Structure* 25(7):1145–1152. e1144. <https://doi.org/10.1016/j.str.2017.05.016>
- Rodriguez CF, Llorca O (2020) RPAP3 C-terminal domain: a conserved domain for the assembly of R2TP co-chaperone complexes. *Cells* 9(5). <https://doi.org/10.3390/cells9051139>
- Roux PP, Ballif BA, Anjum R, Gygi SP, Blenis J (2004) Tumor-promoting phorbol esters and activated Ras inactivate the tuberous sclerosis tumor suppressor complex via p90 ribosomal S6 kinase. *Proc Natl Acad Sci U S A* 101(37):13489–13494. <https://doi.org/10.1073/pnas.0405659101>
- Sancak Y, Peterson TR, Shaul YD, Lindquist RA, Thoreen CC, Bar-Peled L, Sabatini DM (2008) The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* 320(5882):1496–1501. <https://doi.org/10.1126/science.1157535>
- Sarbassov DD, Ali SM, Sengupta S, Sheen JH, Hsu PP, Bagley AF, Markhard AL, Sabatini DM (2006) Prolonged rapamycin treatment inhibits mTORC2 assembly and Akt/PKB. *Mol Cell* 22(2):159–168. <https://doi.org/10.1016/j.molcel.2006.03.029>
- Saxton RA, Sabatini DM (2017) mTOR signaling in growth, metabolism, and disease. *Cell* 168(6):960–976. <https://doi.org/10.1016/j.cell.2017.02.004>
- Scaiola A, Mangia F, Imseng S, Boehringer D, Berneiser K, Shimobayashi M, Stüttfeld E, Hall MN, Ban N, Maier T (2020) The 3.2-Å resolution structure of human mTORC2. *Sci Adv* 6(45). <https://doi.org/10.1126/sciadv.abc1251>
- Schmidt A, Kunz J, Hall MN (1996) TOR2 is required for organization of the actin cytoskeleton in yeast. *Proc Natl Acad Sci U S A* 93(24):13780–13785. <https://doi.org/10.1073/pnas.93.24.13780>
- Serna M, Gonzalez-Corpas A, Cabezero S, Lopez-Perrote A, Degliesposti G, Zarzuela E, Skehel JM, Munoz J, Llorca O (2022) CryoEM of RUVBL1-RUVBL2-ZNHIT2, a complex that

- interacts with pre-mRNA-processing-splicing factor 8. *Nucleic Acids Res* 50(2):1128–1146. <https://doi.org/10.1093/nar/gkab1267>
- Shevchenko A, Roguev A, Schaft D, Buchanan L, Habermann B, Sakalar C, Thomas H, Krogan NJ, Shevchenko A, Stewart AF (2008) Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment. *Genome Biol* 9(11):R167. <https://doi.org/10.1186/gb-2008-9-11-r167>
- Shin SH, Lee JS, Zhang JM, Choi S, Boskovic ZV, Zhao R, Song M, Wang R, Tian J, Lee MH, Kim JH, Jeong M, Lee JH, Petukhov M, Lee SW, Kim SG, Zou L, Byun S (2020) Synthetic lethality by targeting the RUVBL1/2-TTT complex in mTORC1-hyperactive cancer cells. *Sci Adv* 6(31):eaay9131. <https://doi.org/10.1126/sciadv.aay9131>
- Silva STN, Brito JA, Arranz R, Sorzano COS, Ebel C, Douth J, Tully MD, Carazo JM, Carrascosa JL, Matias PM, Bandejas TM (2018) X-ray structure of full-length human RuvB-Like 2 – mechanistic insights into coupling between ATP binding and mechanical action. *Sci Rep* 8(1):13726. <https://doi.org/10.1038/s41598-018-31997-z>
- Silva-Martin N, Dauden MI, Glatt S, Hoffmann NA, Kastitis P, Bork P, Beck M, Muller CW (2016) The combination of X-ray crystallography and cryo-electron microscopy provides insight into the overall architecture of the dodecameric Rvb1/Rvb2 complex. *PLoS One* 11(1):e0146457. <https://doi.org/10.1371/journal.pone.0146457>
- Takai H, Wang RC, Takai KK, Yang H, de Lange T (2007) Tel2 regulates the stability of PI3K-related protein kinases. *Cell* 131(7):1248–1259. <https://doi.org/10.1016/j.cell.2007.10.052>
- Takai H, Xie Y, de Lange T, Pavletich NP (2010) Tel2 structure and function in the Hsp90-dependent maturation of mTOR and ATR complexes. *Genes Dev* 24(18):2019–2030. <https://doi.org/10.1101/gad.1956410>
- Tang HW, Weng JH, Lee WX, Hu Y, Gu L, Cho S, Lee G, Binari R, Li C, Cheng ME, Kim AR, Xu J, Shen Z, Xu C, Asara JM, Blenis J, Perrimon N (2021) mTORC1-chaperonin CCT signaling regulates m(6)A RNA methylation to suppress autophagy. *Proc Natl Acad Sci U S A* 118(10). <https://doi.org/10.1073/pnas.2021945118>
- Toullec D, Elias-Villalobos A, Faux C, Noly A, Lledo G, Seveno M, Helmlinger D (2021) The Hsp90 cochaperone TTT promotes cotranslational maturation of PIKKs prior to complex assembly. *Cell Rep* 37(3):109867. <https://doi.org/10.1016/j.celrep.2021.109867>
- Vaughan CK (2014) Hsp90 picks PIKKs via R2TP and Tel2. *Structure* 22(6):799–800. <https://doi.org/10.1016/j.str.2014.05.012>
- Wang JZ, Zhu H, You P, Liu H, Wang WK, Fan X, Yang Y, Xu K, Zhu Y, Li Q, Wu P, Peng C, Wong CC, Li K, Shi Y, Zhang N, Wang X, Zeng R, Huang Y, Yang L, Wang Z, Hui J (2022) Upregulated YB-1 protein promotes glioblastoma growth through a YB-1/CCT4/mLST8/mTOR pathway. *J Clin Invest* 132(8). <https://doi.org/10.1172/JCI146536>
- Yang H, Rudge DG, Koos JD, Vaidialingam B, Yang HJ, Pavletich NP (2013) mTOR kinase structure, mechanism and regulation. *Nature* 497(7448):217–223. <https://doi.org/10.1038/nature12122>
- Young JC, Agashe VR, Siegers K, Hartl FU (2004) Pathways of chaperone-mediated protein folding in the cytosol. *Nat Rev Mol Cell Biol* 5(10):781–791. <https://doi.org/10.1038/nrm1492>
- Yu Z, Chen J, Takagi E, Wang F, Saha B, Liu X, Joubert LM, Gleason CE, Jin M, Li C, Nowotny C, Agard D, Cheng Y, Pearce D (2022) Interactions between mTORC2 core subunits Rictor and mSin1 dictate selective and context-dependent phosphorylation of substrate kinases SGK1 and Akt. *J Biol Chem* 298(9):102288. <https://doi.org/10.1016/j.jbc.2022.102288>
- Zhang Y, Liu Y, Duan J, Yan H, Zhang J, Zhang H, Fan Q, Luo F, Yan G, Qiao K, Liu J (2016) Hippocalcin-like 1 suppresses hepatocellular carcinoma progression by promoting p21(Waf/Cip1) stabilization by activating the ERK1/2-MAPK pathway. *Hepatology* 63(3):880–897. <https://doi.org/10.1002/hep.28395>
- Zhao R, Davey M, Hsu YC, Kaplanek P, Tong A, Parsons AB, Krogan N, Cagney G, Mai D, Greenblatt J, Boone C, Emili A, Houry WA (2005) Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell* 120(5):715–727. <https://doi.org/10.1016/j.cell.2004.12.024>

- Zhao R, Houry WA (2007) Molecular interaction network of the Hsp90 chaperone system. *Adv Exp Med Biol* 594:27–36. https://doi.org/10.1007/978-0-387-39975-1_3
- Zoncu R, Efeyan A, Sabatini DM (2011) mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol* 12(1):21–35. <https://doi.org/10.1038/nrm3025>

Chapter 18

Valosin-Containing Protein (VCP)/p97 Oligomerization



Guimei Yu, Yunpeng Bai, and Zhong-Yin Zhang

Abstract Valosin-containing protein (VCP), also known as p97, is an evolutionarily conserved AAA+ ATPase essential for cellular homeostasis. Cooperating with different sets of cofactors, VCP is involved in multiple cellular processes through either the ubiquitin-proteasome system (UPS) or the autophagy/lysosomal route. Pathogenic mutations frequently found at the interface between the NTD domain and D1 ATPase domain have been shown to cause malfunction of VCP, leading to degenerative disorders including the inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia (IBMPFD), amyotrophic lateral sclerosis (ALS), and cancers. Therefore, VCP has been considered as a potential therapeutic target for neurodegeneration and cancer. Most of previous studies found VCP predominantly exists and functions as a hexamer, which unfolds and extracts ubiquitinated substrates from protein complexes for degradation. However, recent studies have characterized a new VCP dodecameric state and revealed a controlling mechanism of VCP oligomeric states mediated by the D2 domain nucleotide occupancy. Here, we summarize our recent knowledge on VCP oligomerization, regulation, and potential implications of VCP in cellular function and pathogenic progression.

Keywords VCP/p97 · Oligomerization · Hexamer · Dodecamer · Nucleotide binding · Ubiquitination · IBMPFD · ALS · Cancer

G. Yu

School of Pharmaceutical Science, Hengyang Medical School, University of South China, Hengyang, China

e-mail: guimei.yu@tmu.edu.cn

Y. Bai · Z.-Y. Zhang (✉)

Borch Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN, USA

e-mail: bai62@purdue.edu; zhang-zy@purdue.edu

Biological Function of VCP and Associated Diseases

Valosin-containing protein (VCP, also known as p97) belongs to the ATPase associated with diverse cellular activities (AAA⁺) protein family, which usually assembles into a hexameric ring-shaped structure through the nucleotide-binding domains (NBDs) and extracts substrates with forces generated by the ATPase cycle (Meyer et al. 2012; van den Boom and Meyer 2018). With two tandem NBDs, VCP is a type II AAA+ ATPase (Barthelme and Sauer 2016). VCP is involved in a large number of cellular processes through cooperation with a variety of cofactor proteins (Hanzelmann and Schindelin 2017). For instance, VCP cooperates with the UFD1-NPL4 binary complex to drive the extraction of polyubiquitylated substrates from cellular structures for degradation in multiple proteasomal degradation pathways, including endoplasmic reticulum-associated degradation (ERAD) (Ye et al. 2003; Lee et al. 2013), chromatin-associated degradation (CAD) (Dantuma and Hoppe 2012; Vaz et al. 2013), mitochondria-associated degradation (MAD) (Xu et al. 2011; Wu et al. 2016), and cell cycle progression (Mouysset et al. 2008). VCP also participates in lysosomal hemostasis and autophagy by interacting with the deubiquitinase YOD1 and its C-terminal cofactors PLAA and UBXD1 (Papadopoulos et al. 2017; Bai et al. 2023). In addition, VCP is involved in various membrane trafficking processes, including Golgi reassembly following mitosis and endosomal trafficking, by interacting with multiple cofactors (Meyer 2005; Ritz et al. 2011).

Given the importance of VCP in cellular homeostasis, dysfunction of VCP has therefore been linked to several human diseases, including neurodegenerative disorders and cancers. Mutations in the *VCP* gene cause late-onset degenerative diseases, including inclusion body myopathy with Paget disease and frontotemporal dementia (IBMPFD), amyotrophic lateral sclerosis (ALS), Huntington's disease (HD), and Charcot-Marie-Tooth (CMT) disease (Watts et al. 2004; Johnson et al. 2010; Gonzalez et al. 2014; Meyer and Wehl 2014; Guo et al. 2016; Gite et al. 2020). As a result of the disruption of protein degradation pathways by VCP mutations, the common pathological feature of these neurodegenerative disorders is the accumulation of ubiquitinated aggregates (Meyer and Wehl 2014). On the other hand, VCP is also upregulated in certain cancers, possibly by enhancing the protein degradation efficiency in response to biosynthetic demands caused by uncontrolled proliferation to survive metabolic stress and ensure cellular homeostasis (Deshaies 2014; Costantini et al. 2021). As a result, targeting VCP to induce crisis of protein homeostasis has been considered as a novel strategy to elicit cell death in cancer cells (Anderson et al. 2015).

Molecular Structure of VCP

Human VCP consists of 806 amino acid residues and contains three major domains, including a regulatory N-terminal domain (NTD), followed by two tandem RecA-like ATPase domains (D1 and D2), and a C-terminal extension (CTE) (Fig. 18.1). Similar domain organization is conserved among VCP homologs in mammals and yeast. The NTD (aa 1–187) possesses two globular subdomains formed by a double- Ψ -barrel and a four-stranded β -barrel connected by a loop. Most cofactors bind to VCP through the NTD. D1 and D2 ATPase domains (aa 209–462 and aa 481–762, respectively) share almost identical structures, with two subdomains. The N-terminal RecA-like subdomain features a central β -sheet (β 1–5) sandwiched by multiple α -helices (α 1–4), which is followed by the α -helix bundle subdomain (α 5–8 for D1 α and α 5–9 for D2 α). The adenine base moiety of nucleotides ATP or ADP inserts into the cleft between the RecA-like and D1/2 α subdomains, while other essential motifs including the Walker A motif (Gx(4)GK[TS]), Walker B motif (hhhhDELD), and the arginine finger, stabilize the phosphate groups directly or by coordinating a Mg^{2+} ion (Figs. 18.1 and 18.2). Notably, D2 α contains a unique α 9-helix bridging D2 and CTE (Fig. 18.1), which were found to be critical for both the ATPase activities and the formation of high-order assemblies (Hanzelmann and Schindelin 2016;

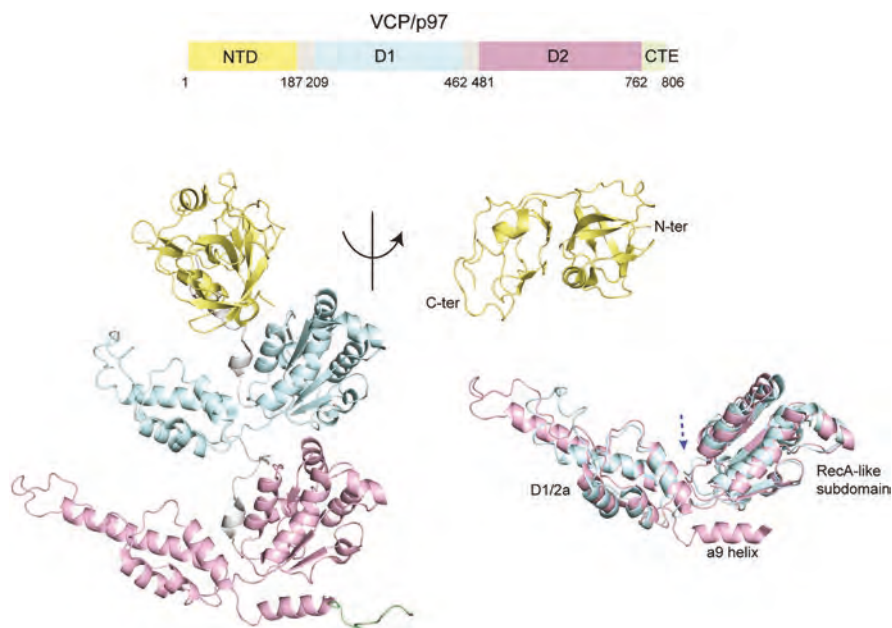


Fig. 18.1 Domain organization and the molecular structure of VCP/p97. VCP contains three major domains (NTD, D1, and D2) and a CTE. Full-length human VCP is composed of 806 residues, with NTD at aa 1–187, D1 at aa 209–462, D2 at aa 481–762, and CTE at 763–806. The arrow points to the nucleotide-binding pockets

Yu et al. 2021). CTE (aa 763–806) is highly flexible, the majority of which is not resolved in the reported structures. Multiple cofactors including UBXD1, PLAA, PNGase, and HOIP directly associate with CTE, where the last ten residues (aa 796–806) mediate these VCP-cofactor interactions (Hanzelmann and Schindelin 2017).

Hexameric Structures of VCPs

As for the other ATPases/helicases, VCP protomers assemble to form a homo-hexamer, which is considered the major functional form of VCP at the present time (Fig. 18.2a). The bilobed (RecA-like subdomain and the D1/2 α) D1 and D2 ATPase domains encompass the RecA-like subdomains in cis, leading to the formation of the barrel-shaped VCP hexamer. The nucleotide-binding pockets reside at the interface of two protomers in both D1 and D2 domains aforementioned, highlighting the importance of hexamerization for VCP function (Figs. 18.1 and 18.2b). Disruption of the hexamerization interface results in loss of VCP ATPase activity. The position of NTDs in VCP hexamers varies in response to the nucleotide-binding status in D1. In the ADP-bound state, NTD is coplanar with D1 (“down” conformation), while upon the dissociation of ADP (the apo state) and the binding of ATP γ S (the mimic of ATP-bound state), NTDs rotate up and switch to an “up” conformation (Banerjee et al. 2016; Hanzelmann and Schindelin 2016; Yu et al. 2021) (Fig. 18.2c). By contrast, the nucleotide-binding status of D2 could regulate the D1–D2 interdomain orientation and the conformation of CTE. While dissociation of ADP did not cause profound conformational change, binding of ATP γ S in D2 led to an $\sim 15^\circ$ rotational twist away from D1, and reorientation of the CTE toward the central pore (Fig. 18.2c) (Banerjee et al. 2016; Hanzelmann and Schindelin 2016; Yu et al. 2021). Although both D1 and D2 domains are capable of ATP hydrolysis, D2 is a more potent ATPase (Song et al. 2003; Chou et al. 2014) and contributes to the majority of the energy required for substrate unfolding and extraction. The CTE is flexible and extends outward relative to the central channel of the hexamer in the ADP-bound state, but is kinked and reoriented toward the channel by ATP (Banerjee et al. 2016; Hanzelmann and Schindelin 2016).

VCPs in different nucleotide-bound states have been resolved as symmetric hexamers. However, the substrate-engaged VCP or Cdc48 (yeast homolog) and VAT (archaea homolog) complexes have been resolved with asymmetric structures, suggesting the breakage of the hexameric symmetry in action (Ripstein et al. 2017; Cooney et al. 2019; Twomey et al. 2019; Pan et al. 2021; Xu et al. 2022a). Four to all six protomers are arranged in a helical conformation to associate with and translocate the substrate threaded to the central pore, which is powered by the accompanied conformational changes triggered by ATP hydrolysis (Fig. 18.3). Pore loop 1 (aa 277–279, aa 550–552) and pore loop 2 (aa 315–318, aa 589–593) from both D1 and D2 in the ATP-bound VCP protomers make tight interactions with the substrate, but that association weakens after ATP hydrolysis and ADP dissociation, leading to ratchet-like motions following the sequential hydrolysis of ATP (Fig. 18.3b).

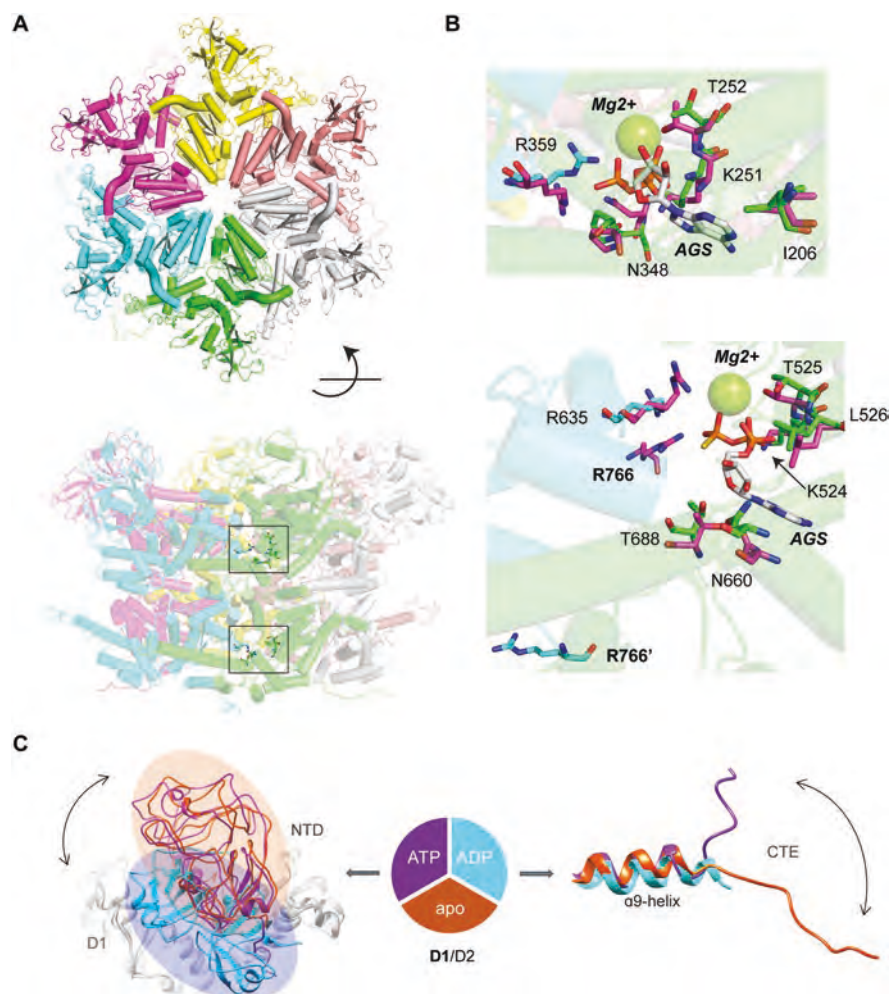


Fig. 18.2 Structure and nucleotide-binding-induced conformational changes of VCP. **(a)** VCP assembles to form a homo-hexamer. The top and side views are displayed (PDBID: 7K59). **(b)** The nucleotide-binding pockets of D1 and D2 ATPase domains reside at the hexamerization interface, constituted by residues from the Walker A/B motifs and the arginine finger of a neighbor. The apo- and ATP γ S-bound nucleotide-binding pockets (PDB ID: 5ftn, in magenta) were superimposed. While no profound conformational changes were observed for other pocket residues, a large-scale relocation of the D2 arginine finger residue R766 was observed. **(c)** The nucleotide-binding status in D1 controls the conformation of NTD, where NTD is coplanar with D1 in ADP-bound state but is rotated up upon ADP dissociation (apo) and ATP γ S binding. By contrast, the nucleotide-binding status of D2 is linked to the D1–D2 interdomain orientation and the conformation of CTE

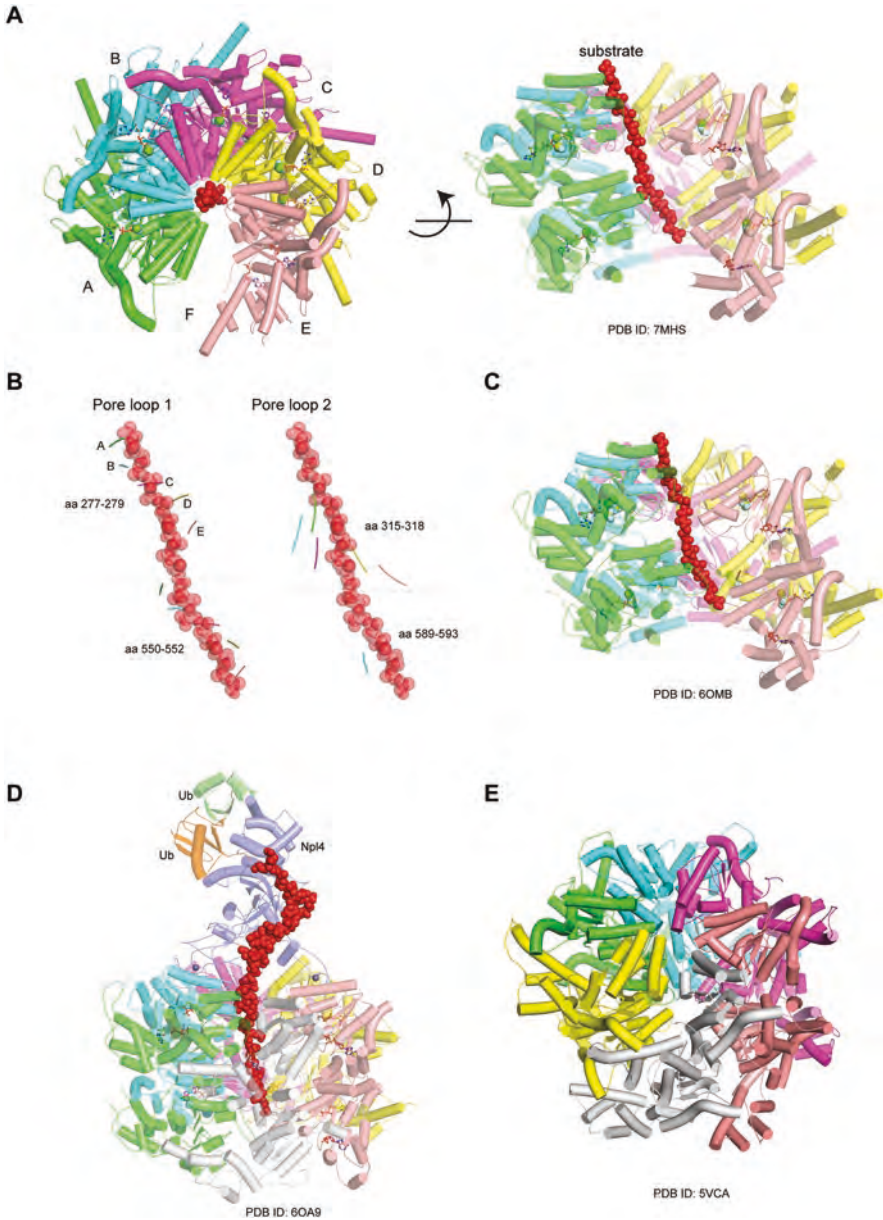


Fig. 18.3 Asymmetric spiral conformations of VCP in the act of substrate processing. (a) A structure of VCP in the process of substrate processing (PDBID: 7MHS). (b) D1 and D2 pore loops (pore loops 1 and 2) form oscillating staircase-like contacts with the substrate accompanying ATP hydrolysis. The four chains bound with ATP (ATP analog, chains A, B, C, and D) form the strongest contacts, while chain E bound with ADP associates less tightly. Chain F potentially in the apo state is highly flexible and not resolved. (c) and (d) are structures of Cdc48 in the substrate processing conformations (PDBID: 6OMB and 6OA9, respectively). (e) Structure of VAT in the substrate processing conformation (PDBID: 5VCA)

Characterization of a New Dodecamer State of VCP

The hexameric model described above can explain most of VCP functions, particularly the involvement of VCP in unfolding and extraction of polyubiquitinated substrates for UPS-mediated protein degradation (van den Boom and Meyer 2018). However, the functional mechanisms of VCP in many cellular processes such as homotypic membrane fusion, autophagy, and endosomal trafficking remain elusive (Barthelme and Sauer 2016; Stach and Freemont 2017; van den Boom and Meyer 2018), which may require alternative functional states of VCP. This notion is further supported by the observation that chemical inhibition of the ATPase activity does not block all of the normal cellular functions mediated by VCP (Anderson et al. 2015).

In an attempt to overexpress and purify human VCP using mammalian cells, we identified and resolved a near-atomic resolution structure of VCP using single-particle cryo-electron microscopy, which, in striking contrast to the commonly observed hexamers, is in a dodecamer state with two hexamers assembled in a

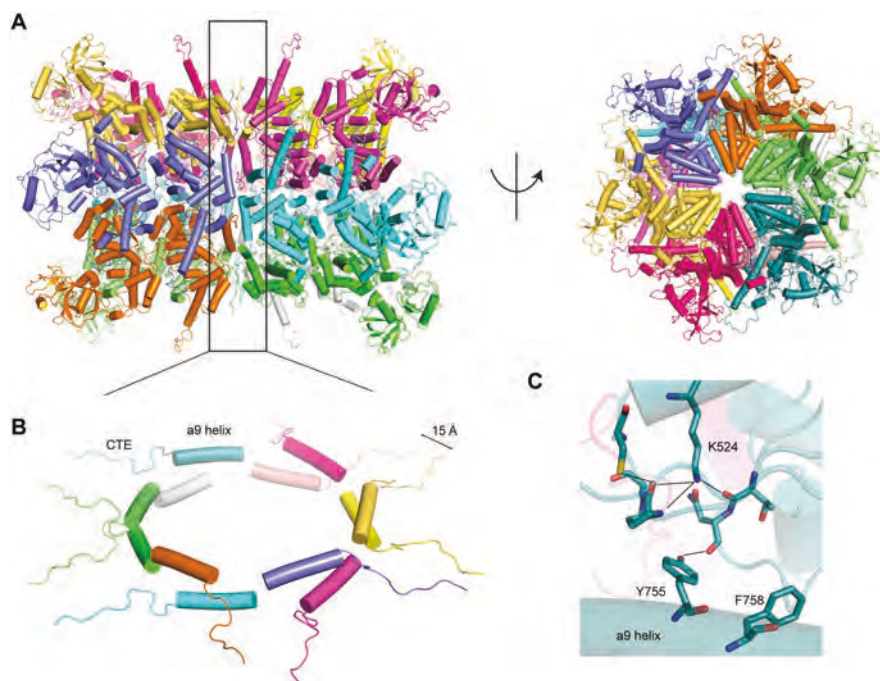


Fig. 18.4 Structure of VCP dodecamer. (a) The top and side views of VCP dodecamer isolated from H1299 mammalian cells overexpressing VCP (PDB ID: 7 K56). (b) The $\alpha 9$ -helices and CTE constitute the major dimerization interface of the two hexamers for dodecamer formation. The dodecamer assembly brings the CTEs from the top and bottom hexamers to close proximity (15 Å), which is a unique feature of the dodecamer state. (c) K524 and Y755 form contact hubs affecting the conformation of $\alpha 9$ -helices and therefore dodecamer formation

tail-to-tail mode (Fig. 18.4) (Yu et al. 2021). In line with this finding, similar dodecamer-like particles had been noted in earlier purifications of VCP and the archaea homology VAT (Peters et al. 1992; Pamnani et al. 1997; Niwa et al. 2012). To understand the structural basis of VCP dodecamer formation, we compared the dodecamer structure with known hexamer structures of VCP, and observed multiple key structural features. The D2 α subdomains constitute the major interaction interface between the two hexamers in the dodecamer state (Fig. 18.4a, b). Particularly, the α 9-helices and CTEs mediate extensive intra- and inter-hexamer interactions, substantially stabilizing the dodecamer structure (Fig. 18.4b). This new assembly brings the last resolved residues of CTE (S775) to close proximity (~ 15 Å), which is unique to the dodecamer state of VCP, and could potentially impact cofactors binding to CTE. In addition, we found the resolved dodecamer structure is in an apo status without nucleotides in neither D1 nor D2, which are distinct from previously reported structures (Zhang et al. 2000; DeLaBarre and Brunger 2003; Davies et al. 2008; Banerjee et al. 2016; Hanzelmann and Schindelin 2016; Schuller et al. 2016). We therefore hypothesized that the apo nucleotide status is potentially linked to VCP dodecamer formation. VCP is typically purified with prebound ADP, which could potentially explain the paucity of dodecamer state (Briggs et al. 2008; Tang and Xia 2016). Supporting this notion, ADP hydrolysis and removal by apyrase treatment substantially enriched the dodecamer state in purified VCP (Yu et al. 2021, 2022). Cryo-EM structural characterization of the bacterially expressed VCP dodecamer generated by apyrase treatment revealed almost identical structure, reinforcing the link between the apo status and dodecamer formation (Yu et al. 2021). To further define the roles of D1 and D2 nucleotide-binding status in VCP dodecamer formation, we generated a series of point mutations in the nucleotide-binding pockets. While mutations in D1 failed in inducing dodecamer formation, mutations T525V and L526A in D2 caused significant increase in VCP dodecamer population, indicating that the apo D2 likely plays a major role in promoting VCP dodecamer formation. Around the same time, another group also reported a dodecamer structure of VCP, which, consistent with our results, also contains apo D2 (Hoq et al. 2021). Distinct from T525 and L526, mutation of another essential residue for nucleotide binding in D2, K524, substantially impaired the dodecamer formation due to the important role of K524 in stabilizing the dodecamer interface. K524 stabilizes the β -phosphate in the nucleotide-bound states but is shifted to build new contacts with the Walker A and the aa 623–628 loops, further bridging contacts with key residues in the α 9-helices such as Y755 and F758 (Fig. 18.4c). These contacts, by inducing conformational changes of the α 9-helices and CTEs, stabilize the dodecamer state. In line with the structural features, mutation of residues within or in close proximity of the interaction network prevented dodecamer formation (Hanzelmann and Schindelin 2016; Yu et al. 2021). Altogether, these results suggest that apo D2, with new contact network from the two hubs by K524 and Y755, promotes the formation of the VCP dodecamer state.

