MATTHEWS, HOLLY K., B.S., May 2023 PHYSICS

DNA SUPPORT STRUCTURES FOR MEMBRANE PROTEIN IMAGING (48 pages)

Thesis Advisor: Thorsten-Lars Schmidt

Researchers believe that protein misfolding could be involved in up to half of all human diseases. Experimentally determining the structure of one especially important class of proteins, called membrane proteins, is essential to the study and treatment of such diseases. A promising approach is the isolation and imaging of these membrane proteins in cryo-electron microscopy which has become one of the most important methods to solve protein structures. The fragile membrane protein can be stabilized in an artificial lipid-bilayer nanodisc. To better capture, stabilize, and orient these nanodiscs, a nanoscale DNA origami support structure was designed. DNA origami is a method of folding a long single-stranded DNA, called a scaffold strand, into engineered shapes using short DNA oligonucleotides called staples. Because of the properties of DNA —including the predictability of Watson-Crick base pairs, the highly studied molecular configuration of DNA, and the self-assembly of DNA into double-stranded helixes—, scaffold strands and staples can be engineered to combine and form very precise and customizable nanoscale structures. Here, we theoretically design, simulate, and experimentally test a DNA origami support structure for membrane protein imaging. The results show favorable stability of the design in simulations and promising self-assembly of the structures during experimental testing. Further studies will continue to develop and optimize the support structure as well as the attachment of nanodiscs. The ability to build a nanoscale support structure to capture these nanodiscs is relevant to the development of an imaging method that could allow 3D reconstruction of essential membrane protein structures.

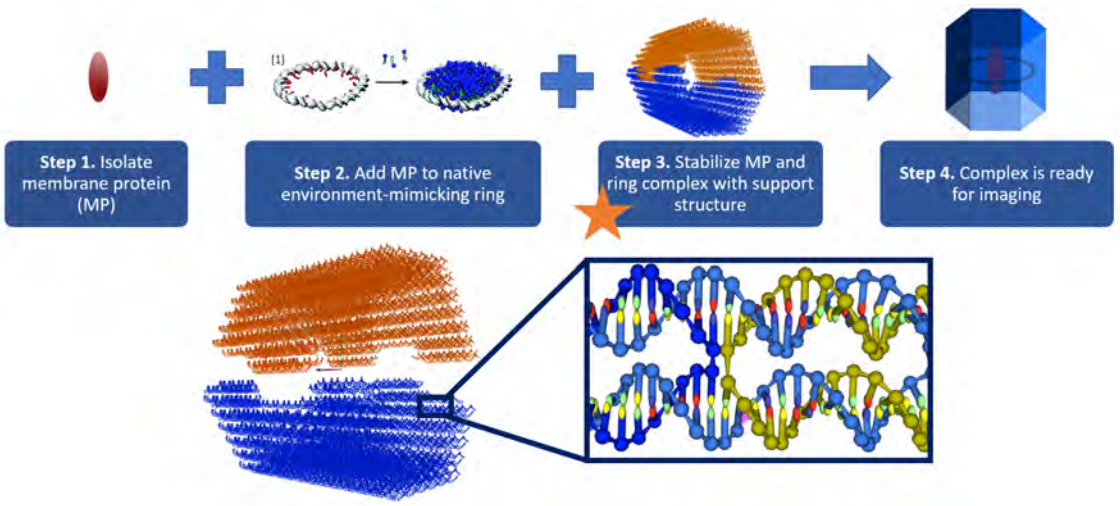

Figure 1: Image depicting the process of membrane protein structure determination. First the protein of interest is extracted from its native environment and placed in a nanodisc containing a native-like lipid bilayer. Then the nanodisc is captured by a DNA support structure for cryoelectron imaging. The focus of this paper is the DNA support structure.

DNA SUPPORT STRUCTURES FOR MEMBRANE PROTEIN IMAGING

A thesis submitted to the
Kent State University Honors College
in partial fulfillment of the requirements
for Departmental Honors

by
Holly K. Matthews
May 2023

Thesis written by

Holly K. Matthews

Approved by

Thorsten-Lars Schmidt, Advisor

Michael Strickland, Chair, Department of Physics

Accepted by

Alison Smith, Dean, Honors College

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Many people have helped me along my undergraduate journey and it would be impossible to thank all of them. Chronologically there are a few exceptionally important people I would like to thank. These are the people, without which, not only would this thesis would not be possible, but also my acceptance and enrollment in the high energy PhD program at Michigan State University. The professors I have taken class with along my journey have also played a huge role in my experience. Thank you especially to Hamza Balci, Mina Katramatou, and Brett Ellman for your advice, amazing classes, and being an inspiration. I've truly learned a lot.

The first people who have helped me on my journey are, of course, my parents. I'd like to thank them for always supporting me and showing me the importance of hard work. Even if they stopped being able to help me with my math homework long ago, I always could count on them to help with anything else. Thank you for always picking up the phone to help me with questions about laundry and to listen about "those boson hadron things." I'd also like to thank them for encouraging me to keep music as a part of my life.

Next I'd like to thank Hend Baza and Professor Oleg Lavrentovich for introducing me to my first real research experience. I am honored to have had the opportunity to work in your lab, learn about data collection, how to read papers, how to present a talk, and how to write an abstract. This experience was invaluable, and I attribute it directly to my opportunity to attend graduate school. Thank you especially to Hend for being a kind, patient, and inspiring mentor to such a shy and nervous intern. I'd like to also thank my advisor Professor Thorsten Schmidt and PhD candidate Praneetha Sundar Prakash for welcoming me into their lab and the world of biophysics to do my thesis. I learned more about pipetting small volumes than I thought possible, and in all seriousness gained really valuable laboratory benchwork skills. I'd like to thank Praneetha for showing me the ropes in the lab, for tolerating my incessant questions, and for taking the time to read and edit this thesis. I appreciate it more than I can express.

In that same vein I'd like to thank my committee members, Hamza Balci, Rajeev Rajaram, and Allison Smith for taking the time out of their busy schedules to read this thesis and participate in

the defense. This is the largest milestone in my career as an undergraduate and it is educators like you that make this possible.

Last but not least I'd like to thank the fellow students who walked this path beside me. Thank you for your company in classes and labs, for the late nights finishing problem sets, and for our bubble tea breaks. Brian especially, who, if you tally all the lab reports and this thesis, kindly and patiently proofread more than 150 pages of my reports. I wish all of you the absolute best. I have enjoyed my time here at Kent State University and I owe it to the great professors, wonderful mentors, and amazing friends I have met here. Thank you!

CHAPTER 1

INTRODUCTION

Molecular self-assembly is a process where molecules adopt particular arrangements without guidance from outside sources. This process occurs due to weak interactions such as hydrogen bonding, hydrophobic interactions, and van der Waals forces (Hong, Zhang, Liu, & Yan, 2017). It governs the formation of lipid molecules into micelles, the formation of DNA double helices, and the tertial structures of RNA (Hong et al., 2017). These self-assembly processes are not only essential for biological function but are also essential to the design of synthetic nanostructures and devices.