Additional VCP Dodecamer Structures

More VCP dodecamer structures have been reported by different labs (Table 18.1). Nandi and colleagues found that the IBMPFD-associated mutation R155H could affect VCP dodecamerization and more than 30% of VCP^{R155H} were present in the dodecamer state in solution (Nandi et al. 2021). In line with our conclusion (Yu et al. 2021, 2022), VCP^{R155H} dodecamer was also resolved in the apo D2 status and could be destabilized by ADP or ATPγS. In addition, Caffrey and colleagues reported two CB-5083-bound VCP dodecamer structures at 2.4 Å and 3.4 Å, respectively (Caffrey et al. 2021). CB-5083 is a potent inhibitor of VCP, which binds to the D2 domain, partially occupying the nucleotide-binding pocket (Tang et al. 2019). The binding of CB-5083, therefore, could displace ADP in the D2 pocket, resulting in the apo nucleotide status of D2 and thereby dodecamer formation (Tang et al. 2019; Caffrey et al. 2021). This is consistent with our proposed apo D2-regulated dodecamer formation mechanism (Yu et al. 2021). Caffrey and colleagues further confirmed that the dodecamer formation was not affected by the nucleotide status of D1, which agrees well with our conclusion. Lastly, another cryo-EM structure of the NMS-873 inhibitor-bound VCP dodecamer has been deposited to the Protein Data Bank (PDB ID: 7WUB), which also contains apo D2 (Liu et al. 2022).

Interestingly, Gao and colleagues reported two dodecamer structures with D1 bound with ATPγS and D2 in ADP or ATPγS-bound state (Gao et al. 2022) (Table 18.1), which appears to contradict the proposed dodecamer formation

Table 18.1 Available structures of VCP dodecamer in the Protein Data Bank (PDB)

PDB ID	EMD ID ^d	Res (Å)	Mutation	Ligand/cofactor	D1 status	D2 status	Reference
6MCK ^a	—	3.8 ^b	NTD truncation	CB-5083	Apo	Apo	Tang et al. 2019
7K56	22,675	3.9	WT	—	Apo	Apo	Yu et al. 2021
7K57	22,676	3.7	WT	—	Apo	Apo	Yu et al. 2021
7L5W	23,191	3.4	R155H	—	Apo	Apo	Nandi et al. 2021
7L5X	23,192	6.1	R155H	—	Apo	Apo	Nandi et al. 2021
7RLI	24,531	3.1	WT	CB-5083	ADP	Apo	Caffrey et al. 2021
7RLJ	24,532	3.8	WT	CB-5083	ATPrS	Apo	Caffrey et al. 2021
7VCS	31,894	3.4	WT	SPRTN ^c	ATPrS	ATPrS ^c	Gao et al. 2022
7VCU	31,896	3.2	WT	SPRTN ^c	ATPrS	ADP ^c	Gao et al. 2022
7WUB	32,827	3.0	WT	NMS-873	ADP	Apo	Liu et al. 2022

^a Dodecamer assembly observed in the crystal space

^b Determined by X-ray crystallography

^c Density was not resolved or weak

^d Electron Microscopy Data Bank

mechanism. One potential cause could be the residual nucleotide-bound hexamers in the classified dodecamer classes, as the nucleotide densities were weak in comparison with the juxtaposed D1 (PDB IDs: 7VCS, 7VCU). Additionally, these two dodecameric structures were isolated through 2D and 3D classification in the single-particle cryo-EM analysis of a VCP-cofactor complex (VCP-SPRTN); although the density of SPRTN was not resolved, it may regulate the dodecamer formation.

What Regulates VCP Dodecamerization?

As VCP proteins were typically copurified with prebound ADP using the bacterial recombinant expression systems (Briggs et al. 2008; Tang and Xia 2016), the hexamer state has been predominantly identified, corroborating the proposed requisite of apo D2 for dodecamer formation (Fig. 18.5). By contrast, a notable fraction of dodecamers were observed when VCP was expressed and purified from the mammalian cells (Yu et al. 2021, 2022), implying that some unknown cellular factors in mammalian cells such as posttranslational modifications, ligands, or cofactors may possibly regulate the hexamer to dodecamer transition of VCP. In line with this hypothesis, Gao et al. discovered the presence of dodecamer states in the VCP-SPRTN complex samples generated from VCP purified from the bacterial expression system (Gao et al. 2022). However, further study needs to be performed to clarify the possible role of SPRTN in potentially triggering and stabilizing VCP dodecamer. Disease-causing mutations may also alter dodecamer formation of VCP. Differing from wild-type VCP, around 40% of the disease-causing R155H mutant VCP was found existing in the dodecamer state (Nandi et al. 2021); however, the functional relevance and the link to pathogenesis and how other disease-causing mutations may influence the oligomer state of VCP remain unknown. In addition, VCP inhibitors might also affect dodecamer formation. Three CB-5083 inhibitor-bound VCP dodecamer structures have been reported by two different research groups (Tang et al. 2019; Caffrey et al. 2021), suggestive of CB-5083 binding-induced dodecamer formation. Additionally, an NMS-873 inhibitor-bound VCP dodecamer structure was also deposited to PDB (Table 18.1). Hoq and colleagues also found VCP particles captured using the UPCDC30245 inhibitor analog-decorated TEM support layer switched to the dodecamer state (Hoq et al. 2021). All of the evidence indicates inhibitor-trapped dodecamer states; however, future studies will be required to determine their functional relationship.

Cellular Function of VCP Dodecamer?

It is well known that VCP plays a vital role in maintaining cellular protein and organelle homeostasis, where it facilitates the unfolding and extraction of misfolded and ubiquitinated proteins from large aggregations or damaged organelles for

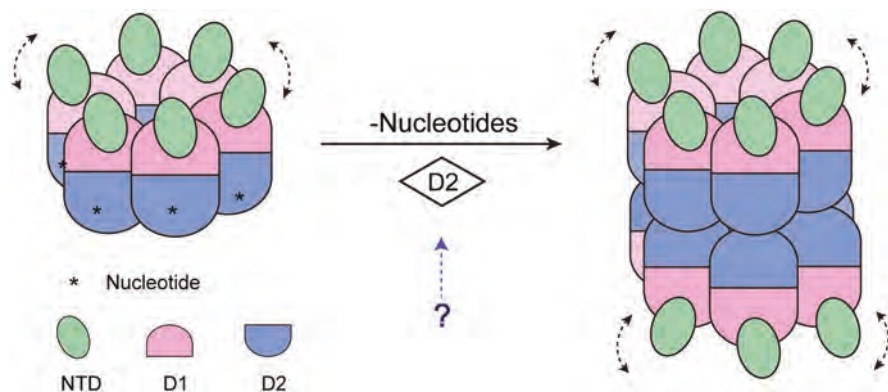


Fig. 18.5 Schematic diagram showing the transition of VCP hexamers to dodecamers. The nucleotide-binding status of D1 controls the conformation of NTD (“down” and “up” conformations as indicated by the curved, dashed arrows). On the other hand, nucleotide-binding status in D2 regulates D2 and CTE conformations. Particularly, when D2 switched to the apo nucleotide-binding status, VCP hexamers further dimerize to form dodecamers. However, the cellular factors and the effects of disease-associated mutations for dodecamer formation remain to be determined

proteasomal or lysosomal degradation (van den Boom and Meyer 2018). The hexamer state is currently considered the major functional state underlying the segregase activity of VCP. Powered by sequential ATP hydrolysis, the pore loops in the ATPase domains of the six protomers form an oscillating spiral staircase, interacting with the substrate threaded to the central pore, to unfold and extract proteins (Fig. 18.3). However, the fact that the dodecamer structure is incompatible with the ATP hydrolysis and substrate processing models, combined with the discovery of VCP inhibitor-trapped dodecamer states, may suggest that VCP dodecamer lacks the segregase activity (Figs. 18.3 and 18.4). On the other hand, the dodecamer assembly may block the complete threading but allow for partial segregation of substrates, which appears to be required in some cellular processes such as membrane fusion (Huang et al. 2016; Stach and Freemont 2017; van den Boom and Meyer 2018). Consistently, alternative functional models of VCP with “holding” or partial substrate threading roles have also been implied (Stach and Freemont 2017; van den Boom and Meyer 2018), which may have a link with the new dodecamer state and deserve further exploration.

Notably, the model could be a shared scheme for some AAA+ ATPases, which may exert different functions in different oligomer states. For instance, hexamer and double-hexamer/dodecamer states have also been identified for other AAA+ ATPases, such as the MCM2–7 complex and the Skd3 ATPase (Xu et al. 2022b; Gupta et al. 2023; Wu et al. 2023). In particular, Gupta et al. reported that Skd3 might fulfill the disaggregase and foldase activities using the hexamer and dodecamer states, respectively (Gupta et al. 2023); while the ATP hydrolysis-fueled ratchet-like motions of the hexamer drive substrate extraction, the large chamber formed by the dodecamer assembly could facilitate the refolding of substrates.

Pharmacological Inhibitors of VCP

As a well-established target for treating cancer and neurodegenerative diseases, VCP has attracted much attention to develop inhibitors for therapeutic intervention (Kilgas and Ramadan 2023). So far, many ATP-competitive inhibitors, covalent inhibitors, as well as allosteric inhibitors for VCP have been discovered (Fig. 18.6). Although developing potent and selective ATPase active site inhibitors is always challenging, a series of VCP ATP-competitive inhibitors have been developed. DBE-Q binds to both D1 and D1 domains of VCP with an IC_{50} value of 1.5 μ M for VCP (Chou et al. 2011), but it also inhibits VPS4 with an IC_{50} of 11.5 μ M (Zhang et al. 2016). ML240 and ML241, derivatives of DBE-Q, were developed to show

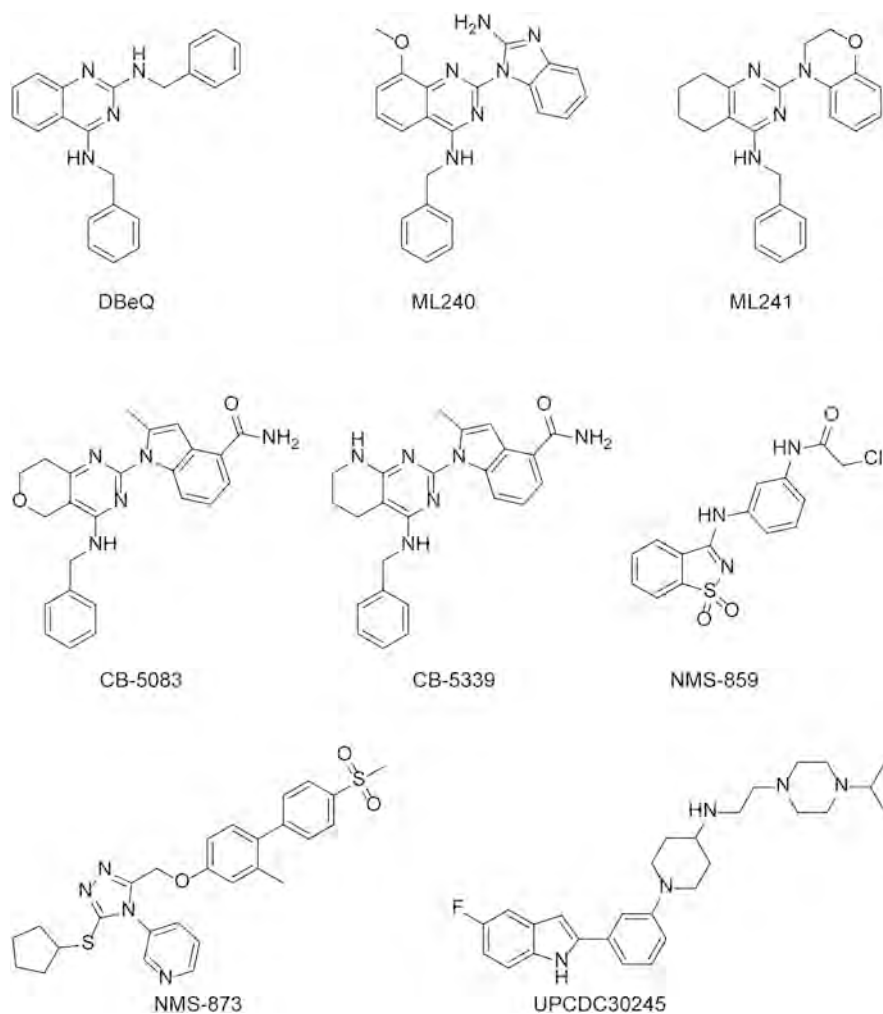


Fig. 18.6 VCP inhibitors. The chemical structures of eight VCP inhibitors are displayed

increased potency with IC_{50} s of 100 nM (Chou et al. 2013). Further optimization lead to the discovery of CB-5083, a potent, selective, and orally bioavailable inhibitor of VCP, selectively targeting its D2 site with an IC_{50} of 11 nM (Zhou et al. 2015). Although CB-5083 demonstrated good antitumor activity by inducing apoptosis through modulation of key VCP-related pathways in a broad range of both solid and hematological xenograft tumor models (Anderson et al. 2015), the phase I clinical trials for cancer were still terminated because of adverse effects on vision, including photophobia and dyschromatopsia, possibly due to its off-target effect on phosphodiesterase 6 (PDE6), a key catalytic enzyme of visual phototransduction (Leinonen et al. 2021). Most recently, CB-5339 was further developed to retain the high on-target potency of CB-5083, but eliminate the off-target effects on PDE6 (Roux et al. 2021), leading to a phase I clinical trial for treating acute myeloid leukemia or myelodysplastic syndrome (NCT04402541). In addition to ATP-competitive inhibitors, covalent inhibitor NMS-859 ($IC_{50} = 0.37 \mu M$), specifically modifying Cys522 in the D2 domain, and allosteric inhibitor NMS-873 ($IC_{50} = 30$ nM), binding to the linker-D2 domain to prevent subsequent conformational changes, were also reported (Magnaghi et al. 2013, Pan et al. 2021). Recently, another allosteric inhibitor, UPCDC30245, was also identified to bind at the D1–D2 linker domains of VCP, with IC_{50} of 27 nM (LaPorte et al. 2018). Development of next-generation inhibitors for VCP is still an active area of research. However, considering the diverse cellular functions of VCP, small molecular inhibitors for VCP may have side effects, especially for VCP-associated neurodegenerative disease, which normally needs long-term treatment to control the disease progression. Therefore, development of inhibitors that specifically disrupt VCP and its cofactor interaction may represent an alternative approach for VCP-based therapeutics.

Summary and Outlook

Despite the recent observation of VCP dodecamer state, many questions remain open. Particularly, it is still not known whether the dodecamer state exists inside the cell. Using the native PAGE assay, we analyzed mammalian cells with or without VCP overexpression, and confirmed the presence of the dodecamer species in the mammalian cell lysates containing overexpressed or endogenous VCP (Yu et al. 2021, 2022). However, more in situ studies will be required to validate the existence and define the cellular localization of VCP dodecamer. Secondly, it remains uncertain whether the dodecamer states have a function in the cell. The presence of nucleotides in the in vitro functional assays (ATPase and substrate unfolding assays) renders it challenging to differentiate the contribution of the dodecamer state from hexamers (Gao et al. 2022). Therefore, development of new in vitro and in vivo functional assays for VCP dodecamers is an urgent task. In this aspect, further characterization of the inhibitor-trapped VCP dodecamers would likely provide essential insights in the separation of the two oligomer states using small molecule compounds. Additionally, the cellular factors regulating dodecamer state formation also await to be characterized, such as PTMs and cofactors. In comparison with

wild-type VCP, the disease-causing R155H mutant was found with increased preference for the dodecamer state (Nandi et al. 2021). However, it is unclear whether this is true for other disease-causing mutants, and whether and how such property is linked to pathogenesis. As discussed above, a partial threading mechanism could operate for some substrates by the VCP dodecamer, particularly those destined to be recycled or reused in processes such as membrane fusion and autophagy. VCP dodecamerization may also potentially serve as a regulatory mechanism to reduce ATP consumption by VCP given its high cellular abundance. Further studies will be needed to delineate the role of VCP dodecamer in regulating diverse cellular functions.

Acknowledgements We thank National Institutes of Health R01CA069202 for financial support. We apologize to our colleagues that their work were not cited here due to length constraints.

References

- Anderson DJ, Le Moigne R, Djakovic S, Kumar B, Rice J, Wong S, Wang J, Yao B, Valle E, Kiss von Soly S, Madriaga A, Soriano F, Menon MK, Wu ZY, Kampmann M, Chen Y, Weissman JS, Aftab BT, Yakes FM, Shawver L, Zhou HJ, Wustrow D, Rolfe M (2015) Targeting the AAA ATPase p97 as an approach to treat cancer through disruption of protein homeostasis. *Cancer Cell* 28(5):653–665
- Bai Y, Yu G, Zhou HM, Amarasinghe O, Zhou Y, Zhu P, Li Q, Zhang L, Nguete Meke F, Miao Y, Chapman E, Tao WA, Zhang ZY (2023) PTP4A2 promotes lysophagy by dephosphorylation of VCP/p97 at Tyr805. *Autophagy* 19(5):1562–1581
- Banerjee S, Bartesaghi A, Merk A, Rao P, Bulfer SL, Yan Y, Green N, Mroczkowski B, Neitz RJ, Wipf P, Falconieri V, Deshaies RJ, Milne JL, Huryn D, Arkin M, Subramaniam S (2016) 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science* 351(6275):871–875
- Barthelme D, Sauer RT (2016) Origin and functional evolution of the Cdc48/p97/VCP AAA+ protein unfolding and remodeling machine. *J Mol Biol* 428(9 Pt B):1861–1869
- Briggs LC, Baldwin GS, Miyata N, Kondo H, Zhang X, Freemont PS (2008) Analysis of nucleotide binding to P97 reveals the properties of a tandem AAA hexameric ATPase. *J Biol Chem* 283(20):13745–13752
- Caffrey B, Zhu X, Berezuk A, Tuttle K, Chittori S, Subramaniam S (2021) AAA+ ATPase p97/VCP mutants and inhibitor binding disrupt inter-domain coupling and subsequent allosteric activation. *J Biol Chem* 297(4):101187
- Chou TF, Brown SJ, Minond D, Nordin BE, Li K, Jones AC, Chase P, Porubsky PR, Stoltz BM, Schoenen FJ, Patricelli MP, Hodder P, Rosen H, Deshaies RJ (2011) Reversible inhibitor of p97, DBE-Q, impairs both ubiquitin-dependent and autophagic protein clearance pathways. *Proc Natl Acad Sci USA* 108(12):4834–4839
- Chou TF, Li K, Frankowski KJ, Schoenen FJ, Deshaies RJ (2013) Structure-activity relationship study reveals ML240 and ML241 as potent and selective inhibitors of p97 ATPase. *ChemMedChem* 8(2):297–312
- Chou TF, Bulfer SL, Weihl CC, Li K, Lis LG, Walters MA, Schoenen FJ, Lin HJ, Deshaies RJ, Arkin MR (2014) Specific inhibition of p97/VCP ATPase and kinetic analysis demonstrate interaction between D1 and D2 ATPase domains. *J Mol Biol* 426(15):2886–2899
- Cooney I, Han H, Stewart MG, Carson RH, Hansen DT, Iwasa JH, Price JC, Hill CP, Shen PS (2019) Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. *Science* 365(6452):502–505

- Costantini S, Capone F, Polo A, Bagnara P, Budillon A (2021) Valosin-containing protein (VCP)/p97: a prognostic biomarker and therapeutic target in cancer. *Int J Mol Sci* 22(18):10177
- Dantuma NP, Hoppe T (2012) Growing sphere of influence: Cdc48/p97 orchestrates ubiquitin-dependent extraction from chromatin. *Trends Cell Biol* 22(9):483–491
- Davies JM, Brunger AT, Weis WI (2008) Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. *Structure* 16(5):715–726
- DeLaBarre B, Brunger AT (2003) Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains. *Nat Struct Biol* 10(10):856–863
- Deshaies RJ (2014) Proteotoxic crisis, the ubiquitin-proteasome system, and cancer therapy. *BMC Biol* 12:94
- Gao H, Li F, Ji Z, Shi Z, Li Y, Yu H (2022) Cryo-EM structures of human p97 double hexamer capture potentiated ATPase-competent state. *Cell Discov* 8(1):19
- Gite J, Milko E, Brady L, Baker SK (2020) Phenotypic convergence in Charcot-Marie-tooth 2Y with novel VCP mutation. *Neuromuscul Disord* 30(3):232–235
- Gonzalez MA, Feely SM, Speziani F, Strickland AV, Danzi M, Bacon C, Lee Y, Chou TF, Blanton SH, Weihl CC, Zuchner S, Shy ME (2014) A novel mutation in VCP causes Charcot-Marie-tooth type 2 disease. *Brain* 137(Pt 11):2897–2902
- Guo X, Sun X, Hu D, Wang YJ, Fujioka H, Vyas R, Chakrapani S, Joshi AU, Luo Y, Mochly-Rosen D, Qi X (2016) VCP recruitment to mitochondria causes mitophagy impairment and neurodegeneration in models of Huntington's disease. *Nat Commun* 7:12646
- Gupta A, Lentzsch AM, Siegel A, Yu Z, Chio US, Cheng Y, Shan SO (2023) Dodecamer assembly of a metazoan AAA(+) chaperone couples substrate extraction to refolding. *Sci Adv* 9(19):eadf5336
- Hanzelmann P, Schindelin H (2016) Structural basis of ATP hydrolysis and Intersubunit signaling in the AAA+ ATPase p97. *Structure* 24(1):127–139
- Hanzelmann P, Schindelin H (2017) The interplay of cofactor interactions and post-translational modifications in the regulation of the AAA+ ATPase p97. *Front Mol Biosci* 4:21
- Hoq MR, Vago FS, Li K, Kovaliov M, Nicholas RJ, Huryn DM, Wipf P, Jiang W, Thompson DH (2021) Affinity capture of p97 with small-molecule ligand bait reveals a 3.6 Å double-hexamer Cryoelectron microscopy structure. *ACS Nano* 15(5):8376–8385
- Huang S, Tang D, Wang Y (2016) Monoubiquitination of Syntaxin 5 regulates Golgi membrane dynamics during the cell cycle. *Dev Cell* 38(1):73–85
- Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ, Gibbs JR, Brunetti M, Gronka S, Wu J, Ding J, McCluskey L, Martinez-Lage M, Falcone D, Hernandez DG, Arepalli S, Chong S, Schymick JC, Rothstein J, Landi F, Wang YD, Calvo A, Mora G, Sabatelli M, Monsurro MR, Battistini S, Salvi F, Spataro R, Sola P, Borghero G, Consortium I, Galassi G, Scholz SW, Taylor JP, Restagno G, Chio A, Traynor BJ (2010) Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* 68(5):857–864
- Kilgas S, Ramadan K (2023) Inhibitors of the ATPase p97/VCP: from basic research to clinical applications. *Cell Chem Biol* 30(1):3–21
- LaPorte MG, Burnett JC, Colombo R, Bulfer SL, Alvarez C, Chou TF, Neitz RJ, Green N, Moore WJ, Yue Z, Li S, Arkin MR, Wipf P, Huryn DM (2018) Optimization of phenyl indole inhibitors of the AAA+ ATPase p97. *ACS Med Chem Lett* 9(11):1075–1081
- Lee JJ, Park JK, Jeong J, Jeon H, Yoon JB, Kim EE, Lee KJ (2013) Complex of Fas-associated factor 1 (FAF1) with valosin-containing protein (VCP)-Npl4-Ufd1 and polyubiquitinated proteins promotes endoplasmic reticulum-associated degradation (ERAD). *J Biol Chem* 288(10):6998–7011
- Leinonen H, Cheng C, Pitkanen M, Sander CL, Zhang J, Saeid S, Turunen T, Shmara A, Weiss L, Ta L, Ton T, Koskelainen A, Vargas JD, Kimonis V, Palczewski K (2021) A p97/Valosin-containing protein inhibitor drug CB-5083 has a potent but reversible off-target effect on Phosphodiesterase-6. *J Pharmacol Exp Ther* 378(1):31–41
- Liu S, Ye X, Liu W, Liu L, Li D, Lin Q, Wang T (2022) Cryo-EM structure of dodecamer human p97 in complex with NMS-873 reveals S(765)-G(779) peptide plays critical role for D2 ring oligomerization. *Biochem Biophys Res Commun* 601:146–152

- Magnaghi P, D'Alessio R, Valsasina B, Avanzi N, Rizzi S, Asa D, Gasparri F, Cozzi L, Cucchi U, Orrenius C, Polucci P, Ballinari D, Perrera C, Leone A, Cervi G, Casale E, Xiao Y, Wong C, Anderson DJ, Galvani A, Donati D, O'Brien T, Jackson PK, Isacchi A (2013) Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nat Chem Biol* 9(9):548–556
- Meyer HH (2005) Golgi reassembly after mitosis: the AAA family meets the ubiquitin family. *Biochim Biophys Acta* 1744(2):108–119
- Meyer H, Wehl CC (2014) The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis. *J Cell Sci* 127(Pt 18):3877–3883
- Meyer H, Bug M, Bremer S (2012) Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat Cell Biol* 14(2):117–123
- Mouyset J, Deichsel A, Moser S, Hoege C, Hyman AA, Gartner A, Hoppe T (2008) Cell cycle progression requires the CDC-48/UDF-1/NPL-4 complex for efficient DNA replication. *Proc Natl Acad Sci USA* 105(35):12879–12884
- Nandi P, Li S, Columbres RCA, Wang F, Williams DR, Poh YP, Chou TF, Chiu PL (2021) Structural and functional analysis of disease-linked p97 ATPase mutant complexes. *Int J Mol Sci* 22(15):8079
- Niwa H, Ewens CA, Tsang C, Yeung HO, Zhang X, Freemont PS (2012) The role of the N-domain in the ATPase activity of the mammalian AAA ATPase p97/VCP. *J Biol Chem* 287(11):8561–8570
- Pamnani V, Tamura T, Lupas A, Peters J, Cejka Z, Ashraf W, Baumeister W (1997) Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon *Thermoplasma acidophilum*. *FEBS Lett* 404(2–3):263–268
- Pan M, Yu Y, Ai H, Zheng Q, Xie Y, Liu L, Zhao M (2021) Mechanistic insight into substrate processing and allosteric inhibition of human p97. *Nat Struct Mol Biol* 28(7):614–625
- Papadopoulos C, Kirchner P, Bug M, Grum D, Koerver L, Schulze N, Poehler R, Dressler A, Fengler S, Arhzaouy K, Lux V, Ehrmann M, Wehl CC, Meyer H (2017) VCP/p97 cooperates with YOD1, UBXD1 and PLAA to drive clearance of ruptured lysosomes by autophagy. *EMBO J* 36(2):135–150
- Peters JM, Harris JR, Lustig A, Muller S, Engel A, Volker S, Franke WW (1992) Ubiquitous soluble Mg(2+)-ATPase complex. A structural study. *J Mol Biol* 223(2):557–571
- Ripstein ZA, Huang R, Augustyniak R, Kay LE, Rubinstein JL (2017) Structure of a AAA+ unfoldase in the process of unfolding substrate. *elife* 6:e25754
- Ritz D, Vuk M, Kirchner P, Bug M, Schutz S, Hayer A, Bremer S, Lusk C, Baloh RH, Lee H, Glatter T, Gstaiger M, Aebersold R, Wehl CC, Meyer H (2011) Endolysosomal sorting of ubiquitylated caveolin-1 is regulated by VCP and UBXD1 and impaired by VCP disease mutations. *Nat Cell Biol* 13(9):1116–1123
- Roux B, Vaganay C, Vargas JD, Alexe G, Benaksas C, Pardieu B, Fenouille N, Ellegast JM, Malolepsza E, Ling F, Sodaro G, Ross L, Pikman Y, Conway AS, Tang Y, Wu T, Anderson DJ, Le Moigne R, Zhou HJ, Luciano F, Hartigan CR, Galinsky I, DeAngelo DJ, Stone RM, Auberger P, Schenone M, Carr SA, Guirouilh-Barbat J, Lopez B, Khaled M, Lage K, Hermine O, Hemann MT, Puissant A, Stegmaier K, Benajiba L (2021) Targeting acute myeloid leukemia dependency on VCP-mediated DNA repair through a selective second-generation small-molecule inhibitor. *Sci Transl Med* 13(587):eabg1168
- Schuller JM, Beck F, Lossl P, Heck AJ, Forster F (2016) Nucleotide-dependent conformational changes of the AAA+ ATPase p97 revisited. *FEBS Lett* 590(5):595–604
- Song C, Wang Q, Li CC (2003) ATPase activity of p97-valosin-containing protein (VCP). D2 mediates the major enzyme activity, and D1 contributes to the heat-induced activity. *J Biol Chem* 278(6):3648–3655
- Stach L, Freemont PS (2017) The AAA+ ATPase p97, a cellular multitool. *Biochem J* 474(17):2953–2976
- Tang WK, Xia D (2016) Mutations in the human AAA(+) chaperone p97 and related diseases. *Front Mol Biosci* 3:79

- Tang WK, Odzorig T, Jin W, Xia D (2019) Structural basis of p97 inhibition by the site-selective anticancer compound CB-5083. *Mol Pharmacol* 95(3):286–293
- Twomey EC, Ji Z, Wales TE, Bodnar NO, Ficarro SB, Marto JA, Engen JR, Rapoport TA (2019) Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science* 365(6452):eaax1033
- van den Boom J, Meyer H (2018) VCP/p97-mediated unfolding as a principle in protein homeostasis and signaling. *Mol Cell* 69(2):182–194
- Vaz B, Halder S, Ramadan K (2013) Role of p97/VCP (Cdc48) in genome stability. *Front Genet* 4:60
- Watts GD, Wymer J, Kovach MJ, Mehta SG, Mumm S, Darvish D, Pestronk A, Whyte MP, Kimonis VE (2004) Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet* 36(4):377–381
- Wu X, Li L, Jiang H (2016) Doa1 targets ubiquitinated substrates for mitochondria-associated degradation. *J Cell Biol* 213(1):49–63
- Wu D, Liu Y, Dai Y, Wang G, Lu G, Chen Y, Li N, Lin J, Gao N (2023) Comprehensive structural characterization of the human AAA+ disaggregase CLPB in the apo- and substrate-bound states reveals a unique mode of action driven by oligomerization. *PLoS Biol* 21(2):e3001987
- Xu S, Peng G, Wang Y, Fang S, Karbowski M (2011) The AAA-ATPase p97 is essential for outer mitochondrial membrane protein turnover. *Mol Biol Cell* 22(3):291–300
- Xu Y, Han H, Cooney I, Guo Y, Moran NG, Zuniga NR, Price JC, Hill CP, Shen PS (2022a) Active conformation of the p97-p47 unfoldase complex. *Nat Commun* 13(1):2640
- Xu N, Lin Q, Tian H, Liu C, Wang P, Suen CM, Yang H, Xiang Y, Zhu G (2022b) Cryo-EM structure of human hexameric MCM2-7 complex. *iScience* 25(9):104976
- Ye Y, Meyer HH, Rapoport TA (2003) Function of the p97-Ufd1-Npl4 complex in retrotranslocation from the ER to the cytosol: dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains. *J Cell Biol* 162(1):71–84
- Yu G, Bai Y, Li K, Amarasinghe O, Jiang W, Zhang ZY (2021) Cryo-electron microscopy structures of VCP/p97 reveal a new mechanism of oligomerization regulation. *IScience* 24(11):103310
- Yu G, Bai Y, Zhang ZY (2022) Purification and cryo-EM structure determination of VCP/p97 dodecamers from mammalian and bacterial cells. *STAR Protoc* 3(2):101339
- Zhang X, Shaw A, Bates PA, Newman RH, Gowen B, Orlova E, Gorman MA, Kondo H, Dokurno P, Lally J, Leonard G, Meyer H, van Heel M, Freemont PS (2000) Structure of the AAA ATPase p97. *Mol Cell* 6(6):1473–1484
- Zhang Y, Li W, Chu M, Chen H, Yu H, Fang C, Sun N, Wang Q, Luo T, Luo K, She X, Zhang M, Yang D (2016) The AAA ATPase Vps4 plays important roles in *Candida albicans* hyphal formation and is inhibited by DBE. *Mycopathologia* 181(5–6):329–339
- Zhou HJ, Wang J, Yao B, Wong S, Djakovic S, Kumar B, Rice J, Valle E, Soriano F, Menon MK, Madriaga A, Kiss von Soly S, Kumar A, Parlati F, Yakes FM, Shawver L, Le Moigne R, Anderson DJ, Rolfe M, Wustrow D (2015) Discovery of a first-in-class, potent, selective, and orally bioavailable inhibitor of the p97 AAA ATPase (CB-5083). *J Med Chem* 58(24):9480–9497

Chapter 19

Oligomeric Structure of Yeast and Other Invertases Governs Specificity



Elena Jiménez-Ortega and Julia Sanz-Aparicio

Abstract Invertases, or β -fructofuranosidases, are metabolic enzymes widely distributed among plants and microorganisms that hydrolyze sucrose and release fructose from various substrates. Invertase was one of the earliest discovered enzymes, first investigated in the mid-nineteenth century, becoming a classical model used in the primary biochemical studies on protein synthesis, activity, and the secretion of glycoproteins. However, it was not until 20 years ago that a member of this family of enzymes was structurally characterized, showing a bimodular arrangement with a β -propeller catalytic domain, and a β -sandwich domain with unknown function. Since then, many studies on related plant and fungal enzymes have revealed them as basically monomeric. By contrast, all yeast enzymes in this family that have been characterized so far have shown sophisticated oligomeric structures mediated by the non-catalytic domain, which is also involved in substrate binding, and how this assembly determines the particular specificity of each enzyme. In this chapter, we will review the available structures of yeast invertases to elucidate the mechanism regulating oligomer formation and compare them with other reported dimeric invertases in which the oligomeric assembly has no apparent functional implications. In addition, recent work on a new family of invertases with absolute specificity for the α -(1,2)-bond of sucrose found in cyanobacteria and plant invertases is highlighted.

Keywords Invertases · β -Fructofuranosidases · Glycosidases · Oligomerization mechanisms · Substrate specificity · Yeasts · 3-D structure · X-Ray crystallography

E. Jiménez-Ortega
Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, Spain
e-mail: ejimenez@cnio.es

J. Sanz-Aparicio (✉)
Department of Crystallography and Structural Biology, Institute of Physical-Chemistry Blas Cabrera. CSIC, Madrid, Spain
e-mail: xjulia@iqf.csic.es

Abbreviations

AkFfase	<i>Aspergillus kawachii</i> β -fructofuranosidase
AtNIN2	<i>Arabidopsis thaliana</i> invertase
BIFFase	<i>Bifidobacterium longum</i> β -fructofuranosidase
BmSUC1	<i>Bombyx mori</i> invertase
FOS	Fructooligosaccharides
FperFfase	<i>Frischella perrara</i> β -fructofuranosidase
GH	Glycoside Hydrolases
InvA	<i>Anabaena</i> sp. Invertase A
InvB	<i>Anabaena</i> sp. Invertase B
RdINV	<i>Rhodotorula dairenensis</i> invertase
ScINV	<i>Saccharomyces cerevisiae</i> invertase
SoFfase	<i>Schwanniomyces occidentalis</i> β -fructofuranosidase
TmINV	<i>Thermotoga maritima</i> invertase
XdINV	<i>Xanthophyllomyces dendrorhous</i> invertase

Invertases: Applications, Sources, and Classification

The effort to promote a green economy is increasing exponentially, and enzymes are introduced as a sustainable and powerful alternative to displace the chemical pollutants in the industry. A large group of enzymes with a wide range of applications are the Glycoside Hydrolases (GHs), whose main activity is the hydrolysis of the glycosidic bond between two carbohydrates, or between a carbohydrate and a noncarbohydrate moiety (Cimini et al. 2016). They are widespread in several organisms and are divided into families based on amino acid sequence similarities (Henrissat 1991). To date, 189 families of GH have been described.

Invertases are metabolic enzymes with a great interest in biotechnology and basic research. They are classified as a type of β -fructofuranosidase (EC.3.2.1.26) due to their ability to catalyze the hydrolysis of the β -D-fructofuranosyl residues at the nonreducing end of various substrates. The main process they carry out is the conversion of the disaccharide sucrose to an equimolar mixture of fructose and glucose (Kulshrestha et al. 2013). The resulting mixture is known as invert sugar syrup because of the inversion in the optical properties, from the dextrorotatory D-sucrose to the dextrorotatory D-glucose and the strongly levorotatory fructose (Alegre et al. 2009). As a result, the invertase product (fructose) is sweeter than sucrose and has lower crystallinity at higher concentrations, which can be preferable in many biotechnological applications. Another important benefit in the food industry is their use to ensure the confectionery products remain fresh and soft after a long period of storage (Du et al. 2013) or their application in the preparation of jams, candies, soft chocolates and biscuits, creams, powder milk for infants, or artificial honey. They can also be used in paper industries, cosmetics, plasticizing agents, production of ethanol and organic acids such as lactic acid, and in electrodes

for sucrose detection (Lincoln and More 2017). Furthermore, invertase activity is commonly used in the alcohol industry to ferment cane molasses into ethanol (Du et al. 2013), among other applications.

In addition, invertases may be exploited in the synthesis of fructooligosaccharides (FOS). These molecules are well-known as prebiotics, being widely consumed by beneficial human gut bacteria such as *Lactobacilli* and *Bifidobacteria* (Hidaka et al. 1986; Tanno et al. 2021). Different types of FOS series have been characterized depending on the linkage between the monosaccharide moieties: ¹F-FOS, derived from inulin-type structure and containing β -(2,1)-linked fructose units (1-kestose or nystose); ⁶F-FOS, containing β -(2,6)-linked fructose units, with a levan-type structure (6-kestose); and ⁶G-FOS or neo-FOS, where a β -(2,6)-link connects fructose chains to a glucosyl moiety (neokestose or neonystose). Compared with ¹F-FOS, β -(2,6)-linked FOS have shown enhanced prebiotic properties and chemical stability.

Depending on the optimum pH, these enzymes are classified into acid invertases (optimum pH of 4.0–5.5) and alkaline/neutral invertases (optimum pH of 6.5–8.0) (Sturm 1999). The specific function of both isoforms is not well-known, but they appear to control the entry of sucrose into different utilization pathways (Alegre et al. 2009). Acid invertases have been the most extensively studied and are widespread in plants, fungi, yeast, and bacteria, a few having been found in animals. However, much less is known about alkaline invertases, which have been only reported from photosynthetic bacteria (cyanobacteria) and plants (Xie et al. 2016; Zhou et al. 2016).

The enzymatic hydrolysis of glycosidic bonds has two possible stereochemical outcomes, with inversion or retention of the anomeric configuration (Davies and Henrissat 1995) depending on whether the reaction proceeds in one or two steps (Fig. 19.1). Acid invertases can utilize sucrose and other β -fructose-containing oligo- and polysaccharides such as raffinose, kestose, and inulin, among others, as substrate (Trollope et al. 2015). Additionally, they are able to carry out the inverse reaction transferring fructose to an acceptor sugar substrate (transglycosylation). The latter reaction occurs in the presence of high sucrose concentrations and produces FOS with either β -(2,1) or β -(2,6) bond linkages, which can be used as prebiotics (short-chain FOS). On the other hand, alkaline/neutral invertases specifically catalyze the hydrolysis of sucrose. Moreover, these two kinds of enzymes show differences in their cellular localization. Whereas acid invertases are mainly localized in cell walls and vacuoles, alkaline/neutral invertases are generally present in the cytosol and organelles such as mitochondria and chloroplast.

According to their family classification, invertases present two different molecular architectures (Lombard et al. 2014; Garron and Henrissat 2019). In general, acid invertases belong to the glycoside hydrolase family 32 (GH32) and are rich in β -sheets. In contrast, the recently discovered alkaline/neutral invertases from glycoside hydrolase family 100 (GH100) are mainly composed of α -helices. As it is known, the quaternary structure of enzymes is determined by noncovalent bonds between complementary surface hydrophobic and hydrophilic regions on the polypeptide monomer subunits. Different acidic and basic side chains can form salt bridges, or covalent stabilization by interchain disulfide bonds can promote their

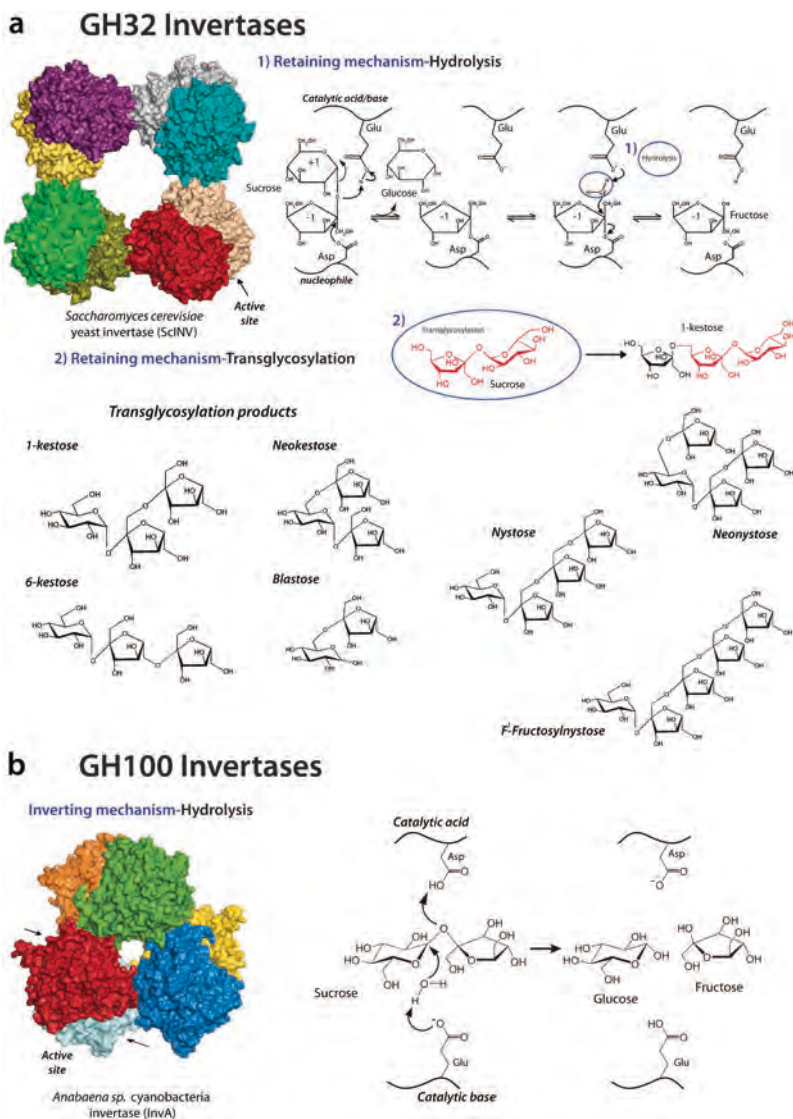


Fig. 19.1 Catalytic mechanisms of invertases. **(a)** A GH32 yeast invertase with retaining mechanism, showing the sucrose processing in two steps. A first attack of the anomeric carbon by the nucleophile catalytic residue frees glucose, generating a covalent enzyme-fructose intermediate. Subsequently, the acid/base catalytic residue deprotonates a water molecule that performs a second nucleophilic attack freeing fructose or, alternatively, activates a secondary sugar molecule (acceptor) generating a longer oligosaccharide (transfructosylation). Figure adapted from (Lammens et al. 2009). By definition, the hydrolysis of the sugar takes place at subsites $-1/+1$ within the enzyme. Common transglycosylation products of invertases and β -fructofuranosidases using different acceptor sugars are shown. Figure adapted from (Zambelli et al. 2014). **(b)** A cyanobacteria GH100 invertase with inverting mechanism, illustrating the hydrolysis of the sucrose in a single displacement step and the anomeric carbon inversion. Figure adapted from (Xie et al. 2016)

particular oligomeric state. Depending on the differences in sequence composition, there are large divergences in their structural features and, consequently, in their assemblies. The different structural elements that contribute in the oligomeric pattern of invertases are discussed in the following sections.

The Bimodular Folding of the GH32 Family

Although invertase activity was one of the earliest described in the mid-nineteenth century, it was not until 2004 that the first three-dimensional crystal structure of a GH32 member, the *Thermotoga maritima* invertase (TmINV), was determined (Alberto et al. 2004). This thermostable enzyme had been previously identified as hydrolyzing sucrose to fructose and glucose and also producing fructose from various substrates such as raffinose and inulin, among others. The 1.9 Å resolution crystal structure showed a catalytic domain folded into a five-bladed β -propeller, each blade made of four (A to D) antiparallel β -strands strongly twisted, which generates an angle of 90° between the first and last β -strands within each blade (Fig. 19.2). This β -propeller domain adopts the classical “W” topology, which surrounds a central negatively charged active site cavity, where an aspartate is the catalytic nucleophile and a

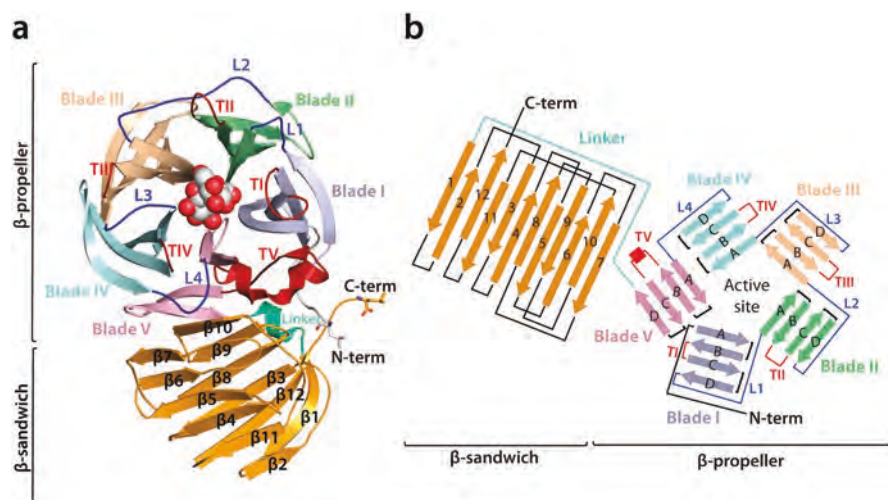


Fig. 19.2 Structure of *Thermotoga maritima* invertase (TmINV). (a) Cartoon representation, and (b) topology scheme, illustrating the bimodular arrangement of TmINV, with the catalytic β -propeller domain showing each blade in a different color, and the β -sandwich domain in orange (PDB code 1UYP). The active site is surrounded by the loops connecting the five blades (L1–L4), in blue, and the turns connecting the β -strands B to C within each blade (TI–TV), in red. By contrast, the loops connecting β -strands A to B (ab) and C to D (cd) are at the opposite face of the propeller. The raffinose molecule bound to the inactivated TmINV (PDB code 1W2T) is represented as spheres

glutamate the acid/base. The catalytic tunnel is open towards the molecular surface and presents a topology fully consistent with the exo-mode of action on the fructose polymer inulin. The presence of insertions is common in this type of β -propeller, with highly variable loops surrounding the catalytic site that connect the different blades (L1–L4) and the turns linking B to C β -strands within each blade (TI–TV).

This β -propeller structure was similar to other glycosidases described from GH43 and GH68 families. However, an extra C-terminal β -sandwich domain, connected by a short ten-residue linker to the catalytic domain, was additionally described (Fig. 19.2). This β -sandwich included two sheets, each composed of six β -strands (Alberto et al. 2004). This bimodular arrangement provided a template not only for invertases but also for all members within family GH32 (including inulinases, levanases, exo-inulinases, sucrose: sucrose 1-fructosyltransferases, and fructan: fructan 1-fructosyltransferases). In general, the β -propeller scaffold is highly conserved, whereas the similarity is low within the β -sheet module, to which a functional role could not be assigned.

Yeast Invertases Are Oligomeric Proteins

Fructans and related compounds play a key role in the metabolism and energy exchange in organisms as plants, where they may be employed as storage polymers. Several microorganisms use them as regular energy and carbon sources by the expression of hydrolytic enzymes. In addition, invertases are essential enzymes for sugar metabolism in plants and yeast, as sucrose plays a paramount role in signaling, being involved in development and stress response. They are classified as acid invertases and belong to the GH32 family that displays the canonical bimodular folding described for TmINV (Fig. 19.2). The substrate specificity of this set of enzymes is mainly regulated by structural elements at three different levels: loop variability, glycosylation pattern, and oligomerization arrangement.

First, the length and the amino acid composition of the loops L1–L4 and TI–TV (Bujacz et al. 2011), which are surrounding the catalytic pocket, have a direct effect in substrate recognition and defining the specificity unique to each enzyme. Moreover, eukaryote invertases may undergo posttranslational modifications, such as N- or O-glycosylation. Glycans are often found to preserve the protein against thermal and chaotropic denaturation (Kern et al. 1992). Indeed, the glycosylation pattern of fungal, yeast, and plant GH32 invertases has been occasionally described to block the catalytic cleft. Thereby, it may undertake a negative effect on the binding of long substrates. Furthermore, most of the invertases deposited in the PDB have been recombinantly expressed in other model yeasts such as *Saccharomyces cerevisiae* or *Pichia pastoris*, or even in *Escherichia coli* that do not present this type of modification. As discussed below, the pattern and the glycosylation degree may have an effect on activity, and has revealed a key role in the oligomerization state of the yeast enzymes.

An additional and essential element controlling the substrate specificity found only in yeast invertases is the oligomerization assembly. Specifically, yeast invertases have been characterized as multimeric, the dimer being the minimum functional unit described. Particularly, two different oligomeric states have been structurally depicted until now, homodimers and homooctamers (Fig. 19.3). In these

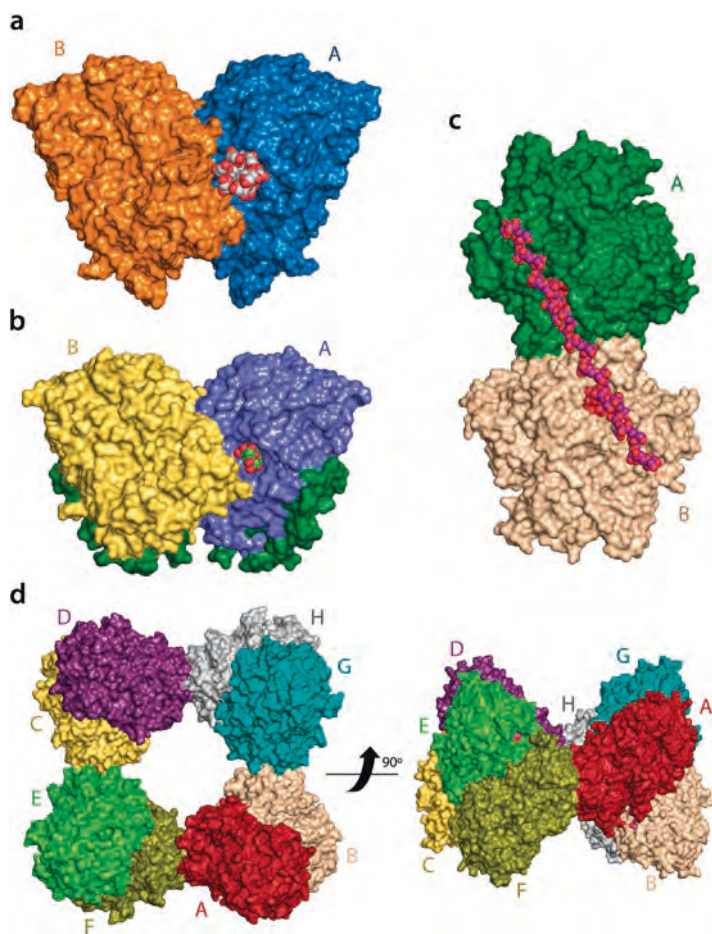


Fig. 19.3 Oligomeric assemblies and substrate specificity of structurally known yeast invertases. (a) *Schwanniomyces occidentalis* β -fructofuranosidase, SoFfase, with a modeled polymeric inulin chain based on its crystal structure in complex with inulin (PDB code 3U14). (b) *Rhodotorula dairenensis* invertase, RdINV, complexed with the trisaccharide raffinose (PDB code 8BEU); the platform created by the inserts is highlighted in green. (c) *Xanthophyllomyces dendrorhous* invertase, XdINV, complexed with the tetrasaccharide nystose (PDB code 5FMD), and with a modeled β -(2,6) fructan chain linking the two bound nystoses; this model may represent the binding of a putative mixed-type fructan polymeric substrate (graminan). (d) Two views of the *Saccharomyces cerevisiae* invertase octamer, ScINV (PDB code 4EQV). Each subunit is represented by a different color in the panels

arrangements, the C-terminal β -sandwich domain, which plays a fundamental role in the oligomerization pattern, is located in the vicinity of the active pocket entrance and controls the binding of substrates. Furthermore, these assemblies have been revealed to be essential in substrate recognition (Álvaro-Benito et al. 2010), regulating their functionality and biological activity, showing how these proteins have developed particular molecular mechanisms to govern fructan metabolism in yeasts (Fig. 19.3).

***Schwanniomyces occidentalis* β -Fructofuranosidase, the Butterfly Model**

The extracellular invertase or β -fructofuranosidase from the nonconventional yeast *Schwanniomyces occidentalis* (SoFfase) presents hydrolytic activity on small fructooligosaccharides such as sucrose, 1-kestose, and nystose, but also a remarkably high activity on the polymer inulin. Moreover, it produces the trisaccharides 6-kestose and 1-kestose by transfructosylation in a 3:1 rate and catalyzes the transference of a fructosyl residue to other biomolecules, such as polyphenols or alditols, to form other glycoconjugates.

Previous studies reported the existence of different oligomeric assemblies depending on the pH, but the unique homodimer state of SoFfase was confirmed by ultracentrifugation. The structure of this glycoprotein, the first yeast invertase reported, was determined in 2009 at 2.9 Å by X-ray crystallography (Álvaro-Benito et al. 2010), its N-deglycosylation treatment with EndoH improving later the diffraction data to 1.9 Å (Álvaro-Benito et al. 2012). A 72 kDa monomer was previously calculated from its amino acid sequence, but the crystal structure showed a 160 kDa dimer, in which two units are associated with a “butterfly” shape to form a tight dimer related by non-crystallographic twofold symmetry and burying 11% of its total molecular surface (Fig. 19.4). Each monomer presents the arrangement common to GH32 enzymes with a catalytic five-bladed β -propeller (residues 24–356) and a C-terminal β -sandwich (residues 365–535) linked by a short 10-residue segment. It also displays five N-glycosylation sites (Asn72, Asn119, Asn219, Asn334, and Asn394), four of them located in the catalytic and one in the β -sandwich domain. A detailed comparison of the catalytic domain with those of previously known GH32 structures defined the regions that are essential in determining substrate specificity among the family members, but also revealed an exceptional active site built up by both subunits.

The dimer dimension is $95 \times 65 \times 64$ Å³, with a total surface area of 36,680 Å², and a buried surface of 3990 Å². An inspection of its interface shows that 60% of the contacts are produced by the interaction of the two catalytic domains, mostly through blade IV from each unit, which is consistent with this blade IV containing the less conserved sequence as compared to GH32 enzymes. This region contains the only two hydrogen bonds that link the main chains of both subunits within the dimer. Other polar links between both catalytic domains are made through loops L3 and c5d5, and through D4

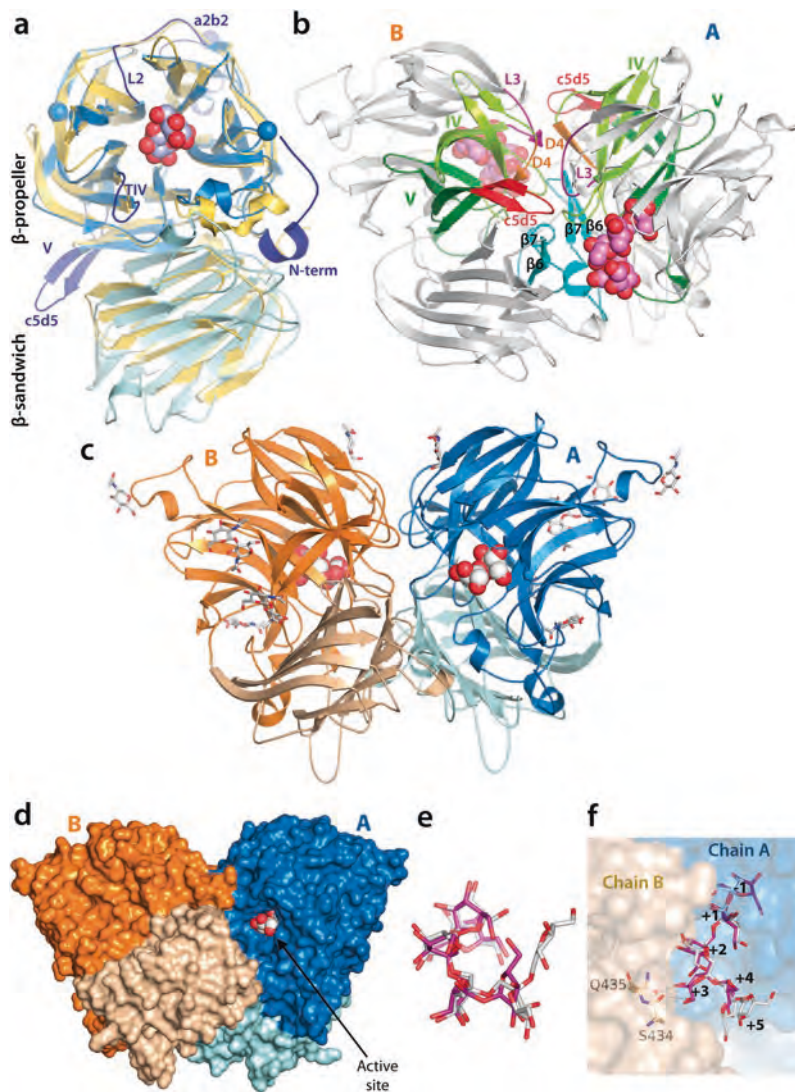


Fig. 19.4 *Schwanniomyces occidentalis* β-fructofuranosidase (SoFfase). (a) Superimposition of SoFfase subunit (PDB code 6S1T, blue) onto the *Thermotoga maritima* invertase (PDB code 1UYP, yellow), highlighting the most different loops of SoFfase at the catalytic domain in a different tone. A sucrose molecule found in the SoFfase crystal is shown as spheres; the N-glycosylation sites are marked with a sphere. (b) The SoFfase dimer is represented in gray cartoon with blade IV and blade V, shaping the interface, in light and dark green, respectively. Loops participating in the interface atomic contacts are in different colors. A fructosylinystose molecule bound to the E230A mutant is in violet spheres (PDB code 3U75). (c) Cartoon and (d) surface representations of the SoFfase dimer, showing subunit A in blue and subunit B in orange tones. In each monomer, the β-propeller is in darker tone than the β-sandwich domain. A sucrose molecule is in spheres. (e) and (f) Details of the active site of SoFfase inactive mutants, with bound fructosylinystose (PDB code 3U75) or inulin (PDB code 3U14), in violet and gray sticks. Both substrates adopt the same conformation upon binding, the sugar unit at subsite +3 making direct polar links to the adjacent subunit

β -strand that is markedly shortened (Fig. 19.4b). Furthermore, the catalytic pocket of one monomer is surrounded by loops from the β -sandwich domain of the other, and thus 16% of the total contacts within the dimer interface arise from the interaction between these two domains from both subunits, through a net of polar interactions stabilizing the interface and shaping the ground of a 25 Å deep cavity. The remaining interface is made by direct interaction between the two β -sandwich domains.

The particular arrangement of the dimer, with the active site cleft shaped by both subunits, and the involvement of the β -sandwich domain in the interface (Fig. 19.4c, d), suggested a direct role of this non-catalytic domain in substrate binding. This issue was first explored by mutagenesis, which showed an effect on substrate binding or the catalytic rate on long substrates upon changes in residues from the β -sandwich domain identified at the interface (Álvaro-Benito et al. 2010). The experimental evidence of the direct role of the non-catalytic domain in substrate binding was obtained in 2012 by crystallographic analysis of two complexes of inactivated SoFfase with fructosyl-nystose or inulin (Álvaro-Benito et al. 2012). This study showed the intricate net of interactions keeping the substrate tightly bound at the active site, with up to five clearly defined subsites, the sugar unit located at subsite +3 being recognized by direct polar interaction with the β -sandwich domain of the adjacent subunit within the dimer. Moreover, despite the fact that long fructooligosaccharide chains are highly flexible, the constricted active site accommodates both substrates into the SoFfase cavity in a very precise conformation that, in fact, closely resembles the helical conformation of the polymer inulin (Figs. 19.3a and 19.4e, f), explaining the dual invertase/inulinase activity of SoFfase, which is in contrast to its GH32 homologues.

Rhodotorula dairenensis Invertase, a Highly Glycosylated Version of the Butterfly Dimer

The invertase from *Rhodotorula dairenensis* (RdINV) is a large highly glycosylated enzyme that presents a hydrolytic preference for short substrates such as sucrose, 1-kestose, nystose, and raffinose, mainly producing 6-kestose by trans-fructosylation, among other FOS. Both native and recombinant RdINV expressed in *Pichia pastoris* show a similar molecular mass of 170 kDa, approximately, which is far from the theoretical 70 kDa calculated based on sequence (675 amino acids). Moreover, it shows an unusual 168-amino acid elongation at its N-terminus with a unique Ser-/Thr-rich region predicted as being hyper-O-glycosylated. This elongation is only present in two other putative *Rhodotorula* invertases (*Rhodotorula toruloides* and *Rhodotorula taiwanensis*, with 60–70% sequence identity), but not in other homologues. Deglycosylation strategies combining EndoH/mannosidase treatment verified that more than 50% of the total mass was associated with glycans, in which only 15% was assigned to N-glycosylation.

The recently determined crystal structure of the recombinant RdINV at 2.05 Å (Fig. 19.5) revealed a similar fold to the other GH32 members (Jiménez-Ortega et al. 2022). This study also confirmed the presence of at least 13 N- and 6 O-attached glycosylation sites, the O-glycosylation sites being concentrated at the N-terminus of its

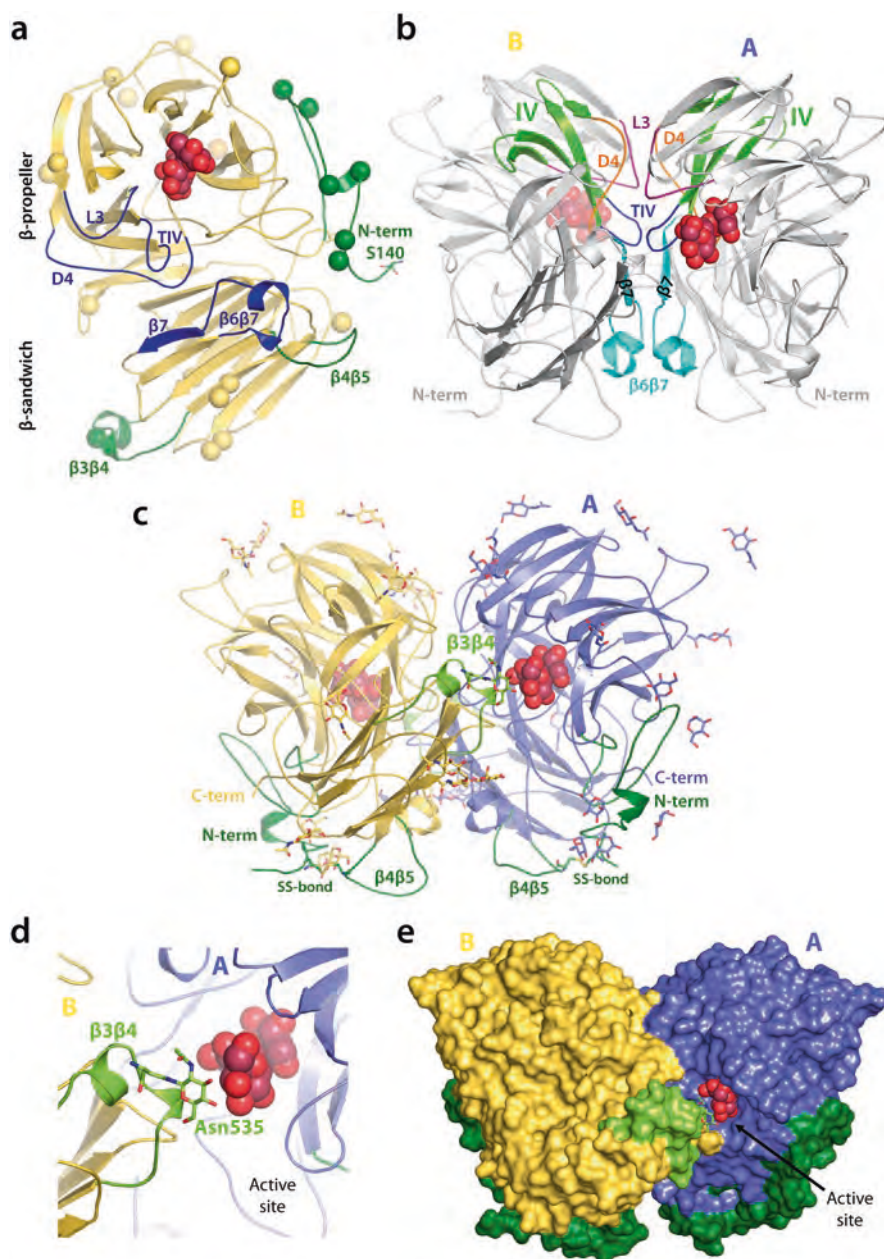


Fig. 19.5 *Rhodotorula dairenensis* invertase (RdINV). (a) Overall folding of the subunit (PDB code 8BEU), with most different regions with respect to SoFase in green, and the interacting loops at the interface in blue. The RdINV glycosylation sites and the raffinose molecule bound at the active site are represented as spheres. (b) Cartoon representation of the RdINV homodimer in gray, with the loops involved in the dimer interface represented in different colors. (c), (d), and (e) cartoon, detail of the interface and molecular surface of the dimer, with the two subunits colored yellow and slate, and the N-terminal region and relevant protruding loops highlighted in green

sequence. The structural comparison of RdINV monomer with the closest homologues from the GH32 family (SoFfase and ScINV) reveals that, apart from the N-terminal extension, greatest discrepancies are concentrated in the unique $\beta 3\beta 4$ (Asn525-Ser544) and $\beta 4\beta 5$ (Pro551-Pro566) inserts within the non-catalytic domain (Fig. 19.5a).