1.1 The DNA Molecule

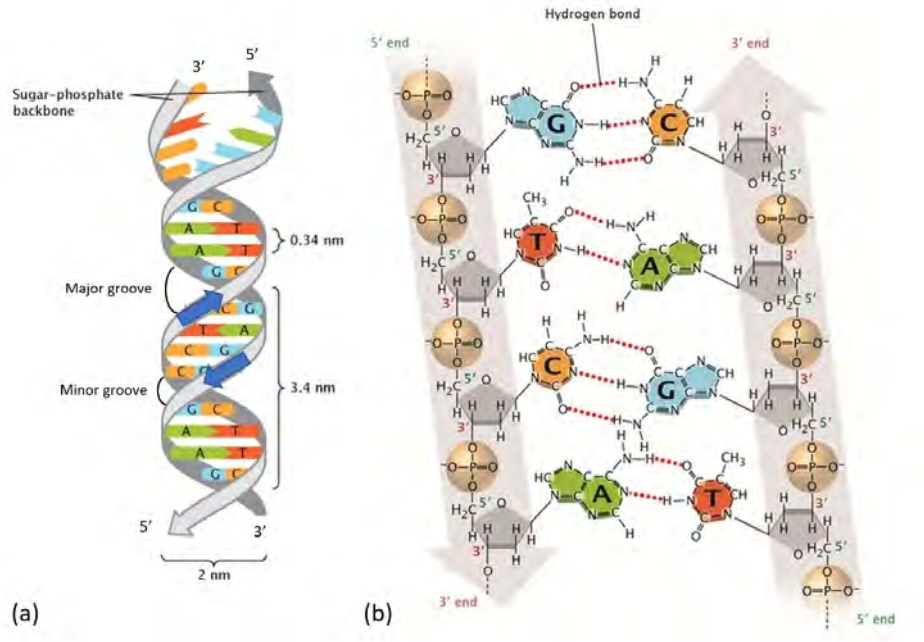

Figure 2: Schematic of the DNA molecule showing the B-form geometry, sugar-phosphate backbone, and base pairing. (a) shows a simple view with the sugar-phosphate backbone represented as a grey ribbon. (b) the right side shows a molecular view of the DNA with hydrogen bonding marked between base pairs. Taken from (Pray, 2008).

The DNA molecule, discovered in conjunction by Friedrich Miescher, Phoebus Levene, Erwin Chargaff, Rosalind Franklin, Maurice Wilkins, as well as James Watson and Francis Crick, is essential to life (Pray, 2008). The molecule constitutes the genetic material that is inside nearly all organisms on Earth (Marenduzzo, 2018). It is a long double-stranded polymer that forms the famous double

helix shape. A sequence of nucleotides runs along the DNA strand. Each nucleotide contains three parts, the sugar (from which DNA gets its name) deoxyribose, a phosphate group, and a base. The sugar and phosphate groups alternate to form the backbone of the DNA strand, while the base which is attached to the sugar extends toward the middle of the structure.

The four bases of DNA, namely, cytosine (C), guanine (G), adenine (A), and thymine (T) exhibit inherent complementarity, that is, C always forms triple hydrogen bonds with G whereas A forms double hydrogen bonds with T. This helps us predict the complementary strand when a sequence of a DNA strand is given. The direction in which the sequence is read, whether up or down, is determined by the orientation of the sugar group. As shown in Figure 2, the 3' (read "three prime") end is indicated by an arrow and represents a strand that ends at the third carbon in the sugar ring (Marenduzzo, 2018). The other end, called the 5' end, represents the side of a strand that ends at the fifth carbon that extends outside the sugar ring. By convention, DNA sequences are always written in the 5' to 3' direction. For example, the sequence of the left stand in Figure 2 (b) is written as GTCA. The bases are made up of aromatic rings that have pi bonds. When stacked, as in the case of DNA structure, these pi bonds overlap and attract each other. The most favorable way this stacking occurs creates the major and minor grooves of DNA (see 2 a).

The geometry of the DNA structure is well-studied. The most common form of DNA that exists in nature is called B-form. B-form DNA is a right-handed double helix with a thickness of two nanometers (nm) and a rise of 0.34 nm per base-pair (Marenduzzo, 2018). The A-form named by Rosalind Franklin represents the geometry of the DNA after being dehydrated and Z-form has a left-handed rather than right-handed twist (Marenduzzo, 2018).

1.2 Structural DNA Nanotechnology

1.2.1 Holliday Junction Crossovers

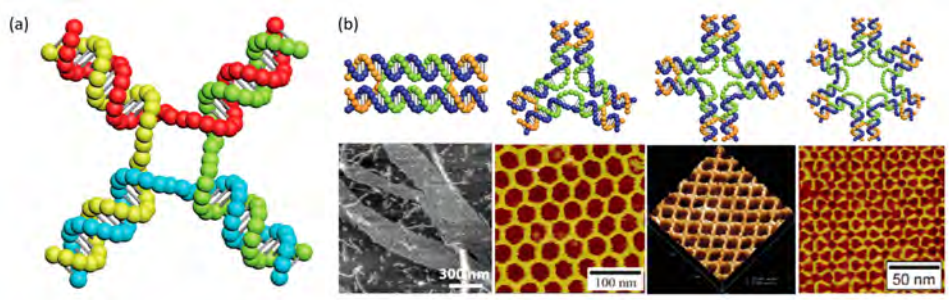

Figure 3: (a) a schematic of a DNA Holiday junction. (b) Schematics of 2,3,4, and 6-arm junctions with resulting lattice assemblies imaged by AFM. Modified from (Hong et al., 2017).

The pioneer of structural DNA nanotechnology, Ned Seeman and his team, made the first breakthrough in synthetic DNA nanostructures by creating the immobile Holliday junction (Hong et al., 2017). A Holliday junction is a connection between two double helix DNA strands that occurs naturally during replication or recombination, see Figure 3 (a). However, as it occurs in nature this junction is unstable and temporary so that the branching point can move positions during DNA replication (Kallenbach, Ma, & Seeman, 1983). In their 1983 paper Kallenbach, Ma, and Seeman proposed a way to break the symmetry of the natural junction with commercially synthesized DNA sequences and prevent the strands from migrating back to a double helix form. This allowed the creation of stable and programmable junctions. Developing 2,3,4, and 6-arm junctions further opened nanostructure design opportunities and enabled them to create different 2D design lattices (Figure 3). Because these junctions were still very flexible, Seeman and team also developed a DNA double crossover design by joining two 4-way junctions. This crossover was rigid and stable allowing the development of 2D DNA arrays with different patterns, setting the groundwork for future DNA nanotechnological advancements (Hong et al., 2017).