The assembly of RdINV closely resembles that reported for SoFfase. It is a dimer with dimensions $90 \times 60 \times 60 \text{ \AA}^3$, with a total surface area of $35,470 \text{ \AA}^2$, and a buried surface of $4010 \text{ \AA}^2$. The interface between monomers is mostly formed by blade IV (β -strand D4, loops L3 and TIV) from each catalytic domain, and loop $\beta 6\beta 7$ from their corresponding β -sandwich domains (Fig. 19.5b) through 12 hydrogens bonds. The dimer interaction shares a high similarity with the *butterfly* model described in SoFfase. However, the RdINV catalytic cleft is outlined by part of the long $\beta 3\beta 4$ loop insert from the adjacent subunit that, in addition, includes a N-glycosylation site at Asn535 (Fig. 19.5c, d, e). This contours a rather protruding region close to the active site, which is in contrast to SoFfase that lacks the $\beta 3\beta 4$ insert and shows markedly retracted loops, therefore enabling inulin binding. The occluded active site of RdINV is compatible with its observed activity against short substrates, and the presence of glycans at the Asn535 N-glycosylation site could possibly suggest some additional mechanism of specificity regulation of the RdINV activity in vivo.

Although the first 139 residues of the sequence (including the signal peptide) were not visible in the crystal structure, the last portion of the RdINV N-terminal elongation was clearly ordered and making a peculiar arrangement with the corresponding $\beta 4\beta 5$ loop inserts within the dimer (Fig. 19.5c, e). These two unique structural elements are connected by a di-sulfur bridge linking Cys143 to Cys555, and, apart from other polar interactions, additional links are expected through the O-glycans attached to Thr146 and Ser147, in which single mannose units have been observed in the crystal structure even after deglycosylation treatment of the enzyme. The polypeptide chain before Ser140, which is predicted as a structurally disorder region, is apparently pointing out from the protein towards the solvent and, through the analysis of a truncated variant, was found to be not essential for activity or thermal stability. Therefore, its high O-glycosylation pattern suggests that it may have a signaling/trafficking role, while the platform formed by the 140–160 regions and the $\beta 4\beta 5$ loops may configure a molecular platform to target the dimeric protein to its receptor cellular compartment, or possibly to the extracellular matrix, a task assisted by the high number of O-attached glycan chains.

Xanthophyllomyces dendrorhous Invertase, an Elongated Dimer Mediated by N-glycosylation

Xanthophyllomyces dendrorhous invertase (XdINV), an atypical highly glycosylated extracellular enzyme, displays hydrolytic activity against disaccharides mostly composed of glucose such as sucrose, maltose, isomaltose, or trehalose. It also catalyzes the synthesis short-chain FOS, especially of the neo-FOS series like neokesotose and neonystose, and the fructosylation of hydroquinone or hidroxytirosol

polyphenols, among other biomolecules of interest. The theoretical molecular weight of the monomer is 72 kDa, corresponding to 665 amino acids, but the biochemistry characterization revealed a 320–380 kDa homodimer with 60–70% being N-attached glycans. After EndoH treatment, the native molecular weight per monomer decreases to 100 kDa, which enabled crystal structure resolution up to 1.81 Å (Ramírez-Escudero et al. 2016). The monomer folding exhibits the expected bimodular arrangement common to GH32 members. However, two distinctive structural features of this yeast enzyme are the presence of two long extensions at its N- and C-terminal regions, respectively, and the existence of up to 17 modeled N-glycosylation sites (Fig. 19.6a).

Two identical subunits, related by a crystallographic twofold axis, make a flat and elongated dimer with dimensions $135 \times 75 \times 45 \text{ Å}^3$ and a total molecular surface of $43,733 \text{ Å}^2$ (Fig. 19.6b, c). Although both the catalytic and β -sandwich domains contribute to dimerization, the created interface is considerably smaller (1037 Å^2) than that observed in the *butterfly* dimer, burying only 2.4% of the total molecular surface. This interaction is focused on blades I and II, through loops L1 (residues 135–141), TII (residues 178–182), and L2 (residues 200–210), the β -sandwich domain (residues 650–662) contributing through the extension at its C-terminal region. Moreover, only 20 residues are 50% buried at the interface, with only 4 hydrogen bonds linking the subunits.

The novel trait found in the XdINV dimer is the essential contribution of glycans to oligomer assembly and activity. Thus, two well-ordered glycan chains were found in the crystal structure after the EndoH treatment at positions Asn58, within the N-terminal extension, and Asn107, located at TI loop within blade I (Fig. 19.6c), which are directly involved and extend the dimerization interface. The Asn107 glycans are making a network of intermolecular hydrogen bonds with L1 and L2 from each subunit (Fig. 19.6c), while the Asn58 glycan chain is packed over each C-terminal region, incrementing the interface area to almost sixfold. The relevance of these glycans was evaluated by analysis of the activity and thermal stability of Asn $\rightarrow$ Ser mutants, which showed an absolute loss of activity and an undetectable protein by SDS in the N107S mutant, and reduced catalytic efficiency and thermostability (by 10°C) upon the N58S change. Therefore, while Asn58 glycosylation makes a significant contribution to protein stabilization, the glycans N-attached to Asn107 are essential for dimer integrity.

Consequently, both, the N- and C-terminal extensions and the Asn58 and Asn107 N-attached glycans, seem to determine the unusual oligomer pattern of XdINV. The concurrence of all these structural elements is unique to this enzyme sequence, which shows similarity levels of only 20–25% identity with respect to other GH32 enzymes, including those from yeast. As a result, the dimer topology is markedly different to that found in SoFfase and RdINV, but, still, the β -sandwich domain keeps a key role in dimer formation and active site building. Furthermore, and contrarily to SoFfase, the two active sites of XdINV are located at the same face of the dimer and seem to be communicated by a 30 Å long crevice. This particular arrangement might suggest the putative ability of the XdINV dimer to accommodate a polymeric long substrate of the mixed-fructan type (Figs. 19.3c and 19.6d), with a

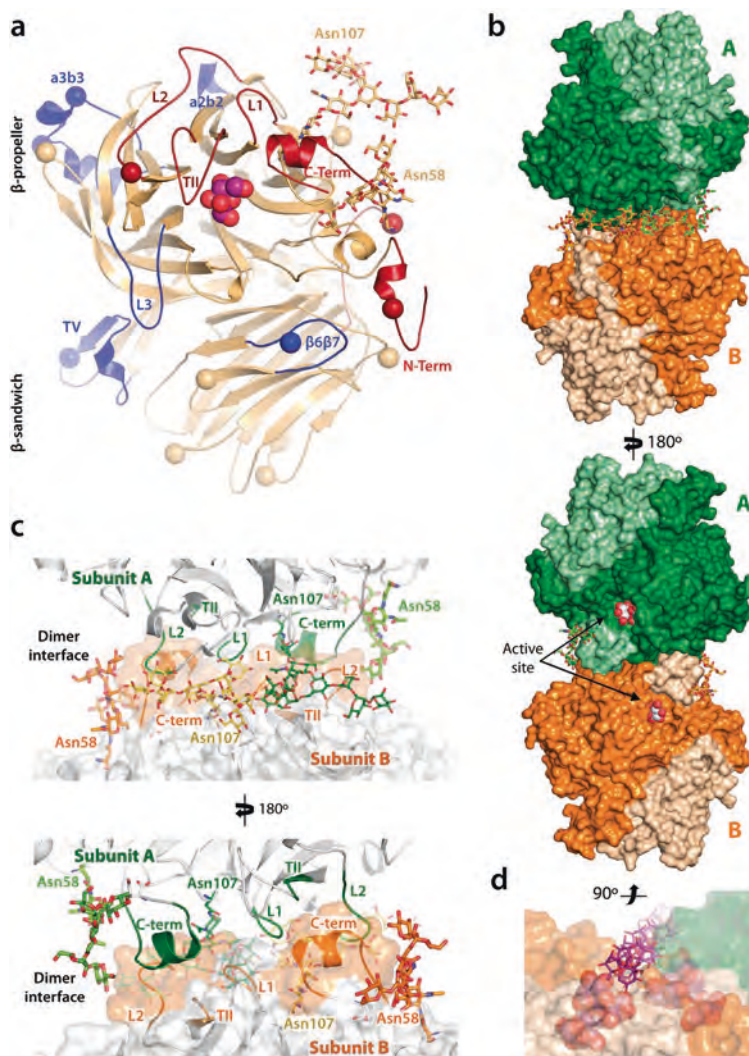


Fig. 19.6 *Xanthophyllomyces dendrorhous* invertase (XdINV). (a) Overall folding of the subunit (PDB code 5FIX) with most different loops as compared to its homologues in blue, and the interacting regions at the interface in red. The XdINV N-glycosylation sites are in spheres and two relevant glycan chains are represented in wheat sticks. The fructose moiety of sucrose bound at the active site is shown in pink spheres. (b) Two views of the molecular surface of the XdINV dimer, showing the sucrose at both active sites as spheres, and the four glycan chains attached to Asn58 and Asn107 from both subunits, which expand the interface. The catalytic domains are highlighted with brighter colors within each monomer. (c) Zoom of the XdINV dimer interface in the same views than B, representing the glycan chains in sticks and the interacting regions in green (subunit A) and orange (subunit B). The transparent molecular surface is shown in subunit B. (d) A close view of the dimer, showing the two nystose molecules bound to the inactivated D80A mutant as spheres (PDB code 5FMD), and with a modeled β -(2,6)-linked fructose chain, in sticks, manually docked into the crevice linking the two active sites. This model may represent the putative binding of a graminan-type substrate, made of β -(2,6)-linked fructosyl units with branches of β -(2,1)-linked fructoses, therefore illustrating a potential cooperativity of the two active sites within the dimer

β -(2,6)-linked fructose chain containing β -(2,1) branches, characteristic of graminans. In agreement with this hypothesis, XdINV showed a small but significant hydrolase activity on fructan polymers, and only a moderate activity on inulin, and also is able to hydrolyze agave fructans, possibly suggesting that these mixed fructans are the targeted substrate of XdINV in its natural habitat. As said above, fructans are essential in plant metabolism, and fungi might have evolved to interact with these biomolecules. Furthermore, microbial fructans have been shown to enhance cellular resistance to environmental stress, which might be related with the fact that the yeast *Xanthophyllomyces dendrorhous* has been isolated from soil samples collected at cold ecosystems. Therefore, although the exact physiological role of XdINV is difficult to envisage, the peculiar topology of its dimeric molecular assembly is a remarkable and distinctive trait within GH32 enzymes, suggesting potential functional implications.

***Saccharomyces cerevisiae* Invertase, a High Level of Structural Complexity Derived from Dimers**

The invertase from *Saccharomyces cerevisiae* (ScINV), commonly known as baker's yeast, is one of the classical and best-studied model enzymes in early biochemistry (Kulshrestha et al. 2013), the whole theory of enzyme kinetics being based on experimental results obtained with yeast invertase. In addition of its historical importance, it has been extensively applied in the industrial confectionary, in the fermentation of cane molasses into ethanol or in the synthesis of fructooligosaccharide (FOS) prebiotics, among others. The yeast produces two invertase isoforms: a heavily glycosylated secreted protein, which represents an evolutionary advantage to avoid the need of sucrose import to the cell, and a non-glycosylated intracellular form. Both are encoded by the same gene but are transcribed as two different mRNAs of different length, translated into polypeptides of different size.

Over the past four decades, several techniques have been applied in order to unravel the complex quaternary structure of this protein. In 1987, electron micrograph studies of both internal and secreted ScINV (Esmon et al. 1987; Tammi et al. 1987) were pioneer in determining its association pattern, described by different oligomeric states as dimers, tetramers, hexamers, and octamers, the last being predominant. Moreover, secreted octamers appeared slightly open on one side, whereas intracellular octamers seemed mostly closed. The variability in quaternary structure was also detected by gel filtration and analytical ultracentrifugation for both the cytosolic and secreted isoforms, both giving different active oligomeric forms. Also, selective release of dimeric glycosylated invertase from intact yeast cells suggested that oligomerization helps in retaining the enzyme in the periplasmic space.

A recombinant *Escherichia coli* isoform, without the 5' end of the signal peptide gene, was determined in 2012 by X-ray crystallography at 3.3 Å (Sainz-Polo et al. 2013), Fig. 19.7). The data revealed a subunit divided into two domains, similarly to GH32 family members, with a catalytic β -propeller domain composed of five blades of four β -strands (residues 1–334), and a β -sandwich domain with two six-stranded

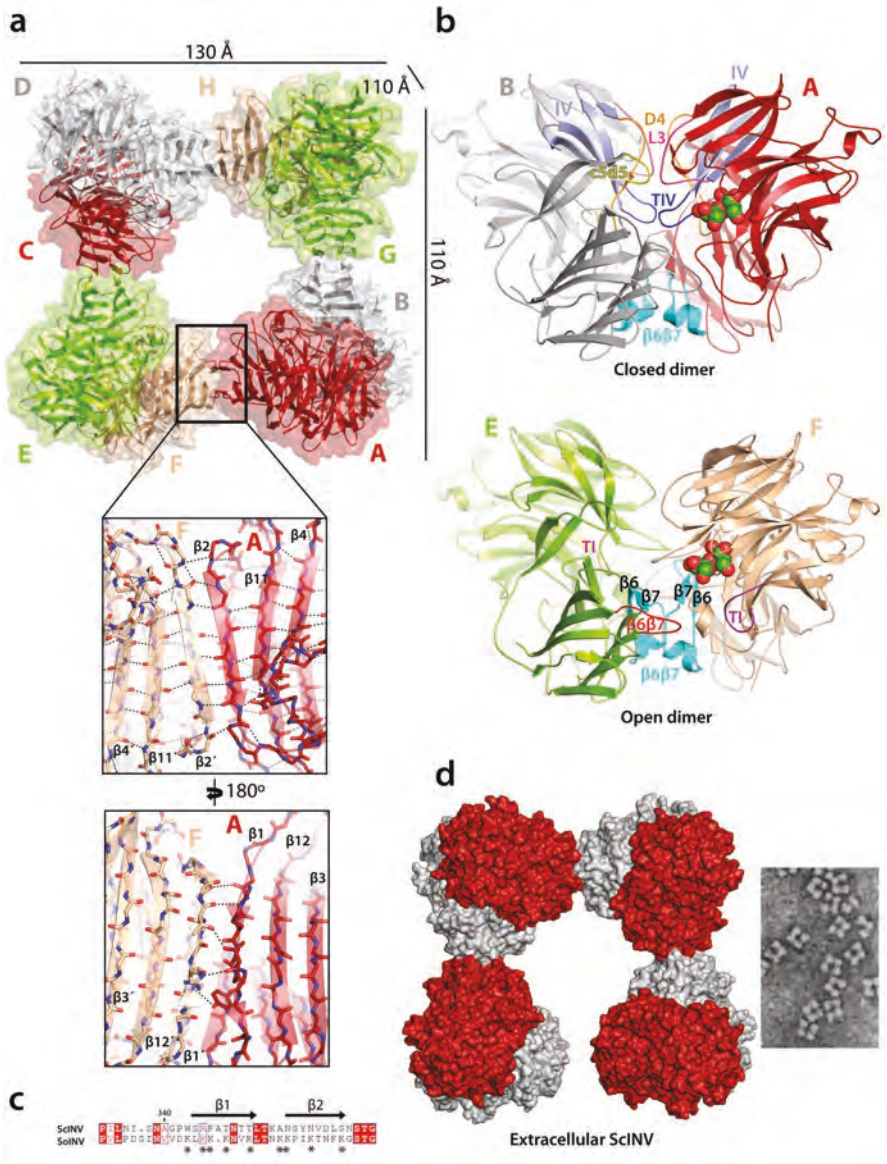


Fig. 19.7 *Saccharomyces invertase* (ScINV). (a) The octamer is a tetramer of alternating closed (red and gray) and open (green and beige) dimers (PDB code 4EQV); the inset shows the interface between both types of dimers at the internal and external walls within the octamer, made by extension of the corresponding β -sheets and mediated by strands $\beta 1$ and $\beta 2$ from adjacent subunits. (b) Closed (AB, CD) and open (EF, GH) dimers, showing the interacting regions at the interface in different colors. The open dimers lose the interaction between the catalytic domains present in the closed dimers, which is mostly mediated by blade IV (slate). (c) Sequence alignment of ScINV onto SoINV at the beginning of the β -sandwich domain; polar links at the ScINV octamer interface are marked with an asterisk and are mostly not conserved. (d) A model of the secreted octameric ScINV based only on the closed dimers is consistent with reported electron micrograph images (Esmon et al. 1987) showing open octamers on one side

antiparallel β -sheets (residues 342–512), linked by a short loop. Compared to other GH32 family members, blade I is the most conserved, whereas blade IV and β -sandwich have the lowest sequence identity. Likewise, these monomers are associated in two different kinds of dimers, which are in turn assembled into an unusual octameric quaternary structure, best described as a tetramer of dimers (Fig. 19.7a). Each monomer (A, B, C, D, E, F, G, and H) presents 512 residues (58.5 kDa) to conform a 428 kDa octamer with $130 \times 130 \times 110$ Å dimensions and 222 point symmetry.

On the one hand, the AB and CD chains, which are located at opposite vertices within the octamer, are tightly associated with a closed arrangement generating dimers resembling those described in SoFfase (Fig. 19.7b), to which it presents 68% sequence homology and 48% identity. 14 of the 32 hydrogen bonds are made between their catalytic domains, mainly through loop L3 and β -strand D4, and also through extended C5 and D5 strands within blade V (residues 316–328) (Fig. 19.7b). The active site of one monomer is surrounded by loops from the β -sandwich domain of the other, with several interactions through TIV and TV to the β -sandwich domains stabilizing the dimer interface.

By contrast, EF and GH are also associated as dimers but present a different interface, losing the interaction between the catalytic domains and generating an open assembly (Fig. 19.7b). Still, the catalytic pocket is reinforced by a new salt bridge between loop TI and the β -sandwich domain from the adjacent subunit, which encloses a wider but well-defined cavity. Therefore, the interface area of open dimers is half that found in closed dimers, as is the number of polar interactions. Furthermore, the closed dimers are energetically more favorable and, consequently, are expected to represent the experimentally detected dimeric form.

To create the higher oligomer, both types of dimers assemble through extension of the two β -sheets involving $\beta 1$ and $\beta 2$ from each β -sandwich domain (Fig. 19.7a). In particular, the extension mediated by $\beta 2$ forms a regular, antiparallel intermolecular β -sheet, similar to that reported in essential biomolecular recognition, processed as DNA base pairing or protein aggregation interaction such as the β -amyloid fibril formation. This interaction is determined by a unique short amino acid sequence at the beginning of the β -sandwich domain (Fig. 19.7c), with none of the residues that provide the polar links for the β -sheet extension through $\beta 1$ and $\beta 2$ being conserved in SoFfase. Furthermore, SoFfase has a high content of Lys residues at the $\beta 1$ to $\beta 2$ region, yielding a positively charged surface that would prevent dimer association along this region due to electrostatic repulsion effects. Consequently, the β -sandwich domain is, again, the responsible for the existence of the higher oligomerization state, unique to this invertase. Moreover, and depending on external conditions, this weaker association may be lost, mediating the disassembly of the octamers and generating the lower oligomeric states. Furthermore, a putative model of secreted octameric ScINV may be made, based on closed dimers linked by the observed intermolecular β -sheets. This model resembles strikingly the electron micrographs described for this isoform, which showed rectangles open in one side (Esmon et al. 1987).

Remarkably, the sophisticated quaternary structure of ScINV determines its substrate specificity.

From the comparative analysis of SoFfase and ScINV, a good fitting between their catalytic and most parts of the β -sandwich domains is evident (Fig. 19.8a). Moreover, the hydrogen link pattern and the topology of the SoFfase dimer are very similar to that in the closed AB/CD ScINV dimers. However, significant differences are obvious at their $\beta 1$ - and $\beta 2$ -strands that build the intermolecular β -sheet within the octamer. These structural differences generate a rearrangement

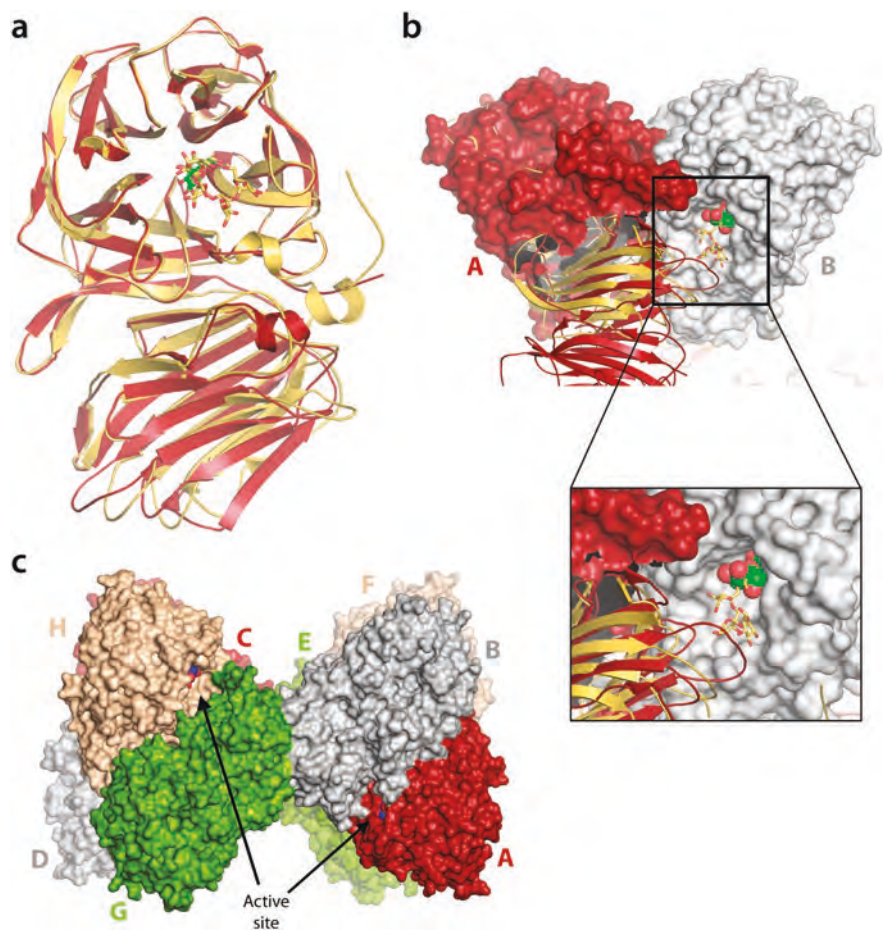


Fig. 19.8 Structural basis of the substrate specificity of ScINV. (a) Structural superimposition of SoFfase subunit (yellow-orange) in complex with fructosylmaltose (yellow sticks, PDB code 3U75) onto that from ScINV (red, PDB code 4EQV) with a modeled sucrose. (b) Surface representation of the ScINV closed dimer (red and gray) superimposed onto SoFfase (yellow), showing the respective β -sandwich domains of one subunit in cartoon representation. The adjacent β -sandwich domain making the extended β -sheet within the ScINV octamer is also shown. Inset: a detail of the closed dimer interface showing that the different arrangement of the β -sandwich domain makes a more constrained active site in ScINV than in SoFfase. (c) Spatial arrangement of the active sites within the octamer

of the loops linking the different β -strands within the non-catalytic domain that are projected towards the adjacent subunit and, consequently set steric constraints limiting the access to the active site to oligosaccharides larger than three or four sugar units (Fig. 19.8b). Moreover, in ScINV, the short-chain residues surrounding the cavity outline a more rigid and therefore less flexible active site to accommodate long and extended oligosaccharides. Therefore, the higher oligomerization pattern, mediated by the non-catalytic domains, plays a determinant role in defining the striking invertase specificity of ScINV. In turn, this large assembly serves to retain the enzyme at the periplasmic space, making a very efficient enzyme in its natural environment.

***Aspergillus kawachii* β -Fructofuranosidase, a Head-to-Head Dimer**

Aspergillus kawachii is a filamentous fungus used to brew alcoholic beverages and whose genome encodes for a β -fructofuranosidase (AkFfase). Apart from sucrose hydrolysis, it was characterized as a high transfructosylating protein, producing several types of FOS, especially 1-kestose. The three-dimensional structure of the recombinant version expressed in *Escherichia coli* was recently solved at 2.1 Å by crystallography, showing the expected GH32 fold (Nagaya et al. 2017). It is composed of four regions, an N-terminal small element lacking defined secondary structure (residues 25–59), a β -propeller domain (residues 60–440), an α -helical linker (residues 441–452), and a C-terminal β -sandwich domain (residues 453–628). Compared to the model bacterial invertase from *Thermotoga maritima*, AkFfase is characterized by the presence of an extra unstructured component at the N-terminal region, and a markedly larger blade V with two extra regions included between 374–388 and 407–428 residues (Fig. 19.9a). AkFfase shares higher structural homology to fungal enzymes such as *Aspergillus japonicus* β -fructosyltransferase (65% sequence identity) but, interestingly, it also presents high topological similarity to yeast XdINV (Fig. 19.9b), despite sharing a low level of 33% sequence identity.

Based on its sequence, the calculated molecular weight of AkFfase is 70 kDa, but a dimer of 157 kDa was determined in solution by the gel filtration technique. The crystal structure showed a dimerization pattern completely different from the oligomeric GH32 yeast enzymes described before, with the AkFfase interface formed only by interaction between both catalytic domains (Fig. 19.9c), in particular through the loop L2, and also L1, TII, and TIII. Thus, despite the topological similarity to XdINV, lack of its C-terminal extension and the essential Asn107 make a different dimeric assembly in AkFfase, in which the β -sandwich domain is not involved. Consequently, this fungal enzyme does not share the central role of the non-catalytical β -sandwich domain in oligomerization and substrate binding described in the yeast enzymes.

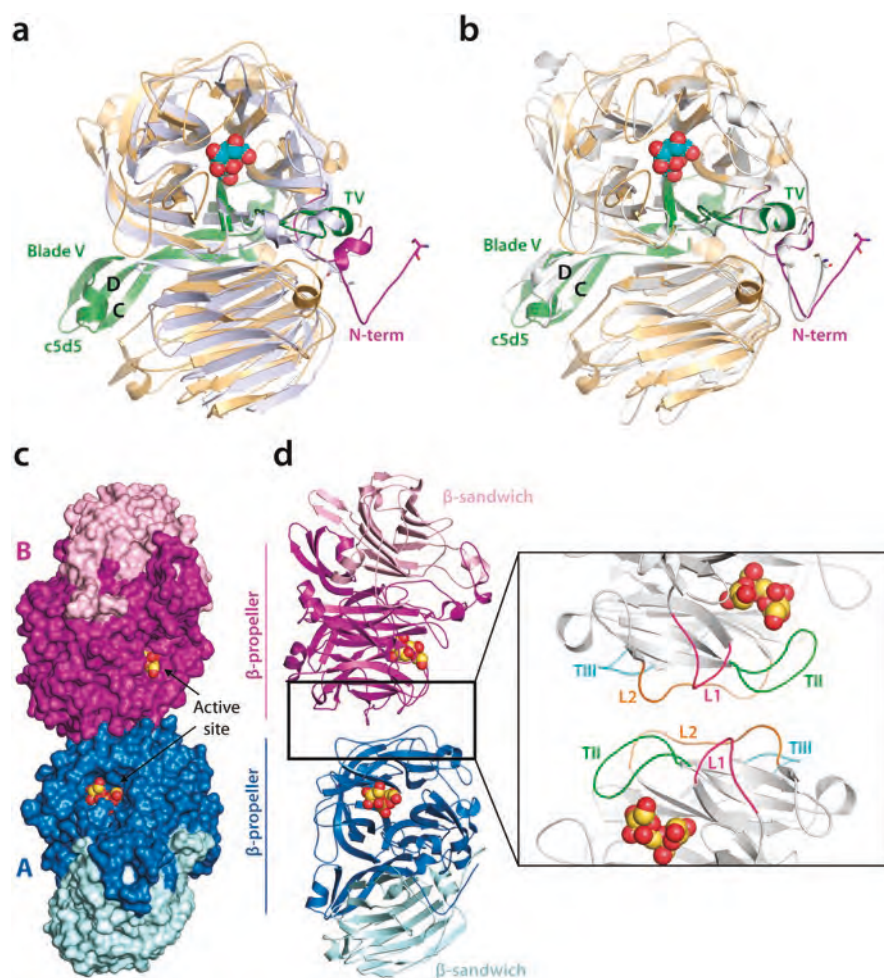


Fig. 19.9 *Aspergillus kawachii* β -fructofuranosidase (AkFfase). **(a)** Superposition of TmINV (light blue, PDB code 1UYP) onto AkFfase (orange, PDB code 5XHA). The N-terminal extension and blade V of AkFfase are in pink and green, respectively, and two fructose molecules at the active site of AkFfase are represented in cyan spheres. **(b)** Superposition of XdINV (gray, PDB code 5ANN) onto AkFfase (orange), showing their similar topology. **(c)** AkFfase dimer (blue and pink) is represented as a surface, and **(d)** as a cartoon; the catalytic domains are highlighted in brighter colors. Two fructose molecules are represented in yellow spheres. A zoom of the dimer interface showing the involved loops is included in the boxed area

***Frishella perrara* β -Fructofuranosidase, a Tail-to-Tail Dimer Containing an Extra Structural Element with Unknown Function**

Frishella perrara is a honeybee gut bacterium that produces a β -fructofuranosidase (FperFfase) able to hydrolyze sucrose and also short FOS, with the peculiarity that is more active on tetrasaccharides than on trisaccharides. Recently, its crystallographic structure has been reported at 1.6 Å (Kubota et al. 2022) showing the expected GH32 folding but, also, an additional N-terminal component (residues 1–26), which forms a long α -helix folded over the β -propeller on blades I and II (Fig. 19.10). This extra structural element was first reported in the β -fructofuranosidase from *Bifidobacterium longum*, BIFFase, which shares 41% sequence identity with FperFfase (Bujacz et al. 2011). Moreover, all β -fructofuranosidases from *Bifidobacteria*, deposited in GenBank, possess this additional N-terminal element, an observation that led to the assumption that is a general mechanism developed in probiotic bacteria adapted to live in a human digestive system, although the precise function of that element is so far unknown. More recently, a similar N-terminal long α -helix has been described in the *Bombyx mori* invertase (BmSUC1, (Miyazaki et al. 2020)), a secreted enzyme expressed in the midgut and silk glands of the domestic silkworm. Considering that the genomes of the animal kingdom do not generally encode GH32 invertases, the existence of sucrose hydrolyzing enzymes in *Bombyx mori* and some coleoptera and lepidoptera insects was attributed to result of lateral gene transfer (Miyazaki et al. 2020).

Furthermore, the elution profile of gel filtration showed that FperFfase formed a dimer in solution, a feature that was observed in the crystal structure (Fig. 19.10c, d). The interface is made essentially through the β -sandwich domains, including 14 hydrogen bonds between the 2 subunits. However, it does not imply an extension of the β -sheets, as observed in the ScINV octamer, and both β -sandwich domains approximate by their convex face. Also, and contrary to yeast XdINV, the great distance and the opposite localization between active sites from both subunits seem to avoid their cooperativity in substrate processing. Interestingly, a similar dimer was first described in the crystal structure of the structurally homologous BIFFase (Bujacz et al. 2011), although it was later discarded since the interface was considered too weak to form a stable dimer, and a monomer had been found in solution. Consequently, the biological relevance of the dimeric assembly of FperFfase remains unknown, the β -sandwich domain mediating the formation of the dimer but lacking an identified role in substrate binding.

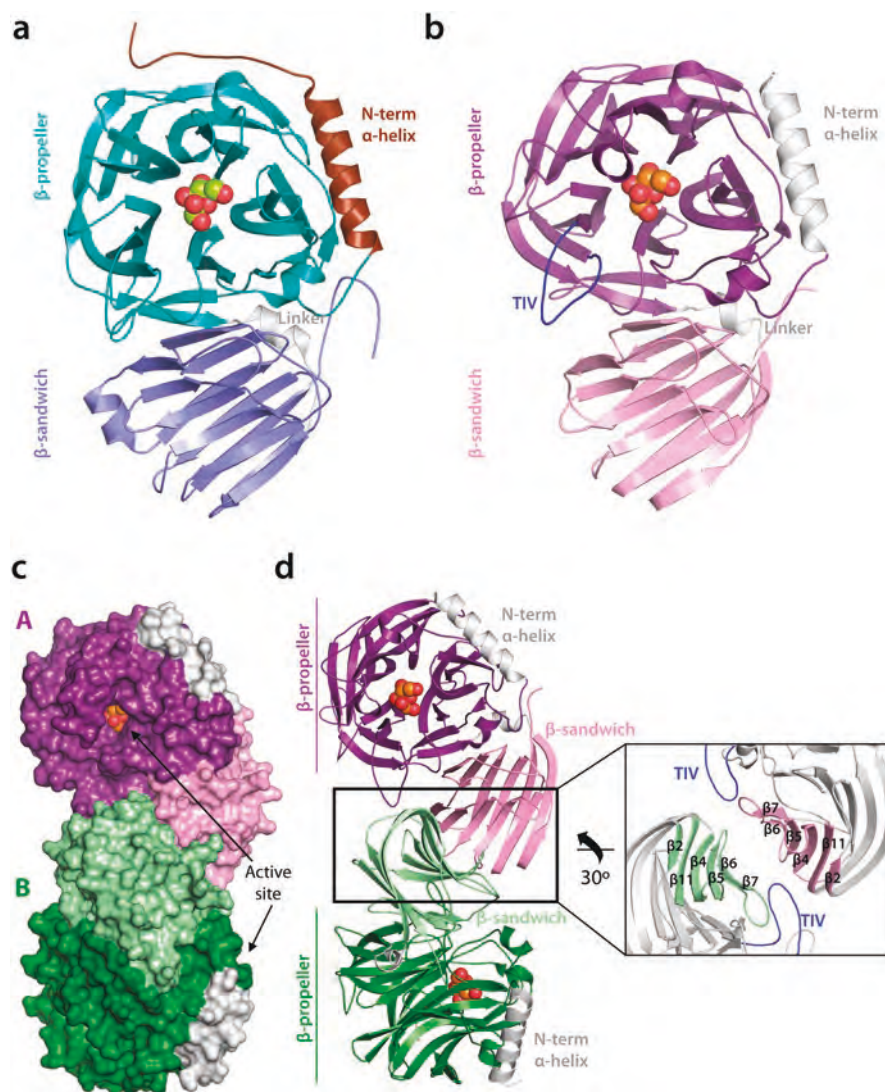


Fig. 19.10 *Frischella perrara* β -fructosyltransferase (FperFfase). (a) *Bifidobacterium longum* invertase (PDB code, 3PIG) showing the N-terminal additional structural element in brown, the catalytic domain in green, the linker in gray, and the β -sandwich domain in slate cartoon. A fructose molecule in the active site is represented in blue spheres. (b) Cartoon of the FperFfase subunit (PDB code 7VCP), showing the N-terminal α -helix and the linker in gray, the catalytic domain in violet, and the β -sandwich domain in orange. A fructose molecule in the active site is represented in pale green spheres. (c) and (d) The FperFfase dimer is made by interaction through the two β -sandwich domains. The inset shows a close view of the interface, which is mostly formed by the external β -sheet from each subunit, with additional participation of loop TIV from the corresponding catalytic domains

The Case of Cyanobacteria and Plant Invertases: A New Family of Glycosidases with Absolute Specificity for the α -(1,2) Bond of Sucrose

Recently, a novel arrangement has been described that is shared by cyanobacteria and plant invertases within the alkaline/neutral class composing the GH100 family. The first GH100 three-dimensional structure reported was that of InvA from the cyanobacteria *Anabaena* sp., which was solved at 2.11 Å resolution (Xie et al. 2016). The monomer is described as an $(\alpha/\alpha)_6$ barrel structure consisting of 12 α -helices, which are arranged in 2 concentric layers (Fig. 19.11a). Very short loops connect these helices at one face of the barrel, whereas long and highly variable segments are defining the other face, including two extended β -hairpins. This fold keeps high structural homology to $(\alpha/\alpha)_6$ -barrel glycoside hydrolases and also to several phosphorylases, the major difference being a 42-residue insert including 3 α -helices beyond the $(\alpha/\alpha)_6$ core.

InvA exists as a hexamer in solution and, in agreement, a hexameric assembly was found in the crystal structure. Its buried surface at the interface is greater within each dimer (about 6400 Å²) than between two dimers (about 2300 Å²), and, thus, the assembly can be described as a trimer of dimers. The dimers interact mainly through the insert mentioned above, which also enables the formation of the InvA hexamer (Fig. 19.11b, c).

The structure of the *Arabidopsis thaliana* invertase, AtNIN2, sharing 56% sequence identity with the *Anabaena* InvA, has been published at 3.4 Å resolution (Tarkowski et al. 2020). This plant neutral invertase resembles closely the hexameric assembly found in InvA, with an r.m.s.d. of only 1.1 Å between 2508 equivalent C α atoms, which points to high evolutionary conservation of neutral invertases (Tarkowski et al. 2020). Nevertheless, *Anabaena* sp. encodes a second alkaline invertase, InvB, which presents a comparable 60% sequence identity with InvA and was found to have a higher catalytic activity (Xie et al. 2018). The crystal structure of InvB, solved at 1.96 Å, showed a marked similarity to the InvA dimeric assembly (Fig. 19.11d), being a dimer with a buried interface area of about 5500 Å². However, a difference of 10° in the inter monomer angles hinders further interaction of two dimers to make a stable hexamer. Interestingly, the only significant structural difference between the InvA and InvB subunits is a slight rearrangement of the insert, this region having lower sequence homology among different alkaline/neutral invertases (Fig. 19.11d).

Other higher oligomeric forms have been previously identified in some alkaline/neutral invertases, but the dimeric association may be considered to be the basic structural template of these enzymes (Xie et al. 2018). As shown in Fig. 19.11e, f, the dimer embraces a rather closed substrate binding pocket, whose access is delineated by both subunits near the interface, which is perfectly complementary to the small sucrose molecule. This feature is different from the generally open catalytic pocket of acidic invertases that, consequently, can accommodate β -fructofuranosides other than sucrose. Moreover, analysis of the *Anabaena* sp. InvA and InvB

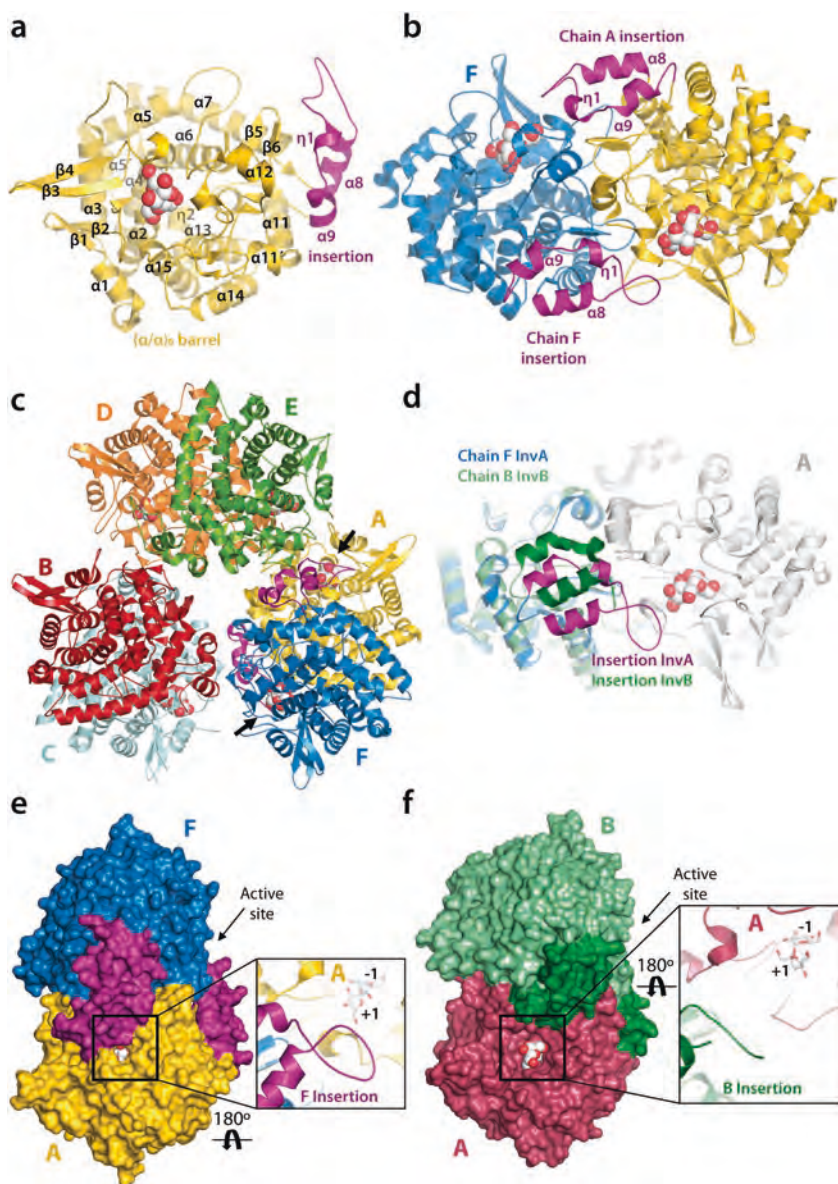


Fig. 19.11 *Anabaena* sp. Alkaline/neutral invertases. (a) Invertase A (InvA) subunit folding, with an $(\alpha/\alpha)_6$ barrel core (yellow) and a 42-residue insert of 3 helices highlighted in violet. A sucrose bound at the active site is shown as spheres (PDB code 5GOP). (b) the InvA tight dimer and (c) the assembly of three dimers into an hexamer. The insert is involved in the oligomerization interface within and between the dimers. (d) Comparison of invertase B (InvB, PDB code 5Z74) and InvA dimers, by superimposition of subunit A from both enzymes. Major difference is a rearrangement of the insert in InvB, which precludes further interaction of the dimers to associate in higher oligomers. (e) Molecular surface of the InvA dimer showing the position of the active site occupied by a sucrose molecule (spheres), close to the interface. The inset shows a close view of the active site, with the sucrose represented in sticks. (f) Molecular surface of the InvB dimer

complexes with sucrose substrate and products showed that the subsites -1 and $+1$ are strictly selective towards glucose and fructose, respectively, and that the hydrolysis occurs at the α -(1,2)-glycosidic bond. Its exclusive ability to accept sucrose as its only substrate makes alkaline/neutral invertases a peculiar and distinctive family among other GHs breaking α -(1,2)-glycosidic bonds.

Conclusions and Outlook

Invertase, or β -fructofuranosidase, is a key metabolic enzyme existing in various forms of life that hydrolyzes β -fructofuranoside residues. Since 1860, in which Berthelot first isolated the invertase from the yeast *Saccharomyces cerevisiae*, the enzyme became a model in biochemical studies and has been one of the most used in industry. Several oligomeric forms were described for both the heavily glycosylated secreted and the non-glycosylated intracellular forms, an issue that was later visualized by electron microscopy by the 1980s. However, it was not until 2004 that structural knowledge was acquired through the reported crystallographic structure of the bacterial invertase from *Thermotoga maritima*, which provided a template for invertases, inulinases, levanases, and fructosyltransferases, all included within the GH32 family. As such, they are bimodular enzymes with a β -propeller catalytic domain that defines the active site. A second non-catalytic β -sandwich domain distinguished this family from related enzymes within family GH68, but no function was attributed to it. Since then, the crystal structures of 22 new enzymes have been reported, 14 of them from eukaryotes, the enzymes from plant, fungi, and bacteria regulating their specificity through a high sequence variability in the loops surrounding the active site at the center of the barrel. However, the enzyme reported from the yeast *Schwanniomyces occidentalis* showed that the active site must be considered in terms of the oligomerization state, and the existence of tight dimers formed by contribution of the β -sandwich domain that also participates directly in substrate binding. This pattern was shared by other enzymes described from the yeasts *Kluyveromyces lactis* (exoinulinase INU1, PDB code 6J0T) and *Rhodotorula dairenensis* (RdINV). Furthermore, the structure of the recombinant invertase from *Saccharomyces cerevisiae* revealed how these tight dimers assemble into higher oligomers by extension of the β -sheets within their non-catalytic domain, illustrating the topology previously described in the intra- and extracellular octameric forms, and the reversible assemble/disassemble into different oligomeric forms upon external conditions. A second class of elongated dimer was described in the invertase from *Xanthophyllomyces dendrorhous* formed through the β -sandwich non-catalytic domain but, also, absolutely dependent on glycans N-attached to a unique Asn site. Thus, all reported enzymes from yeast are oligomers sharing an active site contoured by two subunits, in which the non-catalytic domain is directly involved in substrate binding of the adjacent subunit therefore being essential in fine-tuning specificity. Furthermore, the dimers build a particular arrangement of the active sites within each enzyme that in some cases suggest the putative binding mode of polymeric chains, proposing their cognate substrate in their natural environment. A few dimeric forms have been reported in other fungal and bacterial enzymes,

both through the catalytic (head-to-head) or the β -sandwich domain (tail-to tail), but these assemblies locate the active site far from the interface and, therefore, are not expected to affect the enzymatic activity or specificity. Consequently, the sophisticated oligomeric mechanism depending on the non-catalytic domain and regulating specificity seems unique to the yeast members. Further work on novel enzymes will be necessary to uncover the role of modularity within this important family of enzymes.

Interestingly, the GH100 families recently depicted in cyanobacteria and plant invertases, showing absolute specificity for the α -(1,2)-bond of sucrose, are tight dimers able to associate into higher oligomers. These dimers present their active sites near the interface, although there is no evidence of direct participation of the associated subunit in binding. Nevertheless, the access is delineated by both subunits that make a rather restricted active site corresponding to the small sucrose molecule. This situation is different from the generally open catalytic pocket of acidic invertases that, consequently, can cleave β -fructofuranosides other than sucrose. Thus, dimerization might contribute to support the exclusive invertase activity of this family of enzymes.

Carbohydrate-processing enzymes have developed intricate molecular mechanisms to cope with their highly heterogeneous substrates. Understanding these mechanisms will be of paramount interest to uncover essential biological processes at the molecular level. Also, it will ensure the effective design of novel and superior enzymes for biotechnological purposes.

References

- Alberto F, Bignon C, Sulzenbacher G et al (2004) The three-dimensional structure of Invertase (β -Fructosidase) from *Thermotoga maritima* reveals a Bimodular arrangement and an evolutionary relationship between retaining and inverting Glycosidases. *J Biol Chem* 279:18903–18910. <https://doi.org/10.1074/jbc.M313911200>
- Alegre ACP, Teixeira d L, de Moraes Polizeli M, Terenzi HF et al (2009) Production of thermostable invertases by *Aspergillus caespitosus* under submerged or solid state fermentation using agroindustrial residues as carbon source. *Braz J Microbiol* 40:612–622. <https://doi.org/10.1590/S1517-838220090003000025>
- Álvaro-Benito M, Polo A, González B et al (2010) Structural and kinetic analysis of *Schwanniomyces occidentalis* invertase reveals a new oligomerization pattern and the role of its supplementary domain in substrate binding. *J Biol Chem* 285:13930–13941. <https://doi.org/10.1074/jbc.M109.095430>
- Álvaro-Benito M, Sainz-Polo MA, González-Pérez D et al (2012) Structural and kinetic insights reveal that the amino acid pair Gln-228/Asn-254 modulates the transfructosylating specificity of *Schwanniomyces occidentalis* β -fructofuranosidase, an enzyme that produces prebiotics. *J Biol Chem* 287:19674–19686. <https://doi.org/10.1074/jbc.M112.355503>
- Bujacz A, Jedrzejczak-Krzepkowska M, Bielecki S et al (2011) Crystal structures of the apo form of β -fructofuranosidase from *Bifidobacterium longum* and its complex with fructose. *FEBS J* 278:1728–1744. <https://doi.org/10.1111/j.1742-4658.2011.08098.x>

- Cimini S, Di Paola L, Giuliani A et al (2016) GH32 family activity: a topological approach through protein contact networks. *Plant Mol Biol* 92:401–410. <https://doi.org/10.1007/s11103-016-0515-2>
- Davies G, Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* 3:853–859. [https://doi.org/10.1016/S0969-2126\(01\)00220-9](https://doi.org/10.1016/S0969-2126(01)00220-9)
- Du L, Pang H, Wang Z et al (2013) Characterization of an Invertase with pH tolerance and truncation of its N-terminal to shift optimum activity toward neutral pH. *PLoS One* 8:e62306. <https://doi.org/10.1371/journal.pone.0062306>
- Esmon PC, Esmon BE, Schauer IE et al (1987) Structure, assembly, and secretion of octameric invertase. *J Biol Chem* 262:4387–4394. [https://doi.org/10.1016/s0021-9258\(18\)61360-2](https://doi.org/10.1016/s0021-9258(18)61360-2)
- Garron ML, Henrissat B (2019) The continuing expansion of CAZymes and their families. *Curr Opin Chem Biol* 53:82–87. <https://doi.org/10.1016/j.cbpa.2019.08.004>
- Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 280:309–316. <https://doi.org/10.1042/bj2800309>
- Hidaka H, Eida T, Takizawa T (1986) Effects of Fructooligosaccharides on intestinal Flora and Human health. *Bifidobacteria Microflora* 5:37–50
- Jiménez-Ortega E, Narmontaite E, González-Pérez B et al (2022) Insights into the structure of the highly glycosylated Ffase from *Rhodotorula dairenensis* enhance its biotechnological potential. *Int J Mol Sci* 23:14981. <https://doi.org/10.3390/ijms232314981>
- Kern G, Schulke N, Schmid FX, Jaenicke R (1992) Internal, external, and core-glycosylated invertase from yeast. *Protein Sci* 1:120–131
- Kubota A, Kawai R, Li D et al (2022) Enzymatic and structural characterization of β -fructofuranosidase from the honeybee gut bacterium *Frischella perrara*. *Appl Microbiol Biotechnol* 106:2455–2470. <https://doi.org/10.1007/s00253-022-11863-9>
- Kulshrestha S, Tyagi P, Sindhi V, Yadavilli KS (2013) Invertase and its applications—a brief review. *J Pharm Res* 7:792–797. <https://doi.org/10.1016/j.jopr.2013.07.014>
- Lammens W, Le Roy K, Schroeven L et al (2009) Structural insights into glycoside hydrolase family 32 and 68 enzymes: functional implications. *J Exp Bot* 60:727–740. <https://doi.org/10.1093/jxb/ern333>
- Lincoln L, More SS (2017) Bacterial invertases: occurrence, production, biochemical characterization, and significance of transfructosylation. *J Basic Microbiol* 57:803–813. <https://doi.org/10.1002/jobm.201700269>
- Lombard V, Golaconda Ramulu H, Drula E et al (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:490–495. <https://doi.org/10.1093/nar/gkt1178>
- Miyazaki T, Oba N, Park EY (2020) Structural insight into the substrate specificity of *Bombyx mori* β -fructofuranosidase belonging to the glycoside hydrolase family 32. *Insect Biochem Mol Biol* 127:103494. <https://doi.org/10.1016/j.ibmb.2020.103494>
- Nagaya M, Kimura M, Gozu Y et al (2017) Crystal structure of a β -fructofuranosidase with high transfructosylation activity from *aspergillus kawachii*. *Biosci Biotechnol Biochem* 81:1786–1795. <https://doi.org/10.1080/09168451.2017.1353405>
- Ramírez-Escudero M, Gimeno-Pérez M, Lez BG et al (2016) Structural analysis of β -fructofuranosidase from *Xanthophyllomyces dendrorhous* reveals unique features and the crucial role of N-glycosylation in oligomerization and activity. *J Biol Chem* 291:6843–6857. <https://doi.org/10.1074/jbc.M115.708495>
- Sainz-Polo MA, Ramírez-Escudero M, Lafraya A et al (2013) Three-dimensional structure of *Saccharomyces invertase*: role of a non-catalytic domain in oligomerization and substrate specificity. *J Biol Chem* 288:9755–9766. <https://doi.org/10.1074/jbc.M112.446435>
- Sturm A (1999) Update on biochemistry Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning some common molecular features but differ. *Plant Physiol* 121:1–7
- Tammi M, Ballou L, Taylor A, Ballou CE (1987) Effect of glycosylation on yeast invertase oligomer stability. *J Biol Chem* 262:4395–4401. [https://doi.org/10.1016/s0021-9258\(18\)61361-4](https://doi.org/10.1016/s0021-9258(18)61361-4)

- Tanno H, Fujii T, Hirano K et al (2021) Characterization of fructooligosaccharide metabolism and fructooligosaccharide-degrading enzymes in human commensal butyrate producers. *Gut Microbes* 13:1–20. <https://doi.org/10.1080/19490976.2020.1869503>
- Tarkowski LP, Tsirkone VG, Osipov EM et al (2020) Crystal structure of *Arabidopsis thaliana* neutral invertase 2. *Acta Crystallogr F Struct Biol Commun* 76:152–157. <https://doi.org/10.1107/S2053230X2000179X>
- Trollope KM, Görgens JF, Volschenk H (2015) Semirational directed evolution of loop regions in *Aspergillus japonicus* β -fructofuranosidase for improved fructooligosaccharide production. *Appl Environ Microbiol* 81:7319–7329. <https://doi.org/10.1128/AEM.02134-15>
- Xie J, Cai K, Hu HX et al (2016) Structural analysis of the catalytic mechanism and substrate specificity of *anabaena* alkaline invertase InvA reveals a novel glucosidase. *J Biol Chem* 291:25667–25677. <https://doi.org/10.1074/jbc.M116.759290>
- Xie J, Hu HX, Cai K et al (2018) Structural and enzymatic analyses of *anabaena* heterocyst-specific alkaline invertase InvB. *FEBS Lett* 592:1589–1601. <https://doi.org/10.1002/1873-3468.13041>
- Zambelli P, Fernandez-Arrojo L, Romano D et al (2014) Production of fructooligosaccharides by mycelium-bound transfructosylation activity present in *Cladosporium cladosporioides* and *Penicillium sizovae*. *Process Biochem* 49:2174–2180. <https://doi.org/10.1016/j.procbio.2014.09.021>
- Zhou J, He L, Gao Y et al (2016) Characterization of a novel low-temperature-active, alkaline and sucrose-tolerant invertase. *Sci Rep* 6:1–9. <https://doi.org/10.1038/srep32081>

Chapter 20

Structure, Dynamics and Functional Implications of the Eukaryotic Vault Complex



María González-Álamos, Pablo Guerra, and Núria Verdaguer

Abstract Vault ribonucleoprotein particles are naturally designed nanocages, widely found in the eukaryotic kingdom. Vaults consist of 78 copies of the major vault protein (MVP) that are organized in 2 symmetrical cup-shaped halves, of an approximate size of 70x40x40 nm, leaving a huge internal cavity which accommodates the vault poly(ADP-ribose) polymerase (vPARP), the telomerase-associated protein-1 (TEP1) and some small untranslated RNAs. Diverse hypotheses have been developed on possible functions of vaults, based on their unique capsular structure, their rapid movements and the distinct subcellular localization of the particles, implicating transport of cargo, but they are all pending confirmation. Vault particles also possess many attributes that can be exploited in nanobiotechnology, particularly in the creation of vehicles for the delivery of multiple molecular cargoes. Here we review what is known about the structure and dynamics of the vault complex and discuss a possible mechanism for the vault opening process. The recent findings in the characterization of the vaults in cells and in its natural micro-environment will be also discussed.

Keywords Major vault protein (MVP) · Nanocage · Ribonucleoprotein particles
Vault structure · Vault dynamics

M. González-Álamos · N. Verdaguer (✉)

Structural and Molecular Biology Department, Instituto de Biología Molecular de Barcelona (IBMB-CSIC), Barcelona, Spain

e-mail: mgacri@ibmb.csic.es; nvmcri@ibmb.csic.es

P. Guerra

Cryo-Electron Microscopy Platform – IBMB CSIC, Joint Electron Microscopy Center at ALBA (JEMCA), Barcelona, Spain

e-mail: pglcri@ibmb.csic.es

Historical Background

The vault complex, with a mass of 10 MDa and overall dimensions of 70 nm × 40 nm × 40 nm, is the largest ribonucleoprotein described in the eukaryotic kingdom (Kedersha et al. 1990; Hamill and Suprenant 1997; Herrmann et al. 1997). Despite their prevalence in both human and animal tissues, the existence of vaults went unnoticed until 1986 (Kedersha et al. 1986). The delay in their discovery can be attributed to the limitations of classical electron microscopy techniques, where stains primarily highlighted membranes and nucleic acids, rendering proteins less visible. Nevertheless, the introduction of negative staining electron microscopy made possible the initial identification of vaults as contaminants in clathrin-coated vesicle samples obtained from rat liver tissues. Due to their symmetric barrel-shaped morphology, reminiscent of gothic cathedral ceilings, these structures were dubbed vaults (Rome et al. 1991).

In mammals, vaults contain three proteins, the 100 kDa major vault protein (MVP) (Kedersha and Rome 1986), the 193 kDa vault poly(ADP-ribosyl)ating polymerase vPARP (Kickhoefer et al. 1999a) and the 240 kDa telomerase-associated protein TEP1 (Kickhoefer et al. 1999b), and several small untranslated RNAs of about 100 nt in length (Kedersha and Rome 1986; Kickhoefer et al. 1993; Van Zon et al. 2001). Vaults are uniform in size and morphology when imaged by electron microscopy, presenting a barrel-like structure with an invaginated waist and two protruding caps (Kong et al. 1999). The major vault protein (MVP) is the architect in charge of the construction of this particle since it encodes all the information necessary for the assembly of the vault shell (Kickhoefer et al. 1996; Mikyas et al. 2004). This notion was initially supported by the fact that vault-like particles (VLPs), similar to purified endogenous vaults, were observed when rat MVP was expressed in insect cells as well as in yeast, indicating that MVP was capable of generating the minimal vault structure and the entire outer shell of the vault particle (Stephen et al. 2001; Poderycki et al. 2006; Kickhoefer et al. 2009). The structural integrity of the vault-like particles (VLPs) was further confirmed by comparing the X-ray structures of intact rat vaults (Tanaka et al. 2009; Querol-Audí et al. 2009; Casañas et al. 2013), and the structures of recombinant rat-derived VLPs, determined by cryo-electron microscopy (cryo-EM) (Ding et al. 2018; Guerra et al. 2022). All these structures have shown that the vault shell is folded into 2 identical halves, each consisting of 39 copies of MVP, which assemble to form a barrel-like cage with a large internal cavity, where the minor components are located.

Vaults have been found in a diverse range of species, extending beyond mammals. They have been observed in evolutionarily distant organisms, including protozoa (for instance, *Dictyostelium discoideum* amoeba), molluscs and echinoderms (like the sea urchin *Strongylocentrotus purpuratus*). Vaults are also present in amphibians (*Rana catesbeiana*, *Xenopus laevis*), avian species (*Gallus gallus*) and even in fish (such as the electric ray *Torpedo marmorata*). Interestingly, vaults are notably absent in other species commonly used in molecular research such as *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster* as well as in plants (Berger et al. 2009). The extensive conservation of vaults implies

a fundamental function associated with their distinctive barrel-like structure. Although the precise function of vaults remains unclear, their distinct subcellular localization, the hollow structure and the rapid movement of the vault particles led to a variety of hypotheses relating vaults to intracellular transport, shuttling molecules inside the cell from one location to another (Berger et al. 2009; Frascotti et al. 2021).

On the other hand, the vault particles represent an ideal tool for the targeted delivery of drugs and other cargoes. In fact, since Kickhoefer and colleagues demonstrated that the vPARP-derived peptide INT could be used as a carrier to target heterologous proteins to the vault interior (Kickhoefer et al. 2005), recombinant vaults have undergone significant engineering, including cell-surface receptor targeting (Kickhoefer et al. 2009) and the encapsulation of a wide variety of molecules (Casañas et al. 2012; Rome and Kickhoefer 2013; Wang et al. 2015; Benner et al. 2017; Wang et al. 2018; Fulcher et al. 2019; Nagasawa et al. 2020; Voth et al. 2020). The flexibility and the adaptability of this recombinant system have been greatly enhanced by the increasing amounts of atomic level information and improved comprehension of vault structure and dynamics.

The Major Vault Protein

Since the discovery of the vault complex in the middle of the 1980s, numerous studies have been performed, aimed to understand the structure of this enigmatic particle. The first structure of vaults was obtained by cryo-EM, single-particle analysis (SPA) and 3D reconstruction (Kong et al. 1999). Despite its low resolution, below 30 Å, the barrel-shaped vault morphology with two protruding caps and an invaginated waist was clearly visible (Kong et al. 1999).

Three independent groups were able to crystallize the complete vault particle. The initial crystals were obtained in our lab in 2005, where vaults appeared as contaminants in a human astrovirus preparation (Querol-Audi et al. 2005). Two years later, Anderson and colleagues managed to obtain crystals of recombinant vaults that diffracted near 9 Å resolution (Anderson et al. 2007).

The first high-resolution (3.5 Å) X-ray structure of intact rat vaults was determined by Tanaka and colleagues in 2009, showing that the vault shell structure was organized as a dimer of half-vaults, with each half-vault comprising 39 quasi-identical MVP copies (Tanaka et al. 2009). Each MVP monomer was organized in 12 domains: 9 structural repeat domains at the N-terminus, followed by a shoulder domain, a cap-helix domain and a cap-ring domain at the C-terminus (Fig. 20.1a). The nine MVP repeats share a similar fold, each consisting of five antiparallel β -strands connected by loops of different size and conformation. The shoulder domain folds into a single α/β -domain with a four-stranded antiparallel β -sheet on one side and four α -helices on the other. The large cap-helix domain is formed by a 42-turn-long α -helix that is followed by a U-shaped structure that constitutes the cap-ring domain at the MVP C-terminus (Tanaka et al. 2009). The strongest

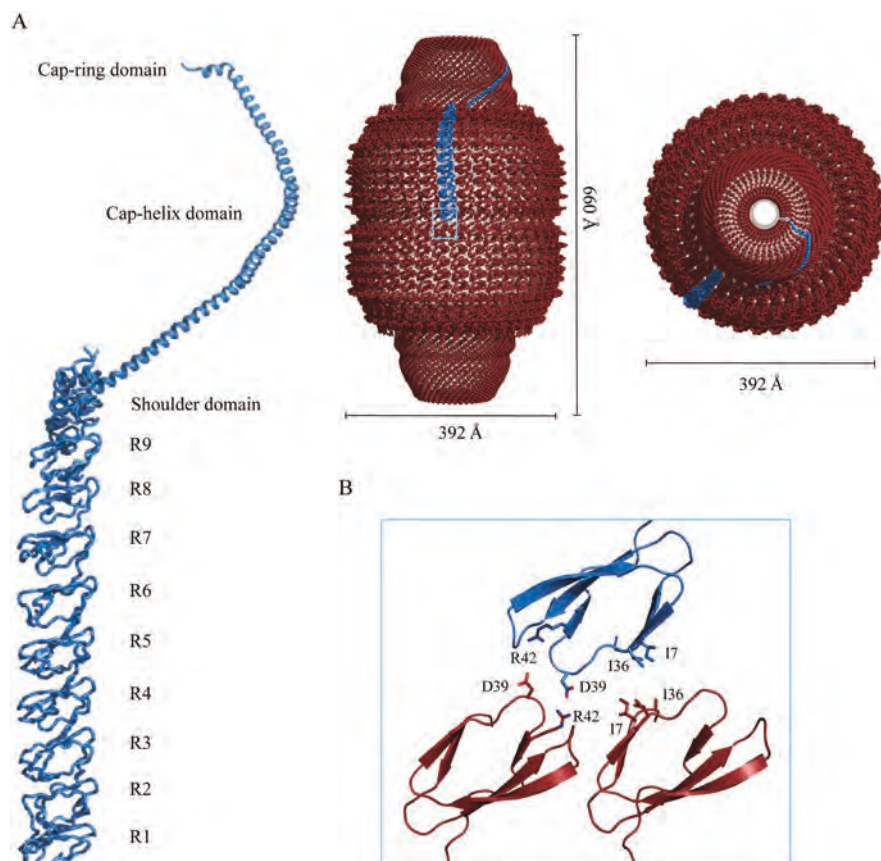


Fig. 20.1 Overall structure of the vault complex and MVP domain organization. **(a)** Side and top views of the vault shell, colored in maroon with one of the 78 MVP copies forming the particle shown in blue (PDB ID: 4HL8; Casañas et al. 2013). The left panel shows the structure of one MVP subunit with its constituent domains explicitly labeled. The MVP is composed of twelve domains: nine repeat domains at the N-terminus (R1 to R9), and α/β shoulder domain, a cap-helix and a cap-ring domain at the C-terminus. **(b)** Detailed view of the R1-R1 contacts at the half vault interface. The reference R1 domain (blues) contacts two consecutive R1 molecules (maroon) through the molecular two-fold axis. Interacting residues are shown in sticks and labelled

MVP-MVP contacts are found between cap-helix domains, where the hydrophobic residues are located on the interface between two adjacent helices and polar residues are exposed to inner or outer surfaces of the particle. Lateral inter-subunit interactions in the barrel region are essentially stabilized by polar bonds (Fig. 20.2).