1.2.2 2-Dimensional DNA Origami

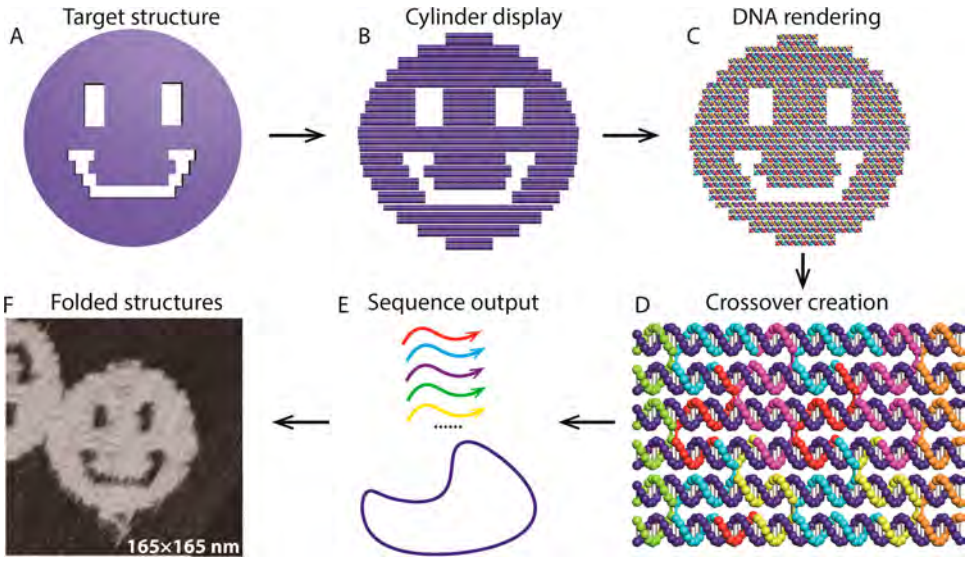

Figure 4: A workflow of DNA origami design and creation. The target structure is first determined, then broken down into raster-style cylinders, then DNA scaffold routing is designed, followed by additions of staple crossovers, and finally, the staples and scaffold strand are be mixed together to self-assemble into the desired structure. From (Hong et al., 2017).

The next milestone in DNA nanotechnology was in 2006 when Paul Rothemund published the DNA origami technique. Rothemund described a method of folding a long, single-strand of DNA called a scaffold, into complex two-dimensional shapes by pinning it with hundreds of short oligonucleotides called staple strands (Figure 4 E) (Rothemund, 2006). To design these nanostructures, a target shape is first determined (Figure 4 a). Then, this shape is broken down into cylinders that represent DNA double helices (Figure 4 b). Next, the DNA sequences and crossover locations are designed (Figure 4 c and d). The scaffold strand is a 7,000 – 8,000 base-pair long single-stranded DNA derived from bacteriophage genomes, typically the M13mp18 type, and has a known sequence (Hong et al., 2017). From this sequence, small stretches (15 – 60 base pairs long) of complementary sequences can be designed to bend the scaffold into the desired shape (Figure 4 e). The scaffold strand and the hundreds of staples are mixed in a salt solution and then heated to 90 °C and then slowly cooled to room temperature. The temperature ramp allows the short staples to find and bind to their specified place on the scaffold (Hong et al., 2017). It also ensures that if a strand binds to an incorrect but similar location, there is enough time and motion for the stand to unbind and find its final destination. Once all the staples are attached, the scaffold will be held in its final

structure, in this case, a smiley face, due to complementary base pairing and pi-pi stacking (Figure 4 (f)). This method allows the design flexibility to engineer a nearly endless number of structures from smiley faces to a map of the world (Rothemund, 2006).

1.2.3 Three-Dimensional DNA Origami

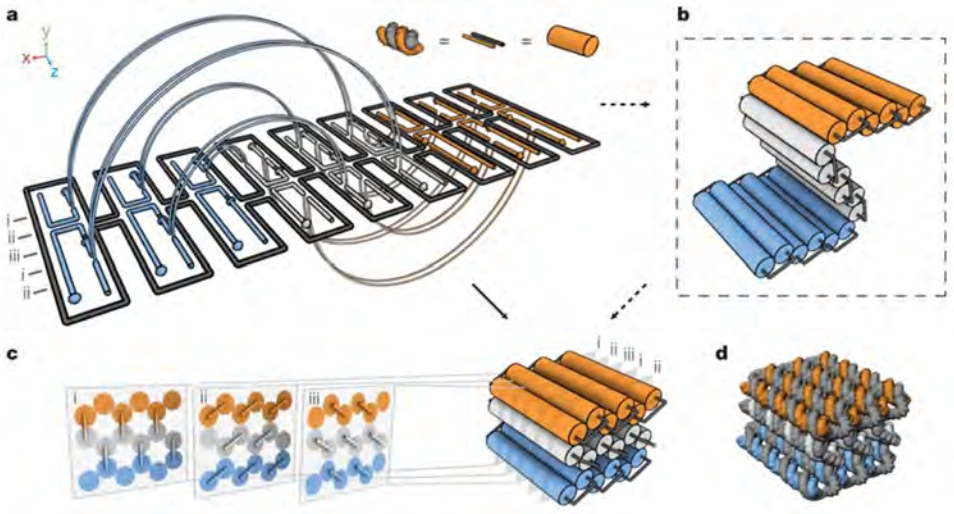

Figure 5: Design of three-dimensional DNA origami. (a) The grey scaffold strand is populated with orange, white, and blue staple strands whose crossovers are shown as arcs. (b) schematic depicting the double helices as cylinders showing the folding process. (c) Model of the final shape with cross-sections showing crossovers. (d) atomic model of the final structure. From (Douglas, Dietz, et al., 2009).

After the development of single-layered two-dimensional designs, the next step in DNA nanotechnology was the development of multi-layered three-dimensional designs. The basic idea was to stack Rothemund's single-layered origami to create rigid 3D structures. In their 2009 paper, Douglas et al. described how custom shapes could be formed as pleated layers constrained to a honeycomb lattice. By using careful placement of the crossover positions with respect to the geometry of DNA, sheets of antiparallel helices can be pulled together to form layers (as shown in Figure 5). This is done by restricting the crossovers to positions where helices are closest together in the honeycomb lattice.

In their study, Douglas et al. found several parameters that optimize 3D origami folding. They suggest a local crossover density of roughly one per 21 base pairs, and staples lengths are between 18 and 49 bases, with a mean between 30 and 42 (Douglas, Dietz, et al., 2009). In addition to design parameters, folding parameters were also experimentally tested in the study. Longer lengths

of thermal ramps and a combination of magnesium chloride ($MgCl_2$) and sodium chloride (NaCl) salt promoted the best folding. The longer lengths of temperature ramps are required because of the increased size and increased kinetic traps during the folding of 3D origami compared to 2D (Douglas, Dietz, et al., 2009). The divalent cation of Mg^{2+} was found to accelerate folding and increase undesired aggregation, whereas the monovalent cation of Na^{+} was found to decrease the rate of folding and decrease undesired aggregation (Douglas, Dietz, et al., 2009). This is because divalent cations can interact with the major groove of DNA and specifically stabilize Holliday-junction crossovers, but also stabilize aggregates through the same interaction. Divalent cations can also stabilize DNA by forming a "bridge" between the minor grooves of two different strands (Luan & Aksimentiev, 2008). Monovalent cations compete with the divalent cations' interaction, and a specific combination of the two optimizes good folding and less aggregation.