Almost in parallel, the X-ray structure of the seven N-terminal repeat domains (R1–R7) that conform the vault waist was determined in our lab at 2.1 Å resolution. The structure of the intact vaults from rat liver was also determined in this work (8 Å resolution), revealing that a combination of electrostatic and hydrophobic interactions governed the association of the two half-vault moieties (Querol-Audi

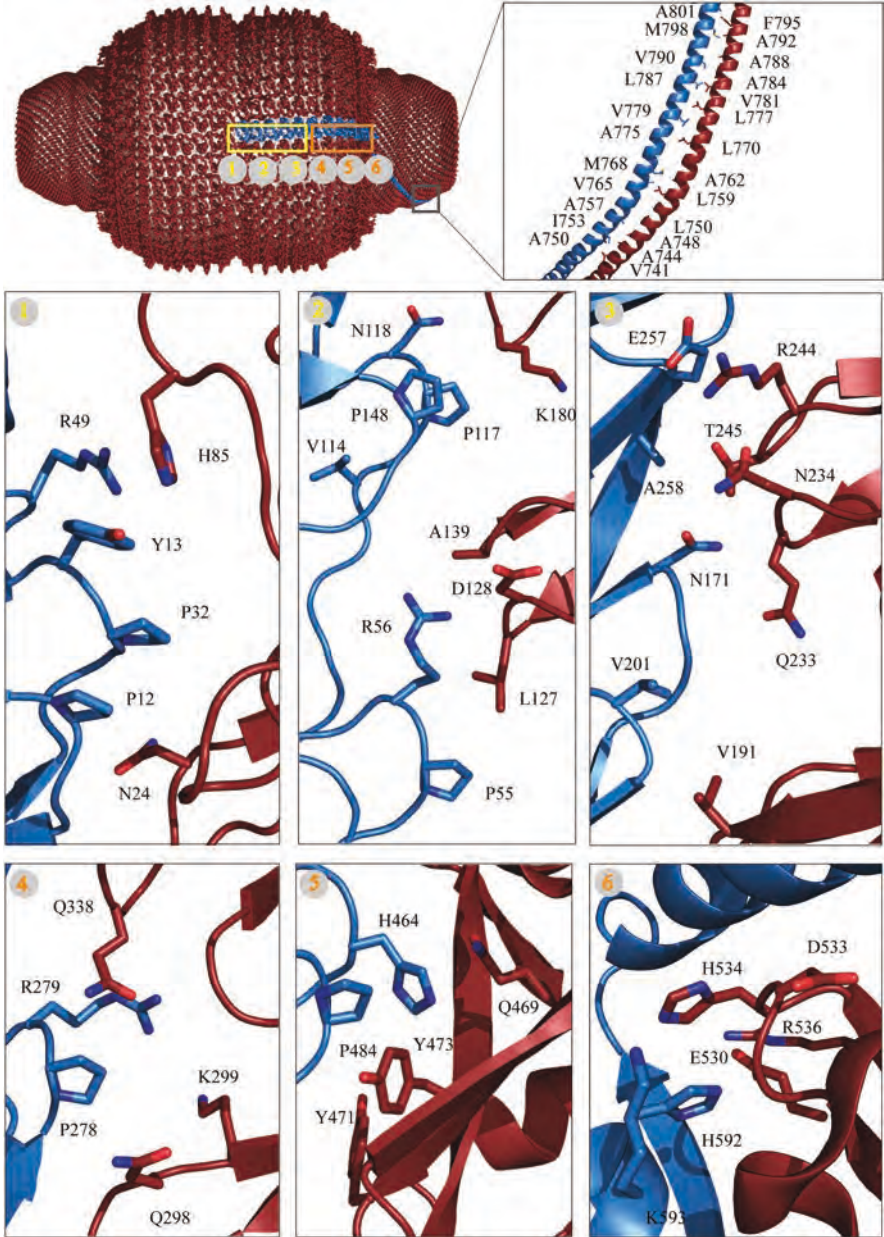


Fig. 20.2 MVP-MVP lateral interactions stabilizing the vault particle. The top left panel shows a cartoon representation of an entire vault (maroon cartoon), with a reference MVP monomer shown in blue. The different panels show representative close up views of the MVP-MVP interactions in the cap-helix (top right panel), in the vault barrel domains (panels 1 to 5) and in the α/β shoulder (panel 6)

et al. 2009) (Fig. 20.1b). The 3.5 Å structure of the entire vault particle was then refined by incorporating the high-resolution information available for the R1–R7, and applying the strict D39 non-crystallographic symmetry (Casañas et al. 2013). The precise examination of the refined vault structure predicted the breakage of the 39-fold symmetry in the cap region that could be formed by only 36 strands defining the narrowest radius (~25 Å), with respect to the central vault axis. The analysis also indicated that MVP could experience significant movements at its N-terminus during the sealing of the two vault-halves. This fact opened the question of what is the behaviour of vaults in solution, in a non-crystalline environment, and whether it undergoes conformational changes (Casañas et al. 2013).

Towards Understanding the Structural Dynamics of the Vault Shell

Very recently, the cryo-EM structures of recombinant vaults have been determined, reaching resolutions similar to those obtained by X-ray diffraction 3.7 Å (Guerra et al. 2022) and ~4.8 Å (Ding et al. 2018). Ding and colleagues were the first to obtain structures of recombinant rat vaults at near-atomic resolution, using single-particle analysis (SPA) techniques and imposing the D39 symmetry. These authors described two different conformations of the vault shell: conformation 1 was almost identical to the crystal structure (PDB: 4HL8) (Casañas et al. 2013), while conformation 2 appeared longitudinally translocated up to 10 Å, resulting in an outward-projected cap and consequently in a vault 14 Å longer than conformation 1, along the D39-symmetry axis (Ding et al. 2018) (Fig. 20.3). The cryo-EM work performed in our lab (Guerra et al. 2022) provided new evidences of multiple vault populations existing in solution, identifying and determining the structures of three different vault states: (i) a fully closed vaults, determined at 3.7 Å resolution by imposing D39 symmetry, closely related to the crystal structure (Figs. 20.3 and 20.4a); (ii) an additional whole vault state (7.9 Å resolution), showing a disruption of the rotational C39 symmetry at the vault midsection (Fig. 20.4b); and (iii) a half-vault state (9.3 Å resolution), showing a distortion of the circular shape of the vault waist (Fig. 20.4c).

Structural comparisons between fully closed vaults (PDB, 7PKR; Guerra et al. 2022) and the crystal structure of native rat vaults (PDB, 4HL; Casañas et al. 2013) showed some differences in the shape and size of the particle but maintaining the overall vault architecture (Figs. 20.3 and 20.4a). The observed differences involved an expansion in the diameter of the barrel region that in the cryo-EM reconstruction was 8 Å bigger than in the crystallized particles (Fig. 20.3). Detailed structural comparisons of the MVP monomers indicated that the overall conformation was maintained, but the orientation of domains R1–R5 relative to the C-terminal helix was changed, increasing the volume of the internal cavity, resulting in a wider, expanded cavity, with a relaxed conformation of the barrel domain (Figs. 20.3 and 20.4a).

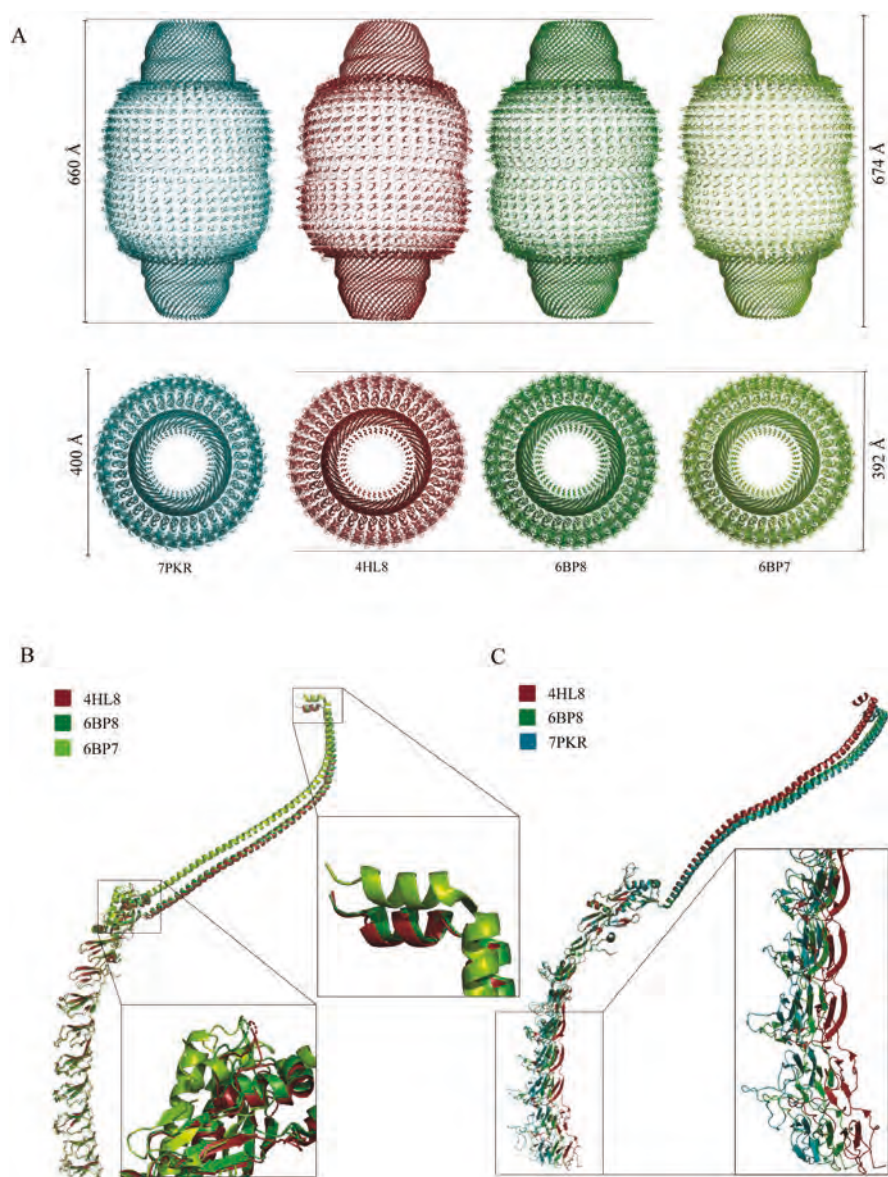


Fig. 20.3 Comparison of particle dimension between the different structures of vaults currently available. **(a)** Cartoon representations of the different vault models: the cryo-EM structure of recombinant rat vaults closed symmetric conformation (blue; PDB ID: 7PKR; Guerra et al. 2022); the X-ray structure of native rat vaults (maroon; PDB ID: 4HL8; Casañas et al. 2013); cryo-EM structures of recombinant rat vaults, conformations 1 (dark green) and 2 (light green) (PDB IDs: 6BP8, 6BP7; Ding et al. 2018). Comparisons show that the longest particle corresponds to the structure 6BP7 and the widest to 7PKR. **(b&c)** Structural superimposition of the X-ray (PDB ID: 4HL8) and cryo-EM structures (PDB IDs: 6BP8, 7PKR), showing that despite the overall conformation of MVP monomers is maintained in all three structures, an important change on the degree of curvature of the vault barrel has been observed, affecting the volume of the internal cavities. It has also been observed a distinctive orientation of the cap helices

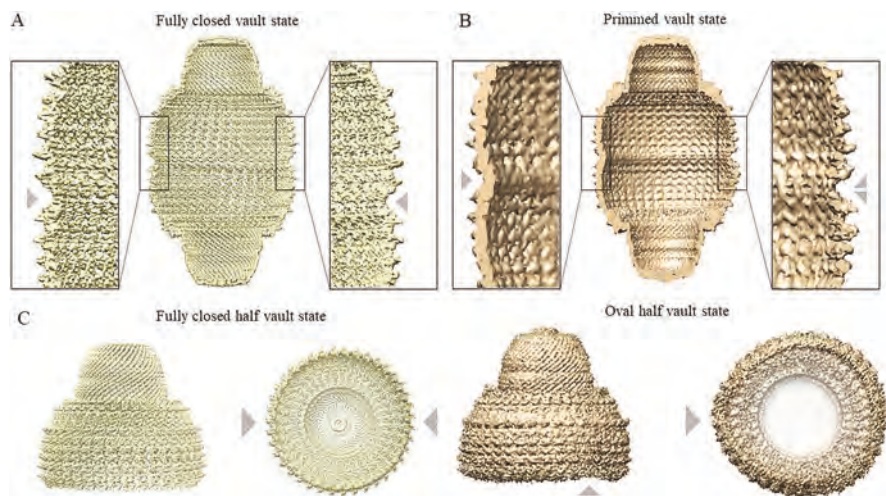


Fig. 20.4 Cryo-EM maps of three vault conformations. (a) The 3D map (3.7 Å resolution) of the symmetric closed vault conformation with (D39 symmetry; map EMD-13478 corresponding to PDB ID: 7PKR; Guerra et al. 2022). The whole particle is represented with a zoom at the midsections of the half-vault interface, shown at the left side. (b) Asymmetrical conformation of closed vaults (map EMD-13483; Guerra et al. 2022). The whole particle is shown in the middle with a zoom at the midsections of the vault interface, showing the R1-R1 domains engaged on the left side and disengaged on the right side. (c) 3D map of free half-vaults (map EMD-13482; Guerra et al. 2022) exhibiting a noncircular, oval contour (right panels). For comparison, similar views from a map derived from the X-ray structure of the whole vault are shown in left panels

The second class of full particles structurally determined (Guerra et al. 2022) showed a local deviation from the symmetry of the ‘closed’ vault state. This deviation was evident in only one region of the vault waist, with R1 domains forming the top and bottom half-vaults engaged on one side and disengaged at the other side (Fig. 20.4b). This structure was hypothesized to represent an initial state of the opening process of the particle (Guerra et al. 2022).

The third state of vaults analysed corresponded to half-vaults (Guerra et al. 2022). These particles showed a distortion of the circular shape of the vault waist, revealing an oval contour configuration rather than the circular one, present in the closed whole vault state (Fig. 20.4c). The reconstructed map revealed the structural changes responsible for the oval waist contour, affecting not only the waist but also the central region of the vault barrel. The main structural variation found was a reorganization of the MVP monomers in 3 clusters of 13 monomers each, while the half-vault overall configuration was maintained. The individual MVP N-terminal domains from R3 to R9 of the barrel, the shoulder and the cap helix and cap ring at the top of the structure matched those of the whole vault state. However, domains R1 and R2 showed great flexibility in the disengaged particle (Fig. 20.4c).

Along with vault history, and because of its curious architecture, several dynamic scenarios have been studied in different attempts to understand the role of vaults in cargo encapsulation and delivery. Early experiments showed a particle

dissociation in two vault-halves in a putative pH-dependent manner. When pH was less than 4.0, closed intact vaults dissociated quickly into half-vaults (Goldsmith et al. 2007). A reversible mechanism of half-vault dissociation induced by a pH change was proposed in light of the observed interactions between the R1 and R1 domains at the vault midsection (Querol-Audi et al. 2009). A salt bridge established between the universally conserved residues Asp39 and Arg42 (Fig. 20.1b) appeared to be a critical contact. In addition, the negatively charged amino acids Glu4, Glu5 and Arg37 and a cluster of hydrophobic residues (Ala6, Ile7 and Ile36), interacting through the twofold axis of the particle, also contributed to the contact interface (Fig. 20.1b). The strong electrostatic character of the interactions observed, with 234 acidic residues at the R1-R1 interface, prompted us to hypothesize that at low pH, these acidic residues would become neutral, leaving a highly electropositive charge and inducing the disassembly of the vault particle by charge repulsion (Querol-Audi et al. 2009). However, this hypothesis was challenged in later studies on the mechanical properties of vaults by atomic force microscopy (AFM), under physiological conditions (Llauró et al. 2014) and also at different pHs (Llauró et al. 2016). In these studies we demonstrated that under physiological conditions, vaults possess an intrinsic fracture/recuperation dynamics in which the particle was able to self-repair over time. Fractures did not occur in the vault midsection but along the lateral contacts between two neighbouring MVPs in the barrel structure, and disappeared over the time (Llauró et al. 2014). Further AFM experiments, performed to monitor the structural changes of individual vault particles while changing the pH in real time, showed that decreasing the pH of the solution destabilized the barrel region, but not the vault midsection, finally leading to particle aggregation. The observed topographical defects suggested that low pH weakened the polar bonds between adjacent proteins at the barrel and shoulder regions (Fig. 20.2) (Llauró et al. 2016).

Our recent studies of vault stabilization by combining site-directed mutagenesis of the R1 residues involved in contacts in the vault midsection (Fig. 20.1b) and different biophysical approaches, including AFM, negative staining EM, and differential scanning fluorimetry (DSF), confirmed the crucial role of both hydrophobic and electrostatic interactions, mediated by the MVP N-terminal residues, in the vault assembly process (Fig. 20.1b) (Guerra et al. 2022).

The role of residues present in the vault midsection was also supported by previous studies by Mrazek and colleagues, where the stabilities of several MVP mutants, containing a series of histidine substitutions at the R1 N-terminus, were analysed. These authors found that the substitutions with four, five and six histidines at amino acid positions from 3 to 8 resulted in the generation of unstable vault particles prone to disassembly (Mrazek et al. 2014). In this work, Mrazek and colleagues also revealed a peculiar aspect of the vault assembly process, showing that the biosynthesis of the MVP subunits and its assembly seem to take place sequentially at the level of polyribosomes, which act like a 3D nanoprinter. Translated chains on a polyribosome contact with their N-termini to form a dimer. Dimers make side-to-side contacts in the polyribosome until the 39 dimers associate and the new vault is released (Mrazek et al. 2014).

The Putative Vault Opening Model

Previous *in vivo* experiments performed at physiological pH by Yang and colleagues supported a model where vault particles were capable of rapidly separating at the particle waist and reassembling back into whole particles. In this way, vaults could release molecules stored inside, when used as recombinant nanocarriers or capture molecules from the environment that could be relevant to their physiological function (Yang et al. 2010).

The recent cryo-EM structures of vaults determined in three different states, closed symmetric, closed asymmetric and half-vaults (Guerra et al. 2022; Fig. 20.4), prompted us to hypothesize a mechanistic model for the vault opening-closing cycle with functional implications for cargo encapsulation, transport and delivery (Fig. 20.5). Fully assembled vaults exhibit a symmetrical configuration with all MVP R1 domains, from top and bottom half-vaults, fully engaged in reciprocal interactions (closed state: Figs. 20.4a and 20.5), rigidifying the waist cavity and allowing a stable inner compartment needed for cargo delivery. The opening process will begin with a waist relaxation, caused by a local loss of interactions between R1 domains of some specific MVPs, yielding a localized distortion that increases the angle between the R1 domains involved (relaxed state: Figs. 20.4b and 20.5). Local distortions of R1 domain interactions are progressively increased, priming the vault towards opening (primed state: Fig. 20.5). When the angle between these R1 domains exceeds a value (committed state: Fig. 20.5), the localized distortion can be easily transmitted beyond the area where it started, generating a break between R1 domains that will eventually trigger full separation of the two half-vaults (Fig. 20.5: opened state). The different conformational states appear to coexist in solution as seen in the cryo-EM studies (Guerra et al. 2022).

The Vault Minor Components

The Poly-(ADP)-Ribose Polymerase (vPARP)

In 1999, Kickhoefer and colleagues identified the 193 kDa vault component as a poly-(ADP)-ribose polymerase (vPARP) (Kickhoefer et al. 1999). Known collectively as members of the PARP family, these related enzymes are capable of catalysing the covalent transfer of ADP-riboses to acidic residues of certain proteins (poly-ADP-ribosylation) (Amé et al. 2004). In addition, the PARP family has a critical role in DNA repair and other genome stability processes, such transcription, replication and chromatin structure regulation (Meyer-Ficca et al. 2005).

The PARP family is formed by at least 18 members that are encoded by different genes, which shared the conserved catalytic domain (Morales et al. 2014). In particular, the vPARP dubbed as PARP-4 is predicted to contain a BRCA1 C-terminal (BRCT) domain at the N-terminus, followed by the catalytic domain, a vault inter-alpha-trypsin domain (VIT), a von Willebrand type a (vWa) domain, a nuclear

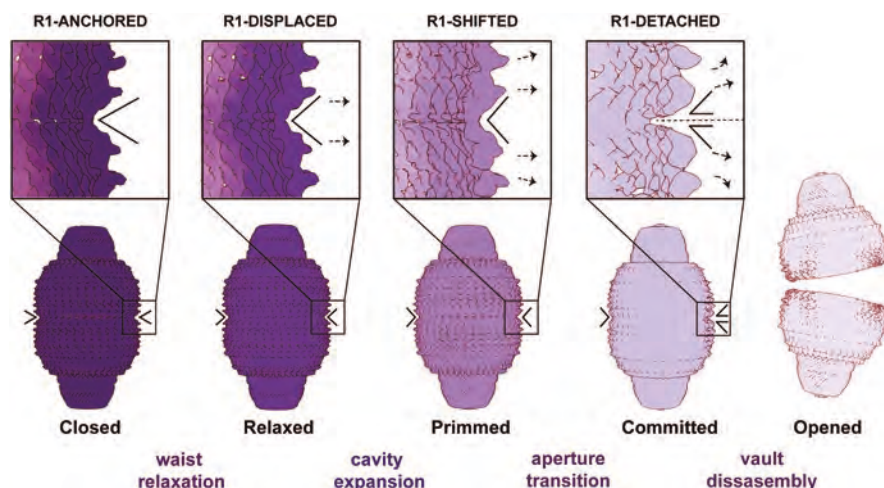


Fig. 20.5 Hypothetical model for the process of vault aperture. Fully assembled vaults exhibit a symmetrical configuration with all MVP R1 domains fully engaged, restricting the flexibility of MVPs, allowing a stable inner compartment needed for cargo delivery (Closed form). The opening process starts with a waist relaxation, caused by local displacement of R1 domains of specific MVPs, yielding a localized distortion that increases the angle between the R1 domains involved (Relaxed form). Local distortions of R1 domain interactions are progressively increased, priming the vault towards opening (Primed form). When the angle between these R1 domains exceeds a value (Committed form), the localized distortion can be easily transmitted beyond the area where it started, generating a “gap” between R1 domains that will eventually trigger full separation of the two half-vaults (Opened form) (Guerra et al. 2022)

localization sequence (NLS) and a MVP-interacting (INT) domain at the C-terminus (Fig. 20.6a). The INT domain binds MVP close to the vault midsection (Yu et al. 2017) and multiple times INT domain was used to direct heterologous proteins to the vault engineered interior (Kickhoefer et al. 2005; Goldsmith et al. 2009).

The number of vPARP molecules per vault remains unclear (4–16 copies); however, it is known that in vitro vPARP is able to poly-(ADP)-ribosylate itself and the MVP (Kickhoefer et al. 1999).

The Telomerase-Associated Protein-1 (TEP1)

As a homolog of *Tetrahymena* p80 protein, one of the protein components of telomerase complex of *Tetrahymena*, the telomerase-associated protein-1 (TEP1), was identified in 1997 by Harrington and colleagues (Harrington et al. 1997). The catalytic subunit of the telomerase complex is the reverse transcriptase (hTERT), and TEP1 is not a core component (Weinrich et al. 1997). Data from a TEP1-deficient mouse model demonstrated that TEP1 is not essential for telomerase activity or telomerase length maintenance in vivo (Liu et al. 2000; Liu et al. 2004). Cryo-EM reconstruction maps of TEP1-deficient mice vaults showed less density in the cap area than previously seen in wild-type vaults, suggesting that TEP1 was located in

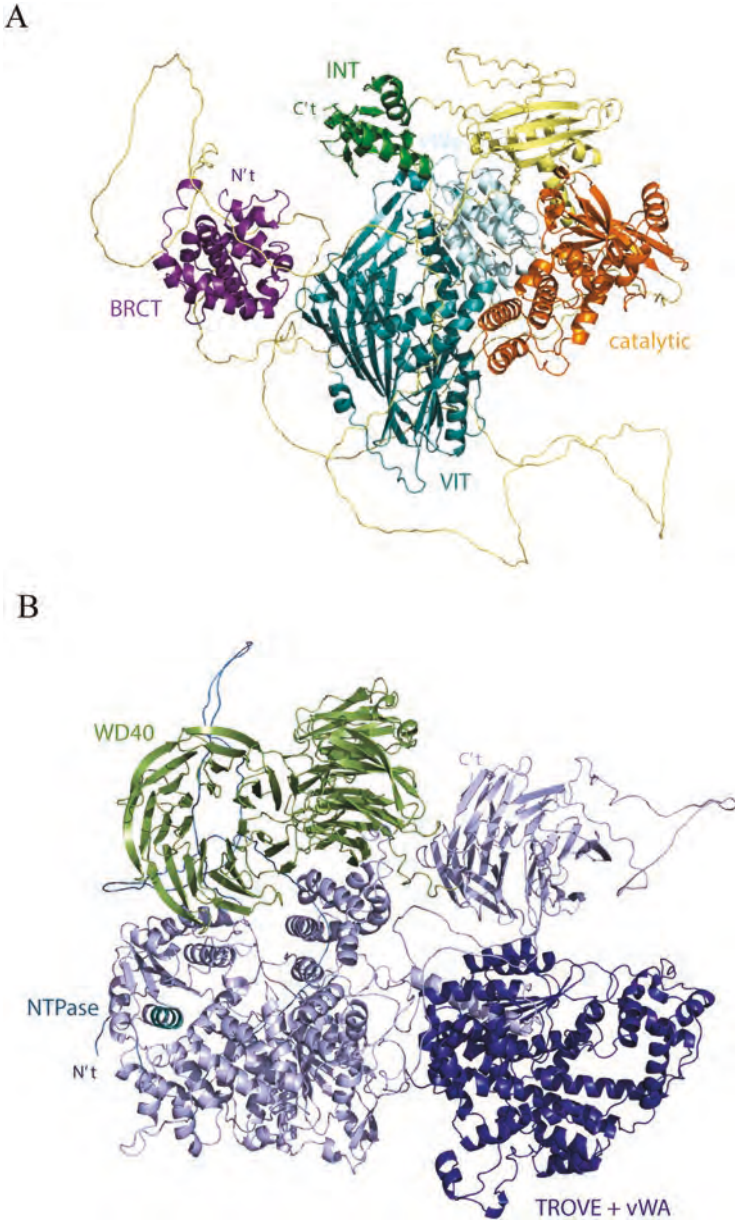


Fig. 20.6 AlphaFold prediction models for the vault minor protein components. **(a)** vPARP (UniProt: Q9UKK3) with the different domains shown in different colors: BRCT (magenta), catalytic (orange), VIT (dark blue), vWa (light blue) and INT (green). **(b)** TEP1 (UniProt: Q99973) with the different domains colored: the four 30-residue N-terminal repeats (dark blue), TROVE vWa domains (dark purple), NTPase (dark blue) and WD40 (green)

this region (Kickhoefer et al. 2001). Additionally, TEP1 interacts with several human vault RNAs (vtRNAs) (Kickhoefer et al. 1999; Poderycki et al. 2005), playing a significant role in localizing and maintaining the vtRNA associated with the vault particles (Kickhoefer et al. 2001).

The structure of the complete 240 kDa TEP1 protein is not yet known. It is predicted to have an N-terminal domain, consisting of four 30-residue repeats, followed by the *Tetrahymena* p80 homology region domain which contains a TROVE (telomerase, Ro and vault) domain and a vWa domain (Fig. 20.6b). RNA binding is carried out by the evolutionarily conserved TROVE domain, which is also found in the telomerase and Ro proteins (Bateman and Kickhoefer 2003). Downstream the vWa domain, there are an NTPase domain (Koonin and Aravind 2000) and a 16-fold WD40 repeat region, located at the C-terminus (Fig. 20.6b) (Harrington et al. 1997; Mikyás et al. 2004).

Vault RNAs (vtRNAs)

The vault particle is described as a ribonucleoprotein particle (RNP) due to its RNA component. Vault RNAs (vtRNAs) are small non-coding RNAs transcribed by the RNA polymerase III (Kickhoefer et al. 1993). Humans express four vtRNA genes of 88–100 nt long (vtRNA1–1, vtRNA1–2, vtRNA1–3, vtRNA2–1) (Stadler et al. 2009). Despite the fact that vtRNAs were the first vault component discovered, only a fraction is associated with vault particles. Outside the vault particle, vtRNAs are involved in different cellular processes: proliferation, cellular differentiation, mRNA regulation, cell-cell communication and cell death (apoptosis, autophagy) (Hahne et al. 2021). With the recent discovery that vaults arrive early in the autophagy process (Carter et al. 2020), and the fact that vtRNA is a regulator of selective autophagy (Horos et al. 2019), the question about the role of the vaults in autophagy has been reopened.

Direct Visualization of Vaults in Cells Is the Most Promising Approach Towards Deciphering the Vault Function

Cryo-electron tomography (cryo-ET) is an emerging technology that allows thin samples, such as macromolecular complexes, to be imaged in 3D in a nearly native state and at high resolution. Cryo-ET can provide insight into the positions and structures of cytoskeleton filaments, cell wall components, motility machines, chemoreceptor arrays, internal compartments and other cellular structures (Murphy and Jensen 2007). In this way, structural information about organelles and individual proteins can be linked with their exact three-dimensional arrangement within the cell. This unique capability holds enormous potential for cell biology, making

cryo-ET a highly promising technology to structurally visualize the vault complex in their natural microenvironment and characterize its molecular sociology.

During the writing of this chapter, we have found three publications, describing the visualization of native vaults in cells, by cryo-ET experiments. Woodward and colleagues showed that these particles were located in the extreme periphery of the cytoplasm, predominantly associated with granule-like structures and actin (Woodward et al. 2015). Vaults imaged within cells exhibited almost identical sizes and shapes to purified and recombinant vaults (Woodward et al. 2015). Authors also realized that the vault structure was not entirely rigid, but was able to adopt a range of conformations without significantly altering its overall dimensions. A number of vaults were seen with a very pronounced indentation at the particle waist and also, a few them, opening up at their waist to interact with closely associated granules (Woodward et al. 2015). The intrinsic plasticity of the vault shell seen in particles directly imaged in intact cells correlated well with the structural changes observed in the cryo-EM structures obtained from purified recombinant particles (Guerra et al. 2022).

Sub-tomogram averaging of intracellular vaults resulted in an averaged map, revealing additional densities within the vault lumen randomly distributed and high density at the vault's cap (Woodward et al. 2015), consistent with the previous vtRNA localization at the vault caps (Kong et al. 2000) (Kickhoefer et al. 2001).

In a second publication, Carter and colleagues captured vaults while studying how exogenous tripartite motif (TRIM) protein 5 α was forming autophagic aggregates *in vivo*. Cryo-tomograms provided three-dimensional views of various structures assembled during macroautophagy in their native state and indicated that vaults arrive early in the process, before the formation of visible aggregates (Carter et al. 2020). Their findings revealed a chronology in the autophagic pathway: colocalization of substrate, arrival of vault, formation of dense substrate aggregates (sequestosomes), creation of phagophores, phagophore closure around substrate to produce autophagosomes and finally autophagosome fusion with either endosomes or lysosomes to produce an autophagic vacuole. Because vaults arrived early, they were seen in all autophagic steps, with sequestosomes, phagophores and autophagosomes, as well as within autophagic vacuoles. The precise function of the vault within the autophagic structures remains unclear. vtRNA acts as a negative regulator of selective autophagy (Horos et al. 2019). Vaults in autophagic structures serve the same purpose as vtRNA. Or, on the contrary, do they have an inverse function? Another mystery surrounds the vault particle's enigmatic biological role is open, in this case, related with autophagy.

Very recently, Cruz-Leon and colleagues captured *in situ* vault-encapsulated ribosomes, supporting the hypothesis that vaults may enclose macromolecular complexes and transport them within the cell (Cruz-León et al. 2023). Further investigations are required to understand whether ribosome encapsulation occurs by transient opening of vaults or during vault biogenesis.

Acknowledgements This work was funded by the Spanish Ministry of Science and Innovation (PID2020-117976GB-I00) to NV. MGA work is supported by a predoctoral contract PRE2018-083964 from Spanish Ministry of Science and Innovation.

References

- Amé J-C, Spenlehauer C, de Murcia G (2004) The PARP super-family. *BioEssays* 26(8):882–893. <https://doi.org/10.1002/bies.20085>
- Anderson DH, Kickhoefer VA, Sievers SA, Rome LH, Eisenberg D (2007) Draft crystal structure of the vault shell at 9-Å resolution. *PLoS Biol* 5(11):2661–2670. <https://doi.org/10.1371/journal.pbio.0050318>
- Bateman A, Kickhoefer V (2003) The TROVE module: a common element in telomerase, Ro and Vault ribonucleoproteins. *BMC Bioinform* 4:49. <https://doi.org/10.1186/1471-2105-4-49>
- Benner NL, Zang X, Buehler DC, Kickhoefer VA, Rome ME, Rome LH, Wender PA (2017) Vault nanoparticles: chemical modifications for imaging and enhanced delivery. *ACS Nano* 11(1):872–881. <https://doi.org/10.1021/acsnano.6b07440>
- Berger W, Steiner E, Grusch M, Elbling L, Micksche M (2009) Vaults and the major vault protein: novel roles in signal pathway regulation and immunity. *Cell Mol Life Sci* 66(1):43–61. <https://doi.org/10.1007/s00018-008-8364-z>
- Carter SD, Mamede JI, Hope TJ, Jensen GJ (2020) Correlated cryogenic fluorescence microscopy and electron cryo-tomography shows that exogenous TRIM5 α can form hexagonal lattices or autophagy aggregates in vivo. *Proc Natl Acad Sci USA* 117:29702–29711. <https://doi.org/10.1073/pnas.1920323117/-DCSupplemental>
- Casañas A, Guerra P, Fita I, Verdager N (2012) Vault particles: a new generation of delivery nanodevices. *Curr Opin Biotechnol* 23(6):972–977. <https://doi.org/10.1016/j.copbio.2012.05.004>
- Casañas A, Querol-Audí J, Guerra P, Pous J, Tanaka H, Tsukihara T, Verdager N, Fita I (2013) New features of vault architecture and dynamics revealed by novel refinement using the deformable elastic network approach. *Acta Crystallogr D Biol Crystallogr* 69(6):1054–1061. <https://doi.org/10.1107/S0907444913004472>
- Cruz-León S, Majtner T, Hoffmann PC, Kreysing JP, Tuijtel MW, Schaefer SL, Geißler K, Beck M, Turoňová B, Hummer G (2023) High-confidence 3D template matching for cryo-electron tomography. *BioRxiv*. <https://doi.org/10.1101/2023.09.05.556310>
- Ding K, Zhang X, Mrazek J, Kickhoefer VA, Lai M, Ng HL, Yang OO, Rome LH, Zhou ZH (2018) Solution structures of engineered vault particles. *Structure* 26(4):619–626.e3. <https://doi.org/10.1016/j.str.2018.02.014>
- Frascotti G, Galbiati E, Mazzucchelli M, Pozzi M, Salvioni L, Vertemara J, Tortora P (2021) Cancers the vault nanoparticle: a gigantic ribonucleoprotein assembly involved in diverse physiological and pathological phenomena and an ideal Nanovector for drug delivery and therapy. *Cancers* 13:707. <https://doi.org/10.3390/cancers>
- Fulcher JA, Tamshen K, Wollenberg AL, Kickhoefer VA, Mrazek J, Elliott J, Ibarrondo FJ, Anton PA, Rome LH, Maynard HD, Deming T, Yang OO (2019) Human vault nanoparticle targeted delivery of antiretroviral drugs to inhibit human immunodeficiency virus type 1 infection. *Bioconjug Chem* 30(8):2216–2227. <https://doi.org/10.1021/acs.bioconjchem.9b00451>
- Goldsmith LE, Yu M, Rome LH, Monbouquette HG (2007) Vault nanocapsule dissociation into halves triggered at low pH. *Biochemistry* 46(10):2865–2875. <https://doi.org/10.1021/bi0606243>
- Goldsmith LE, Pupols M, Kickhoefer VA, Rome LH, Monbouquette HG (2009) Utilization of a protein “shuttle” to load vault nanocapsules with gold probes and proteins. *ACS Nano* 3(10):3175–3183. <https://doi.org/10.1021/nn900555d>

- Guerra P, González-Alamos M, Llauro A, Casañas A, Querol-Audí J, de Pablo PJ, Verdager N (2022) Symmetry disruption commits vault particles to disassembly. *Science*. Advances 8(6):eabj7795. <https://doi.org/10.1126/sciadv.abj7795>
- Hahne JC, Lampis A, Valeri N (2021) Vault RNAs: hidden gems in RNA and protein regulation. *Cell Mol Life Sci* 78(4):1487–1499. Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/s00018-020-03675-9>
- Hamill DR, Suprenant KA (1997) Characterization of the sea urchin major vault protein: a possible role for vault ribonucleoprotein particles in nucleocytoplasmic transport. *Dev Biol* 190(1):117–128. <https://doi.org/10.1006/dbio.1997.8676>
- Harrington L, McPhail T, Mar V, Zhou W, Oulton R, Bass MB, Arruda I, Robinson MO (1997) A mammalian telomerase associated protein. *Science* 275:973–977
- Herrmann C, Zimmermann H, Volknandt W (1997) Analysis of a cDNA encoding the major vault protein from the electric ray *Discopyge ommata*. *Gene* 188:85–90
- Horos R, Büscher M, Kleinendorst R, Alleaume AM, Tarafder AK, Schwarzl T, Dziuba D, Tischer C, Zielonka EM, Adak A, Castello A, Huber W, Sachse C, Hentze MW (2019) The small non-coding vault RNA1-1 acts as a Riboregulator of autophagy. *Cell* 176(5):1054–1067.e12. <https://doi.org/10.1016/j.cell.2019.01.030>
- Kedersha NL, Rome LH (1986) Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA. *J Cell Biol* 103:699–709. <http://rupress.org/jcb/article-pdf/103/3/699/1053878/699.pdf>
- Kedersha NL, Hill DE, Kronquist KE, Rome LH (1986) Subpopulations of liver coated vesicles resolved by preparative agarose gel electrophoresis. *J Cell Biol* 103(1):287–297. <http://rupress.org/jcb/article-pdf/103/1/287/1053272/287.pdf>
- Kedersha NL, Miquel MC et al (1990) Vaults. II. Ribonucleoprotein structures are highly conserved among higher and lower eukaryotes. *J Cell Biol* 110(4):895–901
- Kickhoefer VA, Searles RP, Kedersha NL, Garber ME, Johnson DL, Rome LH (1993) Vault ribonucleoprotein particles from rat and bullfrog contain a related small RNA that is transcribed by RNA polymerase III. *J Cell Biol* 268(11):7868–7873
- Kickhoefer VA, Vasu SK et al (1996) Vaults are the answer, what is the question? *Trends Cell Biol* 6(5):174–178
- Kickhoefer VA, Siva AC et al (1999a) The 193-kD vault protein, VPARP, is a novel poly(ADP-ribose) polymerase. *J Cell Biol* 146(5):917–928
- Kickhoefer VA, Stephen AG et al (1999b) Vaults and telomerase share a common subunit, TEPI. *J Cell Biol* 274(46):32712–32717. <http://www.jbc.org>
- Kickhoefer VA, Liu Y, Kong LB, Snow BE, Stewart PL, Harrington L, Rome LH (2001) The telomerase/vault-associated protein TEPI is required for vault RNA stability and its association with the vault particle. *J Cell Biol* 152(1):157–164. <http://www.jcb.org/cgi/content/full/152/1/157>
- Kickhoefer VA, Garcia Y, Mikyas Y, Johansson E, Zhou JC, Raval-Fernandes S, Minoofar P, Zink JJ, Dunn B, Stewart PL, Rome LH (2005) Engineering of vault nanocapsules with enzymatic and fluorescent properties. *Proc Natl Acad Sci USA* 102(12):4348–4352. <https://doi.org/10.1073/pnas.0500929102>
- Kickhoefer VA, Han M, Raval-Fernandes S, Poderycki MJ, Moniz RJ, Vaccari D, Silvestry M, Stewart PL, Kelly KA, Rome LH (2009) Targeting vault nanoparticles to specific cell surface receptors. *ACS Nano* 3(1):27–36. <https://doi.org/10.1021/nn800638x>
- Kong LB, Siva AC, Rome LH, Stewart PL (1999) Structure of the vault, a ubiquitous cellular component. *Structure* 7(4):371–379. <http://biomednet.com/elecref/0969212600700371>
- Kong LB, Siva AC, Kickhoefer VA, Rome LH, Stewart PL (2000) RNA location and modeling of a WD40 repeat domain within the vault. *RNA* 6(6):890–900
- Koonin EV, Aravind L (2000) the NACHT family—a new group of predicted NTPases implicated in apoptosis and MHC transcription activation. *Trends Biochem Sci* 25:223–224
- Liu Y, Snow BE, Prakash Hande M, Baerlocher G, Kickhoefer VA, Yeung D, Wakeham A, Itie A, Siderovski DP, Lansdorp PM, Robinson MO, Harrington L (2000) Telomerase-associated protein TEPI is not essential for telomerase activity or telomere length maintenance in vivo. *Mol Cell Biol* 20(21):8178–8184

- Liu Y, Snow BE, Kickhoefer VA, Erdmann N, Zhou W, Wakeham A, Gomez M, Rome LH, Harrington L (2004) Vault poly (ADP-ribose) polymerase is associated with mammalian telomerase and is dispensable for telomerase function and vault structure in vivo. *Mol Cell Biol* 24(12):5314–5323. <https://doi.org/10.1128/mcb.24.12.5314-5323.2004>
- Llauró A, Guerra P, Irigoyen N, Rodríguez JF, Verdaguer N, de Pablo PJ (2014) Mechanical stability and reversible fracture of vault particles. *Biophys J* 106(3):687–695. <https://doi.org/10.1016/j.bpj.2013.12.035>
- Llauró A, Guerra P, Kant R, Bothner B, Verdaguer N, de Pablo PJ (2016) Decrease in pH destabilizes individual vault nanocages by weakening the inter-protein lateral interaction. *Sci Rep* 6:34143. <https://doi.org/10.1038/srep34143>
- Meyer-Ficca ML, Meyer RG, Jacobson EL, Jacobson MK (2005) Poly(ADP-ribose) polymerases: managing genome stability. *Int J Biochem Cell Biol* 37(5):920–926
- Mikyas Y, Makabi M, Raval-Fernandes S, Harrington L, Kickhoefer VA, Rome LH, Stewart PL (2004) Cryoelectron microscopy imaging of recombinant and tissue derived vaults: localization of the MVP N termini and VPARP. *J Mol Biol* 344(1):91–105. <https://doi.org/10.1016/j.jmb.2004.09.021>
- Morales JC, Li L, Fattah FJ, Dong Y, Bey EA, Patel M, Gao J, Boothman DA (2014) Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases. *Crit Rev Eukaryot Gene Expr* 24:15–28
- Mrazek J, Toso D, Ryazantsev S, Zhang X, Zhou ZH, Fernandez BC, Kickhoefer VA, Rome LH (2014) Polyribosomes are molecular 3D nanoprinters that orchestrate the assembly of vault particles. *ACS Nano* 8(11):11552–11559. <https://doi.org/10.1021/nn504778h>
- Murphy GE, Jensen GJ (2007) Electron cryotomography. *BioTechniques* 43(4):413–423. <https://doi.org/10.2144/000112568>
- Nagasawa DT, Yang J, Romiyo P, Lagman C, Chung LK, Voth BL, Duong C, Kickhoefer VA, Rome LH, Yang I (2020) Bioengineered recombinant vault nanoparticles coupled with NY-ESO-1 glioma-associated antigens induce maturation of native dendritic cells. *J Neuro-Oncol* 148(1):1–7. <https://doi.org/10.1007/s11060-020-03472-1>
- Poderycki MJ, Rome LH, Harrington L, Kickhoefer VA (2005) The p80 homology region of TEP1 is sufficient for its association with the telomerase and vault RNAs, and the vault particle. *Nucleic Acids Res* 33(3):893–902. <https://doi.org/10.1093/nar/gki234>
- Poderycki MJ, Kickhoefer VA, Kaddis CS, Raval-Fernandes S, Johansson E, Zink JJ, Loo JA, Rome LH (2006) The vault exterior shell is a dynamic structure that allows incorporation of vault-associated proteins into its interior. *Biochemistry* 45(39):12184–12193. <https://doi.org/10.1021/bi0610552>
- Querol-Audí J, Perez-Luque R, Fita I, Lopéz-Iglesias C, Castón JR, Carrascosa JL, Verdaguer N (2005) Preliminary analysis of two and three dimensional crystals of vault ribonucleoprotein particles. *J Struct Biol* 151(1):111–115. <https://doi.org/10.1016/j.jsb.2005.04.002>
- Querol-Audí J, Casãs A, Usón I, Luque D, Castón JR, Fita I, Verdaguer N (2009) The mechanism of vault opening from the high-resolution structure of the N-terminal repeats of MVP. *EMBO J* 28(21):3450–3457. <https://doi.org/10.1038/emboj.2009.274>
- Rome LH, Kickhoefer VA (2013) Development of the vault particle as a platform technology. *ACS Nano* 7(2):889–902. <https://doi.org/10.1021/nn3052082>
- Rome L, Kederasha N et al (1991) Unlocking vaults: organelles in search of a function. *Trends Cell Biol* 1(2–3):47–50
- Stadler PF, Chen JJJ, Hackermüller J, Hoffmann S, Horn F, Khaitovich P, Kretschmar AK, Mosig A, Prohaska SJ, Qi X, Schutt K, Ullmann K (2009) Evolution of vault RNAs. *Mol Biol Evol* 26(9):1975–1991. <https://doi.org/10.1093/molbev/msp112>
- Stephen AG, Raval-Fernandes S, Huynh T, Torres M, Kickhoefer VA, Rome LH (2001) Assembly of vault-like particles in insect cells expressing only the major vault protein. *J Biol Chem* 276(26):23217–23220. <https://doi.org/10.1074/jbc.C100226200>
- Tanaka H, Kato K, Yamashita E, Sumizawa T, Zhou Y, Yao M, Iwasaki K, Yoshimura M, Tsukihara T (2009) The structure of rat liver vault at 3.5 angstrom resolution. *Science* 323(5912):384–388

- Van Zon A, Mossink MH, Schoester M, Scheffer GL, Scheper RJ, Sonneveld P, Wiemer EAC (2001) Multiple human vault RNAs: expression and association with the vault complex. *J Biol Chem* 276(40):37715–37721. <https://doi.org/10.1074/jbc.M106055200>
- Voth BL, Pelargos PE, Barnette NE, Bhatt NS, Chen CHJ, Lagman C, Chung LK, Nguyen T, Sheppard JP, Romiyo P, Mareninov S, Kickhoefer VA, Yong WH, Rome LH, Yang I (2020) Intratumor injection of CCL21-coupled vault nanoparticles is associated with reduction in tumor volume in an in vivo model of glioma. *J Neuro-Oncol* 147(3):599–605. <https://doi.org/10.1007/s11060-020-03479-8>
- Wang M, Abad D, Kickhoefer VA, Rome LH, Mahendra S (2015) Vault nanoparticles packaged with enzymes as an efficient pollutant biodegradation technology. *ACS Nano* 9(11):10931–10940. <https://doi.org/10.1021/acs.nano.5b04073>
- Wang M, Kickhoefer VA, Rome LH, Foellmer OK, Mahendra S (2018) Synthesis and assembly of human vault particles in yeast. *Biotechnol Bioeng* 115(12):2941–2950. <https://doi.org/10.1002/bit.26825>
- Weinrich SL, Pruzan R, Ma L, Ouellette M, Tesmer VM, Holt SE, Bodnar AG, Lichtsteiner S, Kim NW, Trager JB, Taylor RD, Carlos R, Andrews WH, Wright WE, Shay J, Harley CB, Morin GB (1997) Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTERT. *Nat Genet* 17(4):498–502
- Woodward CL, Mendonça LM, Jensen GJ (2015) Direct visualization of vaults within intact cells by electron cryo-tomography. *Cell Mol Life Sci* 72(17):3401–3409. <https://doi.org/10.1007/s00018-015-1898-y>
- Yang J, Kickhoefer VA, Ng BC, Gopal A, Bentolila LA, John S, Tolbert SH, Rome LH (2010) Vaults are dynamically unconstrained cytoplasmic nanoparticles capable of half vault exchange. *ACS Nano* 4(12):7229–7240. <https://doi.org/10.1021/nn102051r>
- Yu K, Yau YH, Sinha A, Tan T, Kickhoefer VA, Rome LH, Lee H, Shochat SG, Lim S (2017) Modulation of the vault protein-protein interaction for tuning of molecular release. *Sci Rep* 7(1):14816. <https://doi.org/10.1038/s41598-017-12870-x>

Chapter 21

Structure and Dynamics of Type 4a Pili and Type 2 Secretion System Endopili



Nadia Izadi-Pruneyre, Yasaman Karami, and Michael Nilges

Abstract Within the highly diverse type four filament (TFF or T4F) superfamily, the machineries of type IVa pili (T4aP) and the type 2 secretion system (T2SS) in diderm bacteria exhibit a substantial sequence similarity despite divergent functions and distinct appearances: T4aP can extend micrometers beyond the outer membrane, whereas the endopili in the T2SS are restricted to the periplasm. The determination of the structure of individual components and entire filaments is crucial to understand how their structure enables them to serve different functions. However, the dynamics of these filaments poses a challenge for their high-resolution structure determination. This review presents different approaches that have been used to study the structure and dynamics of T4aP and T2SS endopili by means of integrative structural biology, cryo-electron microscopy (cryo-EM), and molecular dynamics simulations. Their conserved features and differences are presented. The non-helical stretch in the long-conserved N-terminal helix which is characteristic of all members of the TFF and the impact of calcium on structure, function, and dynamics of these filaments are discussed in detail.

Keywords Integrative structural biology · Pilus · Secretion system · Dynamics

N. Izadi-Pruneyre

Bacterial Transmembrane Systems Unit, Institut Pasteur, Université Paris Cité, CNRS UMR, Paris, France

e-mail: nadia.izadi@pasteur.fr

Y. Karami · M. Nilges (✉)

Structural Bioinformatics Unit, Institut Pasteur, Université Paris Cité, CNRS UMR, Paris, France

e-mail: yasaman.karami@inria.fr; michael.nilges@pasteur.fr

Importance of the Filaments of the Type Four Filament (TFF) Superfamily

Bacteria have evolved fascinating structures to perform a wide variety of biological functions necessary for their survival, social organization, motility, and capacity of infecting host cells. Bacteria frequently engage with their surroundings by utilizing surface polymers known as pili. These pili are constructed through a regular repetition of one or a few protein subunits called pilins. In certain instances, pilins form covalent connections with one another, resulting in the formation of robust and mechanically resilient fibers. This phenomenon is observed in sortase-dependent pili in Gram-positive bacteria (Kang and Baker 2012). Other pili are maintained by noncovalent interactions between subunits to combine mechanical strength with flexibility and dynamics.

The type 4 filament (TFF) superfamily is a vast group of membrane-associated molecular systems in bacteria and archaea. This superfamily is remarkable for the range of function of its members and for their distribution. Members of this superfamily are present in monoderm and diderm bacteria, and archaea (Denise et al. 2019). Within diderm bacteria, the type 4 filament superfamily encompasses type IVa (T4aP) and type IVb (T4bP) pili, the type II secretion system (T2SS), the mannose-sensitive hemagglutinin pilus, and the tight adhesion pilus. In monoderm bacteria, the TFF is represented by the competence system; and, in archaea, by the archaeal flagellum or archaellum. The functions attributed to members of this superfamily include adhesion, protein secretion, DNA uptake, motility, microcolony formation, and interactions with host cells. Significantly, many members are known to be virulence factors or to contribute to bacterial pathogenicity. An analysis of the multifaceted roles and evolutionary significance of this superfamily provides not only deeper insights into microbial survival strategies but also potential avenues for therapeutic interventions and biotechnological applications.

This chapter is devoted to the structural properties of filaments of the T4aP and the T2SS families. Like in all members of the TFF superfamily, the pilins of T4aP and the T2SS are noncovalently linked, allowing them to rapidly assemble and disassemble (Craig et al. 2019). Pilus dynamics are crucial for their biological functions, which include adhesion to host cells, twitching motility, DNA uptake, protein secretion, and microcolony formation. T4aP fibers can reach several micrometers in length and range between 6 and 9 nm in diameter. T4aP are anchored in the inner membrane (IM) and extend beyond the outer membrane (OM) of Gram-negative bacteria through a channel called the secretin. The closest homologue of the T4aP system is the T2SS (Denise et al. 2019), which also needs to readily assemble and disassemble to fulfil its function. Despite the sequence similarity, they exhibit rather different functions and overall structural properties. The T2SS exports folded proteins from the periplasm to the extracellular space and, under physiological conditions, does not extend beyond the OM.

The filaments are part of complex multiprotein machineries that span, in the case of diderm bacteria, the IM and the OM, including the periplasmic space (see

Fig. 21.1 for a schematic overview). This machinery is, among other things, responsible for assembling the filaments, and thus extending the pilus, and for retraction in many cases. The machinery is driven by an ATPase at the base of the system in the intracellular space, to add pilins at the base of the pilus, facilitating filament extension. Notably, T4aP usually feature a secondary ATPase, distinctively tasked with the active retraction of the filament.

Some pili do not extend beyond the OM in systems and uniquely serve for secretion, notably the T2SS and the *Geobacter sulfurreducens* PilA-N-C T4P (Gu et al. 2021). In these systems, growth of the filament stops when it reaches the part of the machinery anchored in the OM. Here, the filament is called an endopilus (historically called pseudopilus for the T2SS) to underline the fact that it does not extend to the extracellular space. Note that the *Geobacter* system is usually annotated as a T4aP in sequence databases but has a T2SS-like function. In a recent phylogenetic analysis (Denise et al. 2019), it is classified as a “generic” TFF. Systems with endopili do not possess a second ATPase for retraction.

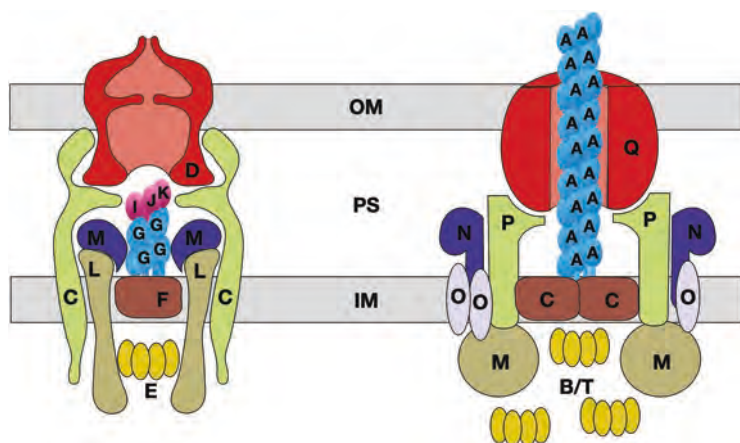


Fig. 21.1 Schematic overview of the overall architecture of the T2SS (left) and the T4aP (right). T2SS endopili do not extend beyond the outer membrane and are confined to the periplasmic space between the inner and outer membrane. The exact length is not known since it is unclear how far it extends into the channel of the secretin in the outer membrane (in red). In total, there are probably around ten pilins in the endopilus, four so-called minor pilins (shown in pink, called PulH, PulI, PulJ, PulK in *K. oxytoca*) that probably form the tip of the endopilus, and five or six major pilins that form the bulk (called PulG in *K. oxytoca*). T4P pili extend beyond the outer membrane, and the bulk of these pili is made up of hundreds of pilins. The T4aP also has four minor pilins, again probably at the tip, which are not shown in Fig. 21.1. For both systems, the names of the individual proteins or subcomplexes are indicated by a single letter. For the T4aP assembly machinery, the identifiers from the *Myxococcus xanthus* system (Chang et al. 2016) are used (“A” corresponds to the protein PilA, etc.), for the T2SS, the full protein names are and PulG, PulI, and PulJ, etc. for *K. oxytoca*. A comprehensive table of protein names of minor pilins in different T4aP and T2SS can be found in Jacobsen et al. (2020). The outer membrane (OM), the periplasmic space (PS), and the inner membrane (IM) are indicated

Central to the stability, flexibility, and functionality of these bacterial filaments are the properties of the pilins that constitute them. The bulk of the filament is composed of so-called major pilins, whose properties and mutual interactions determine the overall structure, stability, rigidity, and flexibility. The mechanical and dynamic properties are crucial for the function of the filaments. In contrast, minor pilins, though fewer in number, play several critical roles. They are likely involved in ensuring the proper assembly of the filament, adding an important layer of orchestration to the process of pilus assembly. They also are involved in key interactions between the filament and its partners, for example, the cargo of the T2SS.

This chapter reviews the current body of knowledge about the structure and dynamics of these filaments, and the physicochemical factors that are responsible. The “resolution revolution” (Kühlbrandt 2014) in structural biology due to the advent of cryo-electron microscopy (cryo-EM) at atomic or near-atomic resolution has had a particularly strong impact. We will review strategies employed for deciphering their three-dimensional structures, focusing on two members where our own contributions have been major. One is the T2SS endopilus from *Klebsiella oxytoca*, which is to date the only T2SS endopilus with structural data, the other one being the T4aP of Enterohemorrhagic *Escherichia coli* (EHEC). Through the comparison of the structures and properties of these members of two different families within the TFF, completed with other data from the literature, we aim to understand the intricate mechanisms underlying their architecture, dynamics, and function.

Structure Determination of TFF Filaments

X-ray crystallographic structural analysis of pilin subunits is challenging because their tendency to form fibers impedes crystallization (Figs. 21.2 and 21.3). Only very few atomic structures of full-length pilins of a type IV pilin have been determined, notably those of the gonococcal (GC) pilin from *Neisseria gonorrhoeae* (Parge et al. 1995; Craig et al. 2006) and the PAK pilin from *Pseudomonas aeruginosa* (Craig et al. 2003). The pilin structure is characterized by a long N-terminal helix, whose sequence is highly conserved within the TFF superfamily, and which protrudes from the globular head domain (see Fig. 21.4c). This helix was predicted to mediate fiber formation via hydrophobic association (Folkhard et al. 1981, Parge et al. 1995). Deciphering the 3D structure of isolated pilins within the TFF superfamily presents a challenge, primarily due to the extreme hydrophobicity of the N-terminal helix that extends from the globular head domain. To facilitate the study of pilins, the first part of the N-terminal helix corresponding to the transmembrane region is usually cleaved off, to prevent aggregation that would hinder a structure determination by X-ray crystallography or NMR.

Early work using fiber diffraction (Folkhard et al. 1981) showed a helical arrangement of the pilins. These data were insufficient to explain how the pili could be at the same time highly flexible but mechanically stable, and how they could fulfil the diverse functions that members of the T4aP and T2SS families are involved

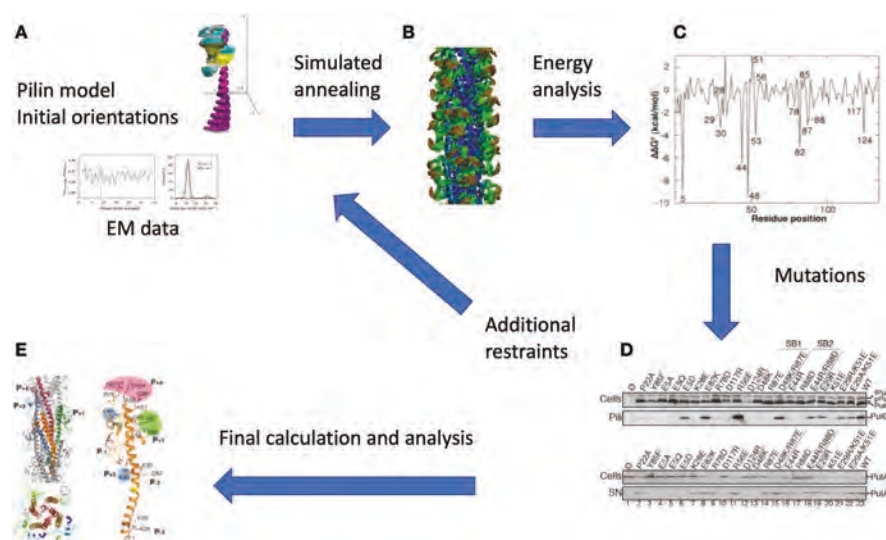


Fig. 21.2 Illustration of the integrative structural biology approach to determine the structure of T2SS pili. (a) Starting point is a homology model of the T2SS pilin together with information on overall symmetry of the pilus derived from electron microscopy. A single pilin is placed in varying orientations close to the overall helix axis. The other pilins are generated during the calculation by applying symmetry operators for rotation and translation corresponding to the helical symmetry of the pilus. (b) This structure is minimized with a simulated annealing protocol. (c) The energetic contribution of each residue in the pilin to the stability of the pilus is then analyzed by the generalized-Born approach. (d) Residues predicted to have the largest contribution are mutated, and the effect on piliation and secretion is analyzed experimentally. Salt bridges predicted from models are validated by single and compensatory charge inversion of two charged residues that are in close proximity in the models. These are the basis of additional distance restraints used in the simulated annealing calculations. (e) After a few cycles, a final set of pilus structures is calculated and analyzed (Campos et al. 2010, Campos et al. 2011). Parts of this Fig. were published elsewhere: the EM data diagram in (a) in Köhler et al. (2004); the energy analysis (c), the gels showing the effects of mutations (d) and the final structure (e) in Campos et al. (2010)

in. Electron microscopy (EM) studies of a T4aP from *Neisseria gonorrhoeae* (gonococcal or GC pilus; Craig et al. 2006) and T2SS filaments from *Klebsiella oxytoca* (Köhler et al. 2004, Campos et al. 2010) started to help us obtain an increasingly precise picture of the molecular architecture at atomic resolution and explain the mechanical and functional properties. T2SS endopili and T4aP pili share a significant sequence similarity. This sequence similarity translated into structural similarity, and for both filaments, the EM reconstructions showed a helical arrangement of the pilins, with dense packing of the pilins without a central cavity.

Achieving an atomic-level model of the filaments exclusively from EM data proved elusive due to insufficient resolution. In the first study (Köhler et al. 2004), not even the handedness of the helical arrangement could be unambiguously determined. All studies relied on the integration of pre-existing subunit structures into the EM data. Two rather distinct strategies were employed for the two systems. For the model of the GC pilus based on 12.5 Å EM data (Craig et al. 2006),

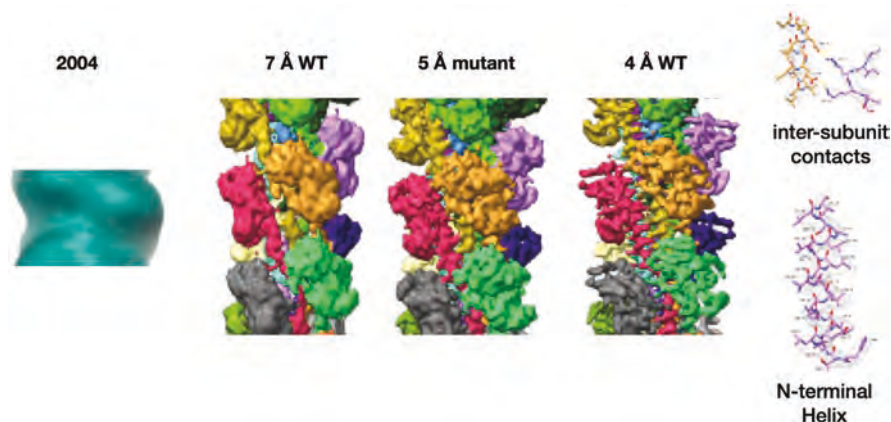


Fig. 21.3 Impact of the development of electron microscopes and detectors on the resolution of the T2SS filaments. All pili shown are from the *K. oxytoca* T2SS. For comparison, the reconstruction from Köhler et al. (2004) is shown on the left. In a normal bacterium, endopili are only a few subunits in length and remain in the periplasm (see Fig. 21.1). To obtain samples that can readily be extracted from bacteria to be characterized by (cryo) electron microscopy, the major pilin was overexpressed, resulting in long pili that extend beyond the bacterial surface (Köhler et al. 2004). The iterative helical real space reconstruction (IHRSR) method (Egelman 2000) was used to determine the helical parameters of the cryo-EM maps. The resolution of each reconstruction is indicated at the bottom. The data for at 7.5 Å (unpublished) and 5 Å (López-Castilla et al. 2017) were obtained on a Titan Krios with a Falcon II direct electron detector, while the one at 4 Å (unpublished) was obtained on a Titan Krios with a K3 direct electron detector. Examples of atomic details from the 4 Å reconstruction, such as a double salt bridge that had been shown by charge inversion experiments (Campos et al. 2010), are shown on the right

the X-ray crystal structure encompassing the full-length pilin of the T4aP from *N. gonorrhoeae* was rigidly docked into the EM reconstruction. In contrast, Campos and colleagues (Campos et al. 2010) based their model on the existing EM data (Köhler et al. 2004) and used an iterative flexible docking procedure, employing strict conformational constraints describing the overall symmetry of the filaments derived from the EM data in addition to distance restraints obtained from mutational studies (Fig. 21.2). The starting point in this case was a 3D model of the pilin derived from the X-ray crystal structure of a truncated pilin for the *K. oxytoca* filament (Köhler et al. 2004), albeit solved in a physiologically meaningless dimeric conformation in the absence of calcium, supplemented by the structure of a truncated pilin of a homologous filament from *E. coli* (Korotkov et al. 2009), and the N-terminal helix of the *N. gonorrhoeae* T4aP pilin (Craig et al. 2006). We think that the method remains of interest even with the much higher resolution obtainable by cryo-EM today, and explain the method in some detail in Fig. 21.2.