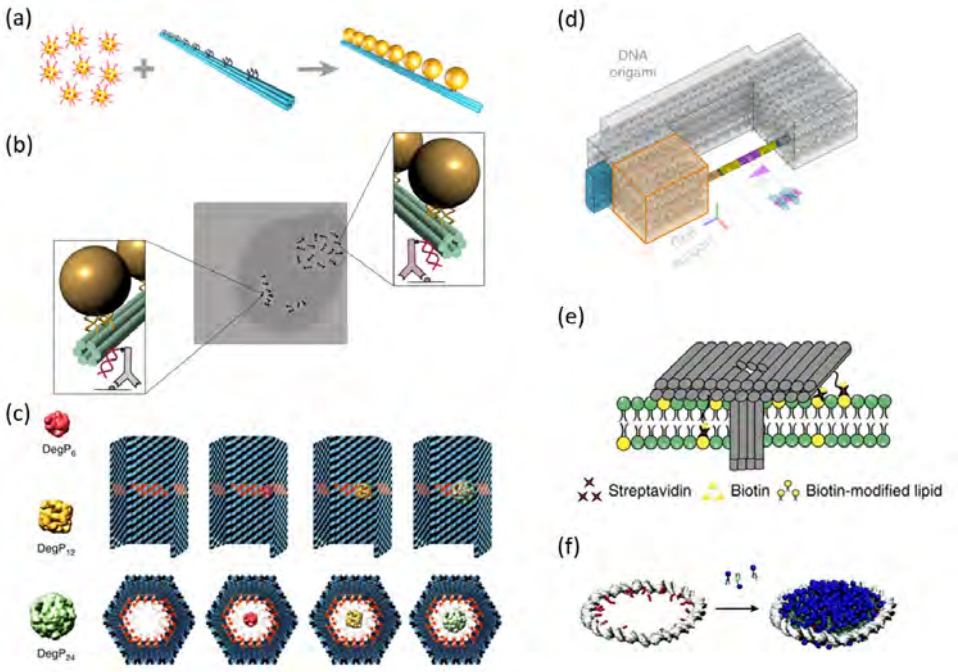

Figure 6: (a) A DNA origami six-helix bundle used to position gold nanoparticles for plasmonic devices from (Gür et al., 2016). (b) six-helix bundles with gold nanoparticles and antibodies for multiplexed immunostaining from (Joshi, 2017). (c) DNA nanostructure for protein encapsulation from (Sprengel et al., 2017). (d) A DNA goniometer from (Aksel et al., 2021). (e) A DNA nanopore from (Krishnan et al., 2016). (f) A DNA encircled lipid-bilayer from (Iric et al., 2018).

The ability to design 3D nanostructures opens an extremely wide variety of applications in physics, biology, and engineering. Figure 6 shows six such applications. The DNA six-helix bundles

origamis can be hybridized to gold nanoparticles (Au NPs) at specific locations through single-stranded extensions attached to origami bundles. The gold nanoparticles are densely covered with the complementary oligonucleotide for optimal hybridization. These Au NP decorated origamis can be used as plasmonic devices for small-scale photonic and electronic applications and also as an immunolabelling method (Gür et al., 2016). For immunolabelling, the origamis were modified to include a single-stranded oligonucleotide extension to bind to the complementary strand attached to the antibody of interest (Joshi, 2017).

Figure 6 (c) shows a DNA structure built to encapsulate proteins. Because origami can be used to make nanoscale objects, DNA containers could be designed to synthetically trap specific proteins in the same way cells metabolize molecules (Sprengel et al., 2017). Figure 6 (d) shows the design of a molecular goniometer (an instrument that precisely orients objects). Because the twist of DNA is so accurately known it is possible to use origami to determine the orientation of objects (Aksel et al., 2021). Figure 6 (e) shows a DNA-based nanopore. Nanopores sit on the edge of cell membranes and have applications in single-molecule biosensing (Krishnan et al., 2016). Because of their small size origamis can be used to program these synthetic devices. Lastly, Figure 6 (f) shows the use of a double-stranded DNA ring to encapsulate a lipid bilayer. This lipid-bilayer mimics the environment of a cell membrane and has applications to membrane protein structure determination, described in more detail in the next section (Iric et al., 2018).

In summary, DNA origami is a method of folding a long single-stranded DNA, called a scaffold strand, into engineered shapes using short DNA oligonucleotides called staples. Because of the properties of DNA —including the predictability of Watson-Crick base pairs, the highly studied molecular configuration of DNA, and the self-assembly of DNA into double-stranded helixes—, scaffold strands and staples can be engineered to combine and form very precise and customizable nanoscale structures. These customizable structures have already been applied to many applications from optics to immunostaining. The method continues as a growing field with more and more applications and DNA structures being developed.

1.3 Protein Structure Determination

One especially promising application of DNA origami is in the structure determination of membrane proteins. Membrane proteins are essential for human life and more than 50% of all drugs

specifically target membrane proteins (Overington, Al-Lazikani, & Hopkins, 2006). Determining the 3D structure of membrane proteins remains a difficult scientific endeavor and it is essential to improve structure determination methods for the development of drugs and the treatment of disease.

Membrane proteins are proteins that are part of, or interact with, cell membranes. These proteins perform many important functions for the cell such as transporting ions, proteins, and RNA into and out of the cell, as well as sending and receiving chemical and electrical signals. Two types of membrane proteins are integral membrane proteins which sit inside the cell membrane and peripheral membrane proteins which attach loosely to the membrane through van der Waals or covalent interactions (Hedin, Illergård, & Elofsson, 2011). Unlike water-soluble proteins which are found in a hydrophilic environment, membrane proteins are found within the hydrophobic lipid bilayer of cell membranes. This difference in the environment greatly affects the structure of the proteins. Thus, determining the structure of these uniquely important membrane proteins also comes with unique challenges.

The first challenge is the small size of most membrane proteins. With sizes on the order of a few nanometers, imaging these proteins requires specialized techniques such as X-ray crystallography and electron microscopy. Another challenge is the expression, purification, and handling of membrane proteins. Conventional X-ray techniques require the proteins to be prepared for imaging by crystallizing milligram amounts of them (Nygaard, Kim, & Mancia, 2020). Because of the low expression yield (the production of proteins by cells) of membrane proteins, it is difficult to generate enough of the target protein to successfully form a crystal (Kermani, 2021). In addition, extracting the membrane proteins that were successfully expressed is also more difficult than typical soluble proteins because of the differences in native environments. Since membrane proteins reside in the hydrophobic cell membranes, they require detergents (soaps) to be extracted. These detergents can damage the structural integrity of fragile membrane proteins. Then, even if the first two expression and extraction steps were successful, it is difficult to find the correct conditions that lead to crystallization, requiring painstaking experimental trials (Kermani, 2021). To combat these challenges there is a call for a new imaging method that has a nanoscale resolution, decreased sample requirements, and the ability to stabilize proteins in a native-like environment.

One promising imaging method is cryo-electron microscopy (cryo-EM). In cryo-EM solutions

of purified proteins are applied to a carbon film and then plunged in liquid ethane (Martin et al., 2016). This creates a thin layer of ice with the protein frozen inside. This frozen sample is then imaged under an electron microscope. The images will contain 2D projections of thousands of identical proteins which can then be fed into 3D reconstruction algorithms to map the structure of the protein (Martin et al., 2016). This technique is particularly promising for membrane protein structure determination because it reduces the sample requirements by three orders of magnitude compared to crystallography objects (Nygaard et al., 2020).

However, cryo-EM presents its own challenges like the high cost of the instrument itself and difficulty in achieving a resolution to detect membrane proteins smaller than 100 kDa. This is because contrast and signal-to-noise ratios are especially low for these nanometer-sized objects (Nygaard et al., 2020). Difficulties in sample preparation also result in nonuniform ice thickness and particle distribution, which harms image resolution and the accuracy of the reconstruction algorithms (Martin et al., 2016). In addition, problems regarding the stability of the membrane proteins outside of their native environment still need to be addressed. Sample preparation from cryo-EM subjects the target proteins to forces during the blotting and freezing which have been shown to denature these fragile complexes (Martin et al., 2016). Forces at the air-water interface during freezing are especially important to membrane proteins because their hydrophobic regions cause them to arrange on the surface. This subjects the membrane proteins to more forces during freezing and may also orient the particles in specific directions. Orientation of the particles all in the same direction is not favorable to reconstructing the 3D structure because entire sides of the protein will not be imaged.

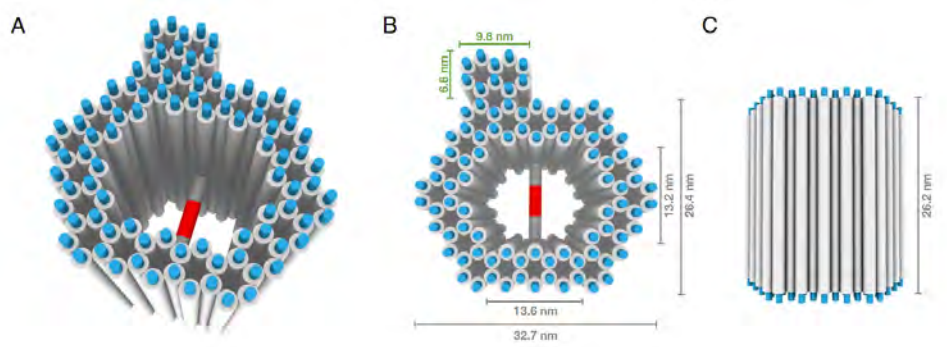

Figure 7: (A) Schematic of a DNA support structure for cryo-electron imaging of a transcription protein from (Martin et al., 2016). (B) A Top-down view and (C) a side-view with the DNA helices represented as cylinders and the transcription protein is represented in red. The purpose of the structure is to protect the protein from damaging forces during freezing as well as orient the protein for 3D reconstruction of its structure.

To address the issues, Martin et al. developed a DNA origami nanocage for capturing DNA-binding proteins. The DNA nanostructure was shaped like a hexagonal prism with an open top and bottom face. A pillar of single-stranded DNA inside of the hexagon was designed to bind to the transcription protein, pictured in red in Figure 7. These nanocages include hydrophobic single-stranded DNA overhangs at the top and bottom faces of the structure that interact with the air-water interface during freezing and orient the structure open-side up. Martin et al. found that the ice layer of the samples prepared with the support structure was more uniform in thickness allowing better resolution. The structure also surrounds the protein, keeping it away from the air-water interface which protects the protein from blotting forces or damage during freezing (Martin et al., 2016). This method is limited, however, in the types of proteins that can be studied. Only a small group of proteins that can bind to DNA can be imaged in this support structure. To extend this technique to membrane proteins another novel device is required.

Membrane proteins, because of their unique hydrophobic environment, are notoriously difficult to study. However, if it is possible to create a synthetic environment, like a small patch of cell membrane, it can serve to house and stabilize the membrane protein and even be placed inside a protective DNA nanocage. This synthetic membrane protein environment can be achieved with nanodiscs. Currently, membrane scaffolding proteins (MSPs), or styrene-maleic acid lipid particles (SMALPs) are used to encircle a patch of lipid-bilayer for these in vitro studies but are limited in the range of sizes that can be produced (Chen, Majdinasab, Fiori, Liang, & Altenberg, 2020).

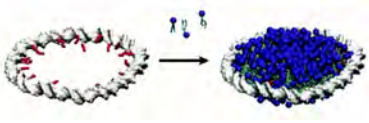

Figure 8: Diagram of a single-stranded ring of DNA being filled with phospholipids forming a lipid bilayer. This lipid bilayer mimics the environment of a cell membrane and is designed to house membrane proteins. From (Iric et al., 2018).

Iric et al. describe the development of DNA-encircled lipid bilayers. Lipids are the molecules that make up cell membranes where membrane proteins reside. By modifying a strand of DNA to form a small ring these lipid bilayers can be captured as shown in Figure 8. These small disks of cell membrane-like structure can house membrane proteins in a native-like environment (Iric et al., 2018). Combining the techniques of a DNA origami support cage, DNA-encircled lipid bilayers, and cryo-EM, current problems in membrane protein reconstruction can be addressed.

1.4 Computer Simulation

To aid in the design of DNA origami structures several models and computer simulation code have been developed to predict the characteristics of structures before experimental testing. Being able to predict the stability, flexibility, and thermodynamic motions of a structure during the design phase allows for design changes without costly and time-consuming experiments. Modeling of DNA can be performed at a variety of levels of resolution, with more detail costing more computation time. During the DNA origami design process, coarser models are used for quick checks and fine models are used at later stages due to this trade-off in computation time.