The models of the T4aP GC pilus from *N. gonorrhoeae* and the T2SS endopilus from *K. oxytoca* largely improved our understanding of the overall architecture and gave some atomic details that are the basis of the properties of these filaments.

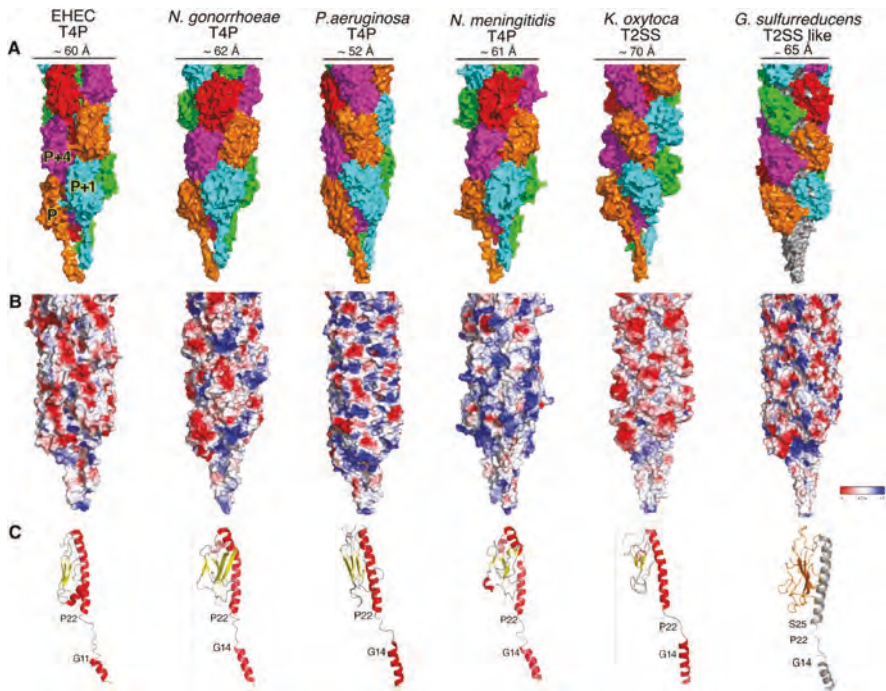


Fig. 21.4 Comparison of structures from different T4aP and T2SS. **(a)** Surface representation of pili, different subunits are colored differently. The *G. sulfurreducens* pilin is composed of two subunits, a globular one and an N-terminal helix (colored in gray). **(b)** Surface representation with electrostatic coloring (blue, positive; red, negative). **(c)** Structure of the pilins in the context of the pilus, clearly showing the conservation of the non-helical region within the N-terminal helix. The positions of “melted helix” between conserved glycine and prolines are indicated. Beta-sheets are in yellow, helices are in red except in *G. sulfurreducens* pilin in which the globular subunit is in brown, and the helical subunit is in gray. The structures are from PDB: EHEC T4P, 6GV9; *N. gonorrhoeae* T4P, 5VXX; *P. aeruginosa* T4P, 5VXY; *N. meningitidis* T4P, 5KUA; *K. oxytoca* T2SS, 5WDA; *G. sulfurreducens* PilA-N-C, 7TGG. In addition to the pili shown here, structures of the two distinct T4P from *Thermus thermophilus* were solved (Neuhaus et al., 2020). Figure adapted from (Bardiaux et al., 2019a)

Notably, the two structures indicated a common architecture for T4aP and T2SS pili and highlighted many residues important for the structure and stability of the filaments.

Despite the insights provided by these two models, and despite the extensive experimental validation for the T2SS endopilus, later higher-resolution structures (Kolappan et al. 2016; Wang et al. 2017; López-Castilla et al. 2017) showed that they were not entirely correct. Certain details were derived from prediction rather than experiment. The underlying assumption that the structure of the isolated pilin is not substantially modified when forming the pilus did not hold completely, although many details were correctly predicted (e.g., the double salt bridge stabilizing directly neighboring subunits shown in Fig. 21.3). That the 3D structure of a

molecule can be altered when it passes from a free isolated state to a complexed state poses a challenge for integrative structural biology methods in general.

The resolution revolution in structural biology has transformed our comprehension of TFF filaments. With advanced cryo-EM and direct electron detectors, a much higher resolution can be obtained, which enables us to visualize TFF filament structures at atomic or near-atomic levels. This higher resolution has unveiled hitherto hidden structural features, shedding light on the intricate arrangement of pilin subunits within these filaments. This detailed structural information obtained due to the resolution revolution has refined our understanding of the architecture and assembly mechanisms of TFF filaments.

This is illustrated by the EM and cryo-EM reconstructions of *K. oxytoca* T2SS endopilus over time (see Fig. 21.3).

The much higher resolution of the cryo-EM data revealed that the pilus models were incorrect in an important detail, the interruption of helicity within the long N-terminal helix (also referred to as “melted helix”), roughly between the conserved residues Pro 22 and Gly 14 (Fig. 21.4c). This came as a surprise since this helix is uninterrupted over more than 50 amino acid residues in the X-ray crystal structure of the full-length pilins (Craig et al. 2003; Craig et al. 2006), as well as in the NMR structure of the entire N-terminal helix of the *G. sulfurreducens* pilin (Reardon and Müller, 2013) obtained under close to native conditions. This system has the peculiarity that the pilin protein is composed of two polypeptides, one being the N-terminal helix, the other the globular head group.

The “melted helix” characteristic was first observed in the *Neisseria meningitidis* T4aP structure (Kolappan et al. 2016). This structural feature is consistently found in the structures of all pili from the T4aP and T2SS families that have been determined so far (Fig. 21.4c), as expected from the sequence conservation. This similarity extends to the pilus from *G. sulfurreducens* (Gu et al. 2021), where the Pro residue is conserved at position 22. Conversely, the T4bP, lacking the Pro residue, does not show the same structural feature (Sonani et al., 2023), despite similarities observed in HDX-MS experiments (Li et al. 2008).

The conservation of the “melted helix” in the pili of both T4aP and T2SS families strongly suggests its structural and/or functional significance. This specific segment likely plays an essential role, potentially contributing to the overall stability, assembly, or functionality of these bacterial appendages. In the case of T4aP, this structural feature is probably responsible for allowing the pili to extend under force (changing the axial rise per subunit from ~ 10 Å to ~ 30 Å) and quickly return to an equilibrium state (Egelman, 2017). Furthermore, it is associated with the flexibility of T4aP, as further described below.

The structure of T4a major pilins is stabilized by a conserved disulfide bridge at their C-terminal extremity. This disulfide bridge does not exist in T2SS major pilins. Instead, in some X-ray structures, a calcium ion was observed (Korotkov et al. 2009). In the case of *K. oxytoca* major pilin, an NMR study revealed that the pilin extracted from bacteria is in a calcium-bound state (López-Castilla et al. 2017). When calcium was removed by a chelating agent, significant structural changes and unfolding of different regions of the protein were observed, together with a drastic

decrease of the protein stability. The pilin calcium-bound state structure displays significant differences compared to the previously reported X-ray structure that had been obtained with proteins prepared in the presence of EDTA calcium chelator as antiprotease (Köhler et al. 2004). The importance of calcium for the *K. oxytoca* endopilus formation and stability was also shown by in vivo and ex vivo studies (López -Castilla et al. 2017). Remarkably, a comparative analysis of the dynamics and solvent accessibility by HDX-MS revealed the high solvent accessibility of the calcium binding site, evident not only in the isolated pilin but also in the assembled pilus (Bardiaux et al. 2019b).

For the EHEC T4aP, in contrast, calcium has a minor effect on pilus formation and stability (Karami et al. 2021). The NMR analysis of calcium effects on a monomeric pilin shows a very different picture than for T2SS. The structural effects are distributed over different areas of the protein containing conserved residues, but none of them can form a calcium binding site if the pilin is seen in isolation. However, in the context of the pilus, these areas are in contact, and at the interface of three pilin subunits form a pocket that can accommodate calcium. An extensive molecular dynamics (MD) analysis (see below) of the pilus along with mutational and in vitro and in vivo stability tests have shown the importance of calcium in T4aP pilus stability and dynamics.

Interestingly, the charge distribution at the surface of T4aP or T2SS pilus (Fig. 21.4b) shows negatively charged stretches of residues that can potentially accommodate calcium. Although the calcium binding residues both in T4P and T2SS pilins are highly conserved among species, further studies on other pilins and (endo)pili are required to confirm the overall role of calcium in pilus assembly and stability.

Dynamical Properties of T4a Pili and T2SS Endopili

As already mentioned, T4aP show remarkable flexibility and strength. Similar properties can be expected for T2SS endopili. The integrative structural biology approach described above aimed at capturing some of the dynamic properties by sampling different conformations. This was later extended with a more aggressive sampling strategy resulting in a very large ensemble of endopilus structures which were then analyzed and clustered with a machine learning approach (Nivaskumar et al. 2014). This led to the identification of several distinct structural states that are predicted to play a role in the assembly of the pili. These pili models were based on pilin subunits with continuous N-terminal helices. Since the non-helical region in the N-terminal helix may have a role in assembly, the implications of this study are likely incomplete.

In later work on structures of the same T2SS endopilus solved with cryo-EM, and thus including the non-helical region, the effects of a mutation on the dynamics were studied by normal mode analysis, together with a detailed NMR analysis of the dynamics of the pilin in isolation. In addition, Hydrogen/Deuterium eXchange

Mass Spectrometry (HDX-MS) was used to compare dynamics and solvent accessibility of isolated and assembled pilins. These complementary approaches provide a comprehensive view of internal and overall dynamics of pili, crucial for their function. The HDX-MS analysis showed that the non-helical stretch of the N-terminal helix (melted helix) was highly accessible to water, despite being buried in the interior of the pilus (Bardiaux et al. 2019b). A similar result had been obtained for a T4bP (Li et al. 2008), due to structural heterogeneity (Sonani et al., 2023).

Molecular dynamics (MD) simulations are the key method to comprehend the extraordinary strength, extensibility, and apparent flexibility exhibited by members of the TFF superfamily at ultimate detail. In one of the first MD simulations focused on a type IVa major pilin (Lemkul and Bevan 2011), the authors conducted MD simulations involving the insertion of PilA from *Pseudomonas aeruginosa* strain K into the membrane to explore the interactions between the pilin and the membrane. Additionally, they assessed the free energy required to extract PilA from the membrane by means of umbrella sampling. A molecular dynamics study of an isolated pilin PulG of the *K. oxytoca* T2SS in a model membrane underscored the role of N-methylation and the conserved Glu 5 in the reduction of polar interactions between the N-terminal domain of PulG and membrane phospholipids (Santos-Moreno et al. 2017).

To date, the sole MD study focusing on a T4aP that is based on a pilus structure with the melted helix stretch is the study of EHEC T4P (Karami et al. 2021). It investigated the conformational dynamics and stability of the pilus, employing 108 μ s MD simulations under various conditions, including different ions, protonation states, salt concentrations, and point mutations validated by experimental studies. The study identified regions influencing pilus flexibility, specifically the significant role of the melted helix in the N-terminal helix in pilus length fluctuations. This study also revealed the structural role of calcium in this T4aP, showing that the spatially separated regions interacting with calcium in isolated pilins, as demonstrated by NMR, form a pocket in the context of the assembled pilus that can stably accommodate calcium (Fig. 21.5). The study also highlighted the importance of other residues. For example, the conserved residue Glu 5 is involved in both intra-subunit interactions (with Phe 1) and inter-subunit interactions (with Thr 2 and Leu 3 of the adjacent subunits).

While insights into dynamical properties can be gained through equilibrium studies, these investigations fall short in directly addressing behavior under substantial forces. To explore the response of the gonococcal type IV pilus from *N. gonorrhoeae* to tension, steered MD simulations (SMD) were conducted (Baker et al. 2013), albeit with a model lacking the melted helix region in the N-terminal helix (Craig et al. 2006). Interestingly, these simulations showed that a 5-residue stretch near the top of the long N-terminal helix, which is in close contact with subunit $n + 4$, became exposed upon filament stretching, similar to experimental observations (Biais et al. 2010). These SMD could only visualize small elongations, making it difficult to relate directly to the experimental results. One would expect major changes in the packing of the N-terminal helix in the interior of the pilus, and additional information on the role of the melted helix stretch during extension. To reduce

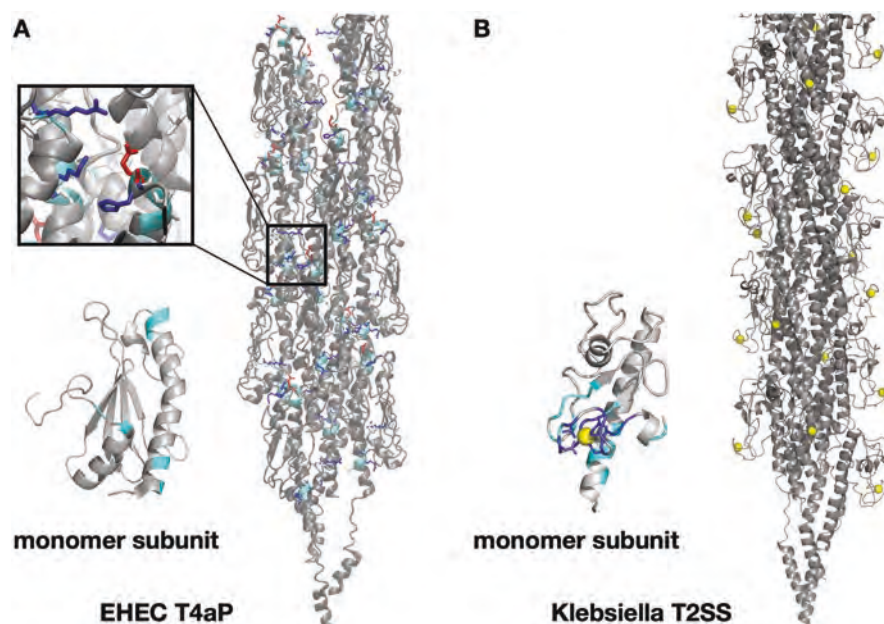


Fig. 21.5 Calcium binding sites in *K. oxytoca* T2SS and in EHEC T4aP pili. (a). EHEC T4a pilus structure (PDB: 6GV9) with a zoom in highlighting the residues of different subunits that form a calcium binding pocket. **Left:** Ribbon representation of the periplasmic domain of the monomer subunit (PDB: 6GMS) with the residues affected by the presence of calcium in cyan, based on their chemical shift perturbation analysis (Karami et al. 2021). (b) *Klebsiella oxytoca* T2SS pilus structure (PDB: 5WDA). **Left:** Ribbon representation of the calcium-bound state of the periplasmic domain of the monomer subunit (PDB: 5O2Y) with the residues affected by the presence of calcium, based on their chemical shift perturbation analysis in dark blue (highest perturbation) and cyan (lower perturbation; López-Castilla et al. 2017). Calcium ions are shown in yellow

the substantial computational cost of performing simulations with long extensions, coarse-grained simulations could be performed. One such study (Zhao et al. 2017) with the MARTINI force field (Marrink et al. 2007) was not analyzed in sufficient atomic detail and was based on the model lacking the melted helix.

Conclusion: Similarities and Differences Between T4aP and T2SS

Despite numerous investigations, our understanding of the TFF structure remains incomplete. This chapter focuses on the “bulk” regions of pili within a T4aP system and a T2SS. However, structural information from the pilus tip and its base, particularly where it inserts into the membrane, is still lacking. The ultimate goal is to obtain a comprehensive picture of entire systems. Cryo-electron tomography (cryo-ET) reconstructions can in principle offer this possibility, but in a cryo-ET study of

the T2SS from *Legionella pneumophila* (Ghosal et al. 2019), the endopilus is not visible, likely due to its dynamic nature. Similarly, a cryo-ET study on a *Myxococcus xanthus* T4aP (Chang et al. 2016) provides no information on the tip or the base of the pilus. At the base, we would expect to obtain important information on the insertion of the pilus into the membrane, the mechanism of insertion of additional subunits, and the coupling of the pilus to the ATPase that provides the energy to drive the whole system. The tip is the likely position of minor pilins (reviewed by Jacobsen et al. 2020), which may serve as a chaperone in assembly of the pilus or endopilus and are responsible for biological functions. Additionally, no structures of macromolecular complexes interacting with a pilus, such as the cargo transported by a T2SS, have been solved to date. Further structural studies are essential, particularly focusing on the (endo)pilus insertion into the membrane and its interactions with different partners, to fully comprehend the assembly and function mechanisms.

T4aP and T2SS share high structural similarity and are clearly evolutionary related. They have evolved to perform rather different functions and consequently adapted the overall structure of the filament. Thus, T2SS endopili are short and are confined to the periplasm. They do not undergo the same kind of external pulling forces as T4aP. The T2SS lacks a specific ATPase for retraction, in contrast to T4aP. Once T2SS endopili have performed their function (export a single folded protein to the extracellular region), they probably disintegrate, to be built up again to export the next protein. How this is triggered and how it happens is not known. Both T4aP and T2SS have a melted helix region in the long N-terminal helix. Given the shortness and function of the T2SS endopilus, neither the overall flexibility nor the ability to stretch and relax are likely to play an important role for an endopilus. Tight packing of the helices in the core of the pilus is a more probable explanation. This could be associated with the force needed to push a cargo protein out of the cell.

The structural significance of calcium in T2SS has been thoroughly documented and experimentally validated (López-Castilla et al. 2017). Together with the fact that T2SS lacks a retraction ATPase, this leads to the speculation that (absence of) calcium may play a role in the disintegration of the endopilus after delivery of the cargo. In the T4aP from EHEC, the presence of calcium and its impact on filament stability *ex vivo* and *in vitro* raises broader questions about the role of calcium in T4aP filaments. The EGTA-dependent dissociation of purified EHEC pili suggests a direct role of calcium in stabilizing the pilus core. Additionally, the conservation of polar residues near the calcium binding pocket raises the possibility of calcium playing a similar role in stabilizing the T4aP core structure in other species. Finally, the identified calcium binding pocket may suggest a binding site to serve as a pocket for drug design. Recent anti-virulence strategies targeting T4aP ATPases at the cytoplasmic base have emerged, leading to pilus disassembly (Ozcan et al. 2023; Denis et al. 2019; Duménil, 2019).

Acknowledgments We thank Olivera Francetic, Ed Egelman, and Benjamin Bardiaux for a long-standing collaboration on structures of the T2SS and T4aP systems. Data for the unpublished 4 Å resolution image of the T2SS endopilus was recorded at the NanoImaging Core Facility of Institut Pasteur, headed by Matthijn Vos which is partially financed by CACSICE (Equipex ANR-11-EQPX-0008).

References

- Baker JL, Biais N, Tama F (2013) Steered molecular dynamics simulations of a type IV pilus probe initial stages of a force-induced conformational transition. *PLoS Comput Biol* 9:e1003032. <https://doi.org/10.1371/journal.pcbi.1003032>. Epub 2013 Apr 11. PMID: 23592974; PMCID: PMC3623709
- Bardiaux B, de Amorim GC, Luna Rico A, Zheng W, Guilvout I, Jollivet C, Nilges M, Egelman EH, Izadi-Pruneyre N, Francetic O (2019a) Structure and assembly of the enterohemorrhagic *Escherichia coli* Type 4 Pilus. *Structure* 27:1082–1093.e5. <https://doi.org/10.1016/j.str.2019.03.021>. Epub 2019 May 2. PMID: 31056419; PMCID: PMC7003672
- Bardiaux B, Cordier F, Brier S, López-Castilla A, Izadi-Pruneyre N, Nilges M (2019b) Dynamics of a type 2 secretion system pseudopilus unraveled by complementary approaches. *J Biomol NMR* 73:293–303. <https://doi.org/10.1007/s10858-019-00246-4>. Epub 2019 May 23. PMID: 31124002; PMCID: PMC6692295
- Biais N, Higashi DL, Brucic J, So M, Sheetz MP (2010) Force-dependent polymorphism in type IV pili reveals hidden epitopes. *Proc Natl Acad Sci USA* 107:11358–11363. <https://doi.org/10.1073/pnas.0911328107>. Epub 2010 Jun 3. PMID: 20534431; PMCID: PMC2895099
- Campos M, Nilges M, Cisneros DA, Francetic O (2010) Detailed structural and assembly model of the type II secretion pilus from sparse data. *Proc Natl Acad Sci U S A* 107:13081–13086. <https://doi.org/10.1073/pnas.1001703107>. Epub 2010 Jul 2. PMID: 20616068; PMCID: PMC2919960
- Campos M, Francetic O, Nilges M (2011) Modeling pilus structures from sparse data. *J Struct Biol* 173:436–444. <https://doi.org/10.1016/j.jsb.2010.11.015>. Epub 2010 Nov 27. PMID: 21115127
- Chang YW, Rettberg LA, Treuner-Lange A, Iwasa J, Sogaard-Andersen L, Jensen GJ (2016) Architecture of the type IVa pilus machine. *Science* 355:aad2001. <https://doi.org/10.1126/science.aad2001>. Epub 2016 Mar 10. Erratum in: *Science*. 352(6282). pii: aaf7977. 10.1126/science.aaf7977. PMID: 26965631; PMCID: PMC5929464
- Craig L, Taylor RK, Pique ME, Adair BD, Arvai AS, Singh M, Lloyd SJ, Shin DS, Getzoff ED, Yeager M, Forest KT, Tainer JA (2003) Type IV pilin structure and assembly: X-ray and EM analyses of vibrio cholerae toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Mol Cell* 11:1139–1150. [https://doi.org/10.1016/s1097-2765\(03\)00170-9](https://doi.org/10.1016/s1097-2765(03)00170-9). PMID: 12769840
- Craig L, Volkman N, Arvai AS, Pique ME, Yeager M, Egelman EH, Tainer JA (2006) Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol Cell* 23:651–662. <https://doi.org/10.1016/j.molcel.2006.07.004>. PMID: 16949362
- Craig L, Forest KT, Maier B (2019) Type IV pili: dynamics, biophysics and functional consequences. *Nat Rev Microbiol* 17:429–440. <https://doi.org/10.1038/s41579-019-0195-4>. PMID: 30988511
- Denis K, Le Bris M, Le Guennec L, Barnier JP, Faure C, Gouge A, Bouzinba-Ségard H, Jamet A, Euphrasie D, Durel B, Barois N, Pelissier P, Morand PC, Coureuil M, Lafont F, Join-Lambert O, Nassif X, Bourdoulous S (2019) Targeting type IV pili as an antivirulence strategy against invasive meningococcal disease. *Nat Microbiol* 4:972–984. <https://doi.org/10.1038/s41564-019-0395-8>. Epub 2019 Mar 25. PMID: 30911127
- Denise R, Abby SS, Rocha EPC (2019) Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biol* 17:e3000390. <https://doi.org/10.1371/journal.pbio.3000390>. PMID: 31323028; PMCID: PMC6668835
- Duménil G (2019) Type IV pili as a therapeutic target. *Trends Microbiol* 27:658–661. <https://doi.org/10.1016/j.tim.2019.05.005>. Epub 2019 Jun 7. PMID: 31182345
- Egelman EH (2000) A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* 85:225–234. [https://doi.org/10.1016/s0304-3991\(00\)00062-0](https://doi.org/10.1016/s0304-3991(00)00062-0). PMID: 11125866

- Egelman EH (2017) Cryo-EM of bacterial pili and archaeal flagellar filaments. *Curr Opin Struct Biol* 46:31–37. <https://doi.org/10.1016/j.sbi.2017.05.012>. Epub 2017 Jun 10. PMID: 28609682; PMCID: PMC5683938
- Folkhard W, Marvin DA, Watts TH, Paranchych W (1981) Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. *J Mol Biol* 149:79–93. [https://doi.org/10.1016/0022-2836\(81\)90261-8](https://doi.org/10.1016/0022-2836(81)90261-8). PMID: 6118441
- Ghosal D, Kim KW, Zheng H, Kaplan M, Truchan HK, Lopez AE, McIntire IE, Vogel JP, Cianciotto NP, Jensen GJ (2019) In vivo structure of the legionella type II secretion system by electron cryotomography. *Nat Microbiol* 4:2101–2108. <https://doi.org/10.1038/s41564-019-0603-6>. Epub 2019 Nov 21. Erratum in: *Nat Microbiol* 2020 Apr;5(4):651. PMID: 31754273; PMCID: PMC6879910
- Gu Y, Srikanth V, Salazar-Morales AI, Jain R, O'Brien JP, Yi SM, Soni RK, Samatey FA, Yalcin SE, Malvankar NS (2021) Structure of *Geobacter* pili reveals secretory rather than nanowire behaviour. *Nature* 597:430–434. <https://doi.org/10.1038/s41586-021-03857-w>. Epub 2021 Sep 1. PMID: 34471289; PMCID: PMC9127704
- Jacobsen T, Bardiaux B, Francetic O, Izadi-Pruneyre N, Nilges M (2020) Structure and function of minor pilins of type IV pili. *Med Microbiol Immunol* 209:301–308. <https://doi.org/10.1007/s00430-019-00642-5>. Epub 2019 Nov 29. PMID: 31784891; PMCID: PMC7248040
- Kang HJ, Baker EN (2012) Structure and assembly of gram-positive bacterial pili: unique covalent polymers. *Curr Opin Struct Biol* 22:200–207. <https://doi.org/10.1016/j.sbi.2012.01.009>. Epub 2012 Feb 16. PMID: 22342454
- Karami Y, López-Castilla A, Ori A, Thomassin JL, Bardiaux B, Malliavin T, Izadi-Pruneyre N, Francetic O, Nilges M (2021) Computational and biochemical analysis of type IV pilus dynamics and stability. *Structure* 29:1397–1409.e6. <https://doi.org/10.1016/j.str.2021.07.008>. Epub 2021 Sep 13. PMID: 34520738
- Köhler R, Schäfer K, Müller S, Vignon G, Diederichs K, Philippsen A, Ringler P, Pugsley AP, Engel A, Welte W (2004) Structure and assembly of the pseudopilin PilG. *Mol Microbiol* 54:647–664. <https://doi.org/10.1111/j.1365-2958.2004.04307.x>. PMID: 15491357
- Kolappan S, Coureuil M, Yu X, Nassif X, Egelman EH, Craig L (2016) Structure of the *Neisseria meningitidis* type IV pilus. *Nat Commun* 7:13015. <https://doi.org/10.1038/ncomms13015>. PMID: 27698424; PMCID: PMC5059446
- Korotkov KV, Gray MD, Kreger A, Turley S, Sandkvist M, Hol WG (2009) Calcium is essential for the major pseudopilin in the type 2 secretion system. *J Biol Chem* 284:25466–25470. <https://doi.org/10.1074/jbc.C109.037655>. Epub 2009 Jul 29. PMID: 19640838; PMCID: PMC2757947
- Kühlbrandt W (2014) The resolution revolution. *Science* 343:1443–1444. <https://doi.org/10.1126/science.1251652>. PMID: 24675944
- Lemkul JA, Bevan DR (2011) Characterization of interactions between PilA from *Pseudomonas aeruginosa* strain K and a model membrane. *J Phys Chem B* 115:8004–8008. <https://doi.org/10.1021/jp202217f>. Epub 2011 Jun 1. PMID: 21630674
- Li J, Lim MS, Li S, Brock M, Pique ME, Woods VL Jr, Craig L (2008) *Vibrio cholerae* toxin-coregulated pilus structure analyzed by hydrogen/deuterium exchange mass spectrometry. *Structure* 16:137–148
- López-Castilla A, Thomassin JL, Bardiaux B, Zheng W, Nivaskumar M, Yu X, Nilges M, Egelman EH, Izadi-Pruneyre N, Francetic O (2017) Structure of the calcium- dependent type 2 secretion pseudopilus. *Nat Microbiol* 2:1686–1695. <https://doi.org/10.1038/s41564-017-0041-2>. Epub 2017 Oct 9. PMID: 28993624; PMCID: PMC5705324
- Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* 111(27):7812–24. <https://doi.org/10.1021/jp071097f>
- Neuhaus A, Selvaraj M, Salzer R, Langer JD, Kruse K, Kirchner L, Sanders K, Daum B, Averhoff B, Gold VAM (2020) Cryo-electron microscopy reveals two distinct type IV pili assembled

- by the same bacterium. *Nat Commun* 11:2231. <https://doi.org/10.1038/s41467-020-15650-w>. PMID: 32376942; PMCID: PMC7203116
- Nivaskumar M, Bouvier G, Campos M, Nadeau N, Yu X, Egelman EH, Nilges M, Francetic O (2014) Distinct docking and stabilization steps of the Pseudopilus conformational transition path suggest rotational assembly of type IV pilus-like fibers. *Structure* 22:685–696. <https://doi.org/10.1016/j.str.2014.03.001>. Epub 2014 Mar 27. PMID: 24685147; PMCID: PMC4016124
- Ozcan A, Keskin O, Sariyar Akbulut B, Ozbek P (2023) Piperidine-based natural products targeting type IV pili antivirulence: a computational approach. *J Mol Graph Model* 119:108382. <https://doi.org/10.1016/j.jmgm.2022.108382>. Epub 2022 Nov 26. PMID: 36463631
- Parge HE, Forest KT, Hickey MJ, Christensen DA, Getzoff ED, Tainer JA (1995) Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature* 378:32–38. <https://doi.org/10.1038/378032a0>. PMID: 7477282
- Reardon PN, Mueller KT (2013) Structure of the type IVa major pilin from the electrically conductive bacterial nanowires of *Geobacter sulfurreducens*. *J Biol Chem* 288:29260–29266. <https://doi.org/10.1074/jbc.M113.498527>. Epub 2013 Aug 21. PMID: 23965997; PMCID: PMC3795228
- Santos-Moreno J, East A, Guilvout I, Nadeau N, Bond PJ, Tran Van Nhieu G, Francetic O (2017) Polar N-terminal residues conserved in type 2 secretion Pseudopilins determine subunit targeting and membrane extraction steps during fibre assembly. *J Mol Biol* 429:1746–1765. <https://doi.org/10.1016/j.jmb.2017.04.005>. Epub 2017 Apr 17. PMID: 28427876
- Sonani RR, Sanchez JC, Baumgardt JK, Kundra S, Wright ER, Craig L, Egelman EH (2023) Tad and toxin-coregulated pilus structures reveal unexpected diversity in bacterial type IV pili. *Proc Natl Acad Sci USA* 120:e2316668120. <https://doi.org/10.1073/pnas.2316668120>. Epub 2023 Nov 27. PMID: 38011558
- Wang F, Coureuil M, Osinski T, Orlova A, Altindal T, Gesbert G, Nassif X, Egelman EH, Craig L (2017) Cryoelectron microscopy reconstructions of the *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* type IV pili at sub-nanometer resolution. *Structure* 25:1423–1435.e4. <https://doi.org/10.1016/j.str.2017.07.016>. PMID: 28877506; PMCID: PMC8189185
- Zhao X, Liu Y, Guo Z, Zhang Y, Li Y, Liu W (2017) Mechanical response and deformation mechanics of type IV pili investigated using steered coarse-grained molecular dynamics simulation. *J Biomech* 56:97–101. <https://doi.org/10.1016/j.jbiomech.2017.03.013>. Epub 2017 Mar 18. PMID: 28365063