1.4.1 CanDo

CanDo is a very quick but coarse-grained model of the mechanical properties of DNA nanostructures. Each double-stranded helix of DNA is represented as an elastic rod that can stretch, twist, and bend. It was originally developed by Prof. Do-Nyum as a postdoctoral research associate at MIT (Kim, Kilchherr, Dietz, & Bathe, 2012), experimentally calibrated and verified by Castro et al. (Castro et al., 2011) and updated to simulate Brownian Dynamics by Dr. Keyao Pan (Sedeh et al., 2016). The main purpose of the CanDo model is to predict the 3D solution shape and flexibility of DNA-based nanostructures. This is done by modeling DNA as elastic rods with the same geometry (base-pair rise of 0.34 nm, diameter of 2.25 nm) and properties (1.100 pN stretch modulus, 230

pN nm^2 bend modulus, and 460 pN nm^2 twist modulus). Crossover connections between helices are modeled as springs. The software uses these values to calculate root-mean-square fluctuations (RMSF) of the deformed 3D structure in thermal equilibrium using standard statistical mechanics equations and the worm-like chain model of polymers (Castro et al., 2011). The model ignores sequence-dependent properties, interhelical electrostatic repulsion, and the major-minor groove of DNA. An example of a CanDo simulation is presented in Figure 21. These ignored parameters, especially the major-minor groove, can affect origami structure. Thus, before a design is ready for experimentation, it is necessary to check these effects with a finer model.

1.4.2 oxDNA

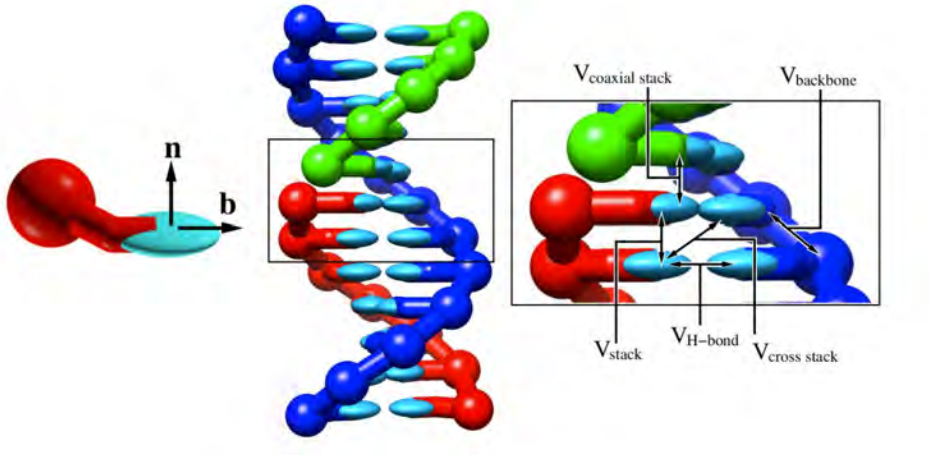

Figure 9: Schematic of the oxDNA model. Nucleotides are modeled as rigid bodies with interaction sites. Each nucleotide is defined by its position **r**, base vector **b**, and normal vector **n**. From (Doye et al., 2020)

A more detailed model of DNA dynamics is used in the simulation code called oxDNA developed by T. E. Ouldridge, J. P. K. Doye, and A. A. Louis at the University of Oxford (Doye et al., 2020). This model considers nucleotide level interactions as shown in Figure 9. The electrostatic potential of the negatively charged backbone, and the potentials of direct, coaxial, cross stacking, and the hydrogen bond between bases are all considered. This level of resolution allows the model to accurately describe the important features of origami designs including bending and twisting, splaying at Holliday junctions, and the overall global structure, all within experimental resolution (Doye et al., 2020). In addition, it can predict internal stresses, elastic properties, the effect of tension, and the thermodynamics of the origami design. The model can even describe the

hybridization and self-assembly processes of smaller designs.

While being much more detailed than mechanical models such as CanDo, oxDNA still has several limitations. Most relevant to DNA origami is electrostatic interactions. DNA origami structures are often highly sensitive to salt concentrations, with too little salt resulting in low stability of the structure and too much salt resulting in aggregation. This is because of the negatively charged backbone of DNA which repels the neighboring helices in origami structures. Most often concentrations of sodium or magnesium ions in the millimolar range are added. OxDNA is designed to model the electrostatic interactions on sodium ions only and has been calibrated to reproduce hybridization thermodynamics for sodium concentrations greater than 0.1 M (Doye et al., 2020). However, magnesium, which is a divalent cation, interacts with the DNA differently and is often used in origami solutions because of the particular stabilization of junctions, but is not modeled here. A workaround suggested by Doye et al. is to use a sodium ion concentration of 1 M which is approximately representative of the behavior of 12.5 mM $MgCl_2$ conditions. This means that the oxDNA simulation can still be used to check the stability and fluctuations of an origami design, but salt concentrations need to be optimized experimentally.

A feature of oxDNA especially important for DNA origami simulation is the relaxation of the initial geometry step. When an origami design is first loaded into oxDNA the configuration is very unnatural. Some DNA backbones are too long, and some nucleotides may overlap. If a simulation is run in this configuration, the origami structure would be subject to extremely large forces that cause the simulation to fail (Doye et al., 2020). Before running the molecular simulation, it is important to relax the origami configuration. This is done using a modified backbone potential.

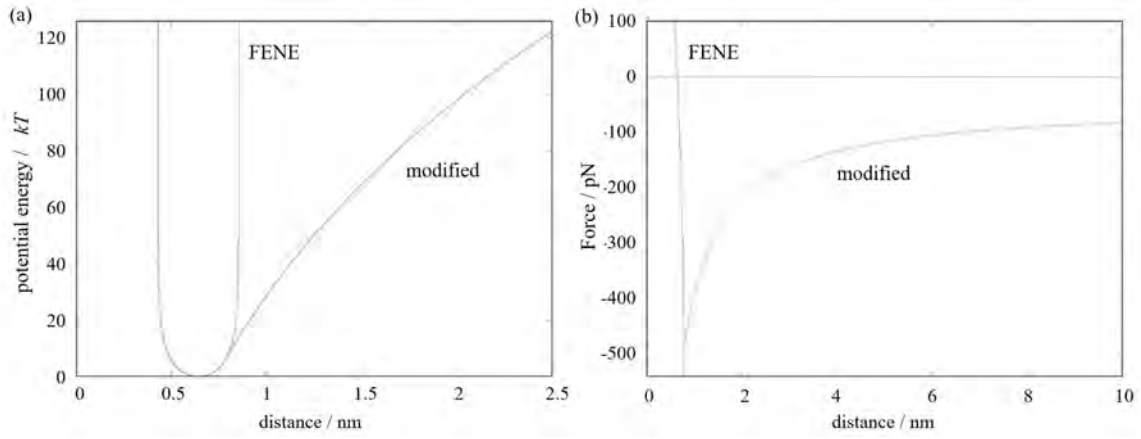

Figure 10: (a) The FENE backbone potential and the modified potential used for relaxation of origami configurations before molecular dynamics simulations. (b) the resulting forces from the FENE and modified potential. Note how the FENE potential diverges at 0.9 nm and results in extremely large forces, while the modified potential smooths this divergence allowing overstretched bonds to relax without being ripped apart. From (Doye et al., 2020).

The potential used in the simulation is a FENE "Finitely Extensible Nonlinear Elastic" potential. This potential is used to describe long-chained polymers by representing the polymer as a sequence of beads with nonlinear springs (Doye et al., 2020). This potential quickly diverges, however, at around 0.87 nm, resulting in extremely large forces on overstretched bonds. This is what breaks the origami apart and causes the simulation to fail (Doye et al., 2020). If the starting configuration is highly unnatural due to the earlier design pipeline of 2D structures, the modified potential serves to remove this divergence and allow stretched bonds to relax without experiencing damaging forces. The modified potential is as follows:

$$V_{mod} = V_{FENE} = -\frac{\epsilon}{2}\ln\left[1 - (\frac{r - r_0}{\Delta})^2\right] \text{ for } r \leq r_{max}$$

$$V_{mod} = Ar + B\log(r) + C \text{ for } r > r_{max}$$

Where r is the position of the backbone. r_{max} is the distance at which the force due to FENE is equal to a value F_{max} which is set to the force that would break the backbone. A is the limiting value of the force at large r and is also set in the simulation input files. B and C are chosen so that the modified potential is continuous and differentiable (Doye et al., 2020). More details on the parameters of the FENE potential as well as the stacking potentials used can be found in Appendix A of Ouldridge et al. (Ouldridge, Louis, & Doye, 2011).

1.5 Thesis Objective

Membrane protein structure reconstruction is notoriously difficult. First, the nanoscale size requires high resolution imaging techniques. Second, the hydrophobic native environment adds unique challenges in extracting and stabilizing the proteins. Third, low expression rates increase difficulty in using high sample requirement techniques like x-ray crystallography. Finally, low sample requirement techniques like cryo-EM have drawbacks like low resolution from nonuniform ice thickness, strong denaturing forces during freezing, and unfavorable orientation due to hydrophobic regions of the proteins interacting with the air-water interface.

To address these problems the aim of this project is the computer-aided design, simulation, and experimental verification of a DNA origami support structure to capture membrane protein mimetic systems. The structure will be designed to capture a nanodisc that can house membrane proteins in a native-like environment. It will also protect the protein from damaging forces in cryo-electron microscopy sample preparation, help form uniform ice layers, as well as orient the protein to aid in 3D reconstruction. This project combines the promising technique of electron microscopy with the novel method of DNA lipid nanodiscs. The aim is to provide a method for higher resolution 3D reconstruction studies of membrane proteins. Mapping 3D structures would allow improvements such as better development of binding mechanisms in drugs, which are essential to the treatment of human disease.

CHAPTER 2

DESIGN AND SIMULATION

The development of a DNA nanoscale device has two main phases: design and simulation, and experimental testing. The first phase, described here, consists of designing the shape, size, and capture method of the nanostructure as well as testing the stability of the design through computer simulation. Two separate structures are discussed. The first is an original design by the author and the second is a similar design created by Jing Huang and Barbara Saccà that is modified for this project (Huang, Gambietz, & Saccà, 2022). Ultimately, the second design was chosen to move on to the experimental phase, because the self-assembly protocols had already been optimized.

The design process consists of four steps. First, the cross-sectional shape and dimensions of the device had to be determined. Second, the routing of the long scaffold through the structure was decided. Finally, the complementary staples were added, and then stability of the design was tested using computer simulation.

2.1 Original Design

Initially, the DNA support structure of interest was designed using caDNAno, a computer-aided design software developed by Shawn Douglas (Douglas, Dietz, et al., 2009) (Douglas, Marblestone, et al., 2009). Because the length and sequence of hundreds of DNA building blocks need to be determined before strands can be synthesized and mixed to self-assemble, computer-aided design is an essential tool. In this section, the design software caDNAno is introduced. Then, the requirements for the overall dimensions of the structure are discussed.

2.1.1 Software Introduction

In the caDNAno software, the DNA structure is schematically laid out in two panels, the slice panel, and the path panel. The slice panel depicts a cross-section of the design, while the path panel is a schematic side view. Each colored line represents a single strand of DNA. The relationship between these views and a 3D rendering of a simple example design is shown in the following Figure 11.

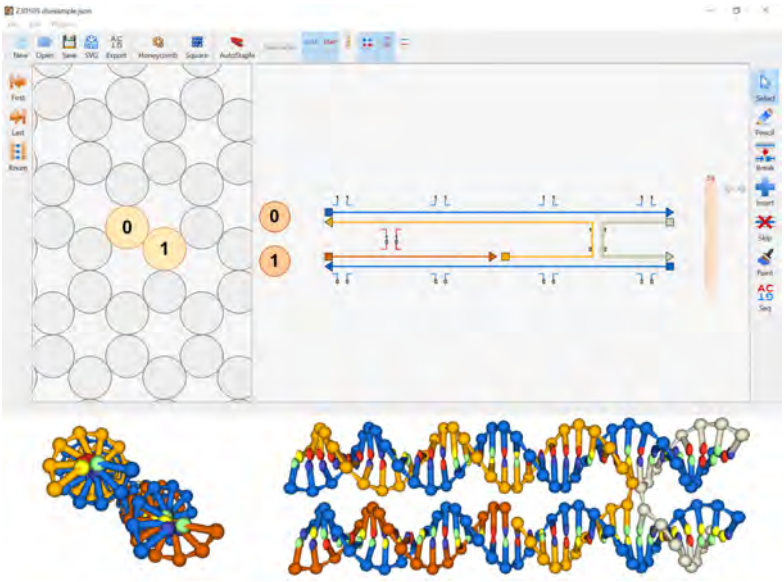

Figure 11: Image depicting the relationship between caDNAno software schematics and 3D renders for an example structure with two double-stranded helices. The slice panel depicts a cross-section of the structure (LEFT) while the path panel shows a side-on view (RIGHT).

The slice panel allows additional double-stranded helices to be added to a design file. It has two modes, hexagonal and square. Hexagonal mode is used to design 3D structures and links adjacent helices with Holliday-junction crossovers in a hexagonal shape. A rectangular mode is also available, which is better suited for flat designs.

The path panel contains much of the important information in a design. The dark blue strands represent the scaffold strand, the long DNA strand that is folded in DNA origami. The colorful shorter stands grouped immediately above or below the scaffold strand are the staple strands. The squares represent the 5' end of the DNA strand and triangles represent the 3' end, as DNA sequences are always read in the 5' to 3' direction (Douglas, Dietz, et al., 2009).

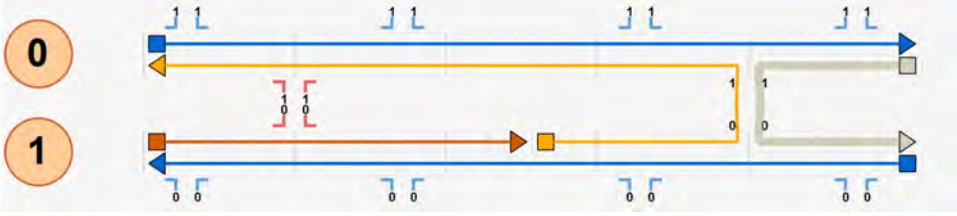

Figure 12: A close-up of the path panel with the example two helix structure, showing staple crossover locations (RED) and scaffold crossover locations (BLUE).

The most useful tool in this software is the crossover location markers. Since every 10.5 base pair

constitutes a helical twist, the closest proximity between the backbones of neighboring strands is restricted to 10.5 base-pair intervals. The design software restricts available crossovers to locations where backbones are the closest and where the angle between neighboring helices is 120°. For example, Figure 12 shows two helices labeled 0 and 1. On the right is an existing crossover where the yellow strand starts on helix 1 and switches to helix 0 and the grey strand from 0 to 1. On the left side, the next potential crossover is marked in red and is 21 spaces away.

Using this software, first, the cross-sectional shape of the structure was designed using the slice panel. Then the scaffold routing between each of the helices was determined and complementary staples were added. This was the most essential step in the design process because the scaffold routing and staple distributions greatly influence how well a structure folds and how stable it is.

Dimensions and Shape

Since the barrels are designed to capture nanodiscs, the cavity of the structure needs to be big enough to encapsulate them, as well as tall enough to prevent exposure of the membrane protein to the air-water interface during cryo-electron microscopy, thus preserving its structure. The design can be roughly described as an open hexagonal prism, which is referred to as a "barrel". Since the nanodiscs are 15 nm wide, the barrel has been designed to have an inner radius of 22.5 nm, be 46 nm tall, and have a side length of 26 nm (Figure 13).

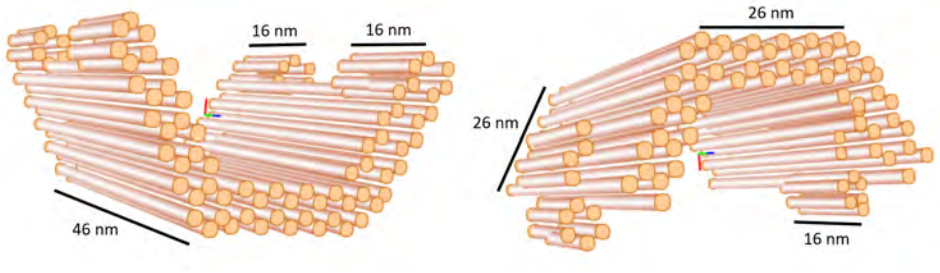

Figure 13: A schematic view of the two halves of the design with double helices represented as cylinders. The dimensions were chosen to capture and protect the 15 nm diameter nanodisc and membrane protein complex for cryo-electron imaging.

Because the purpose of the barrel is to encapsulate the nanodisc, how the nanodisc is captured is very important. The barrel and the nanodisc will be added together in solution and entropically, the chances of a nanodisc entering and attaching inside a closed barrel are not favorable. This means that a very low number of nanodiscs would combine with the barrels to form the desired

product. To solve this, the two halves will be constructed separately, the nanodisc will be attached to one half (say the bottom half) and then the top half will close around the nanodisc-left half pair as shown in the following figure.

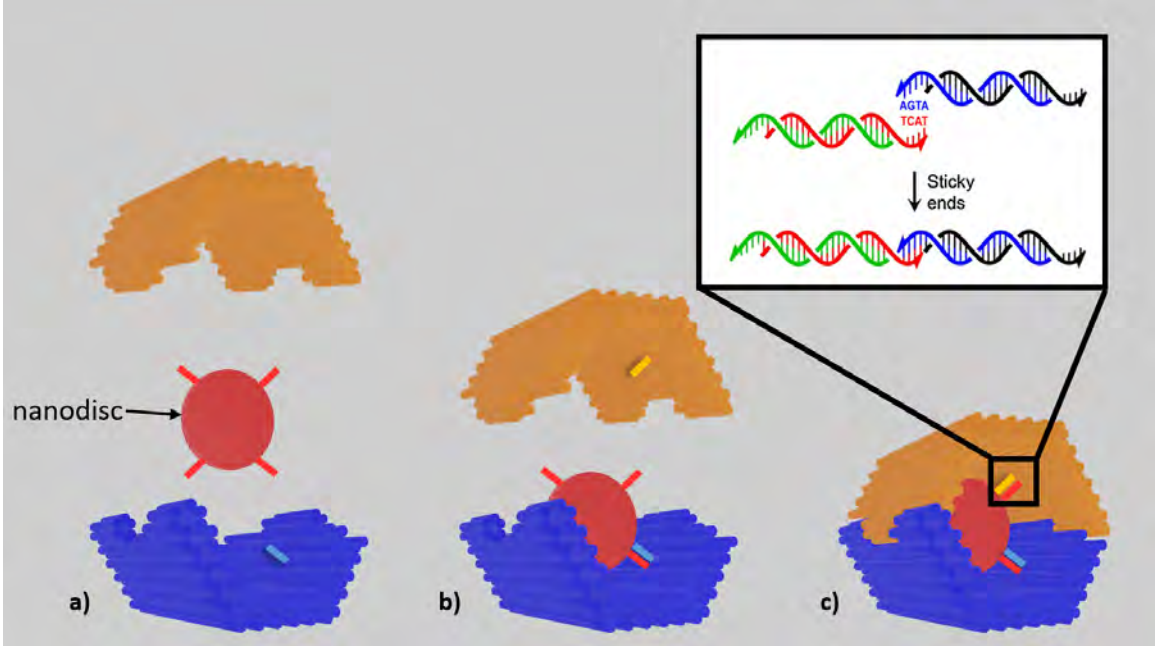

Figure 14: A schematic view of the two halves of the design with double helices represented as cylinders. The dimensions were chosen to capture and protect the 15 nm diameter nanodisc and membrane protein complex for cryo-electron imaging. The close-up shows the sticky-end interactions between the single-strand extensions binding the nanodisc and barrel, figure adapted from (Chandrasekaran & Zhuo, 2016).

The nanodisc binds to the inside of the barrel using DNA. Single-stranded extensions from the nanodisc will be programmed to be complementary to single-stranded extensions on the barrel. As shown in the close-up in Figure 14 these two stands will bind together forming a double helix. This joining of two structures with DNA is commonly referred to as a "sticky-end" interaction. While the close-up only shows four base pair sticky-ends, more base pairs will be used to increase the strength of the interaction and hold the nanodisc inside the barrel.

2.1.2 Early Designs

During the design process, several factors were investigated and optimized. The functionality of the protein image studies, the DNA structure stability, and the cost of the DNA were all considered. The two halves of the barrel are identical except for the edges where they fit together. One half

is called “the plug” because of the two protrusions resembling an electrical plug. The other half is referred to as “the socket” for its resemblance to an electrical socket. Thus, the two halves fit together as a plug fits into an electrical socket.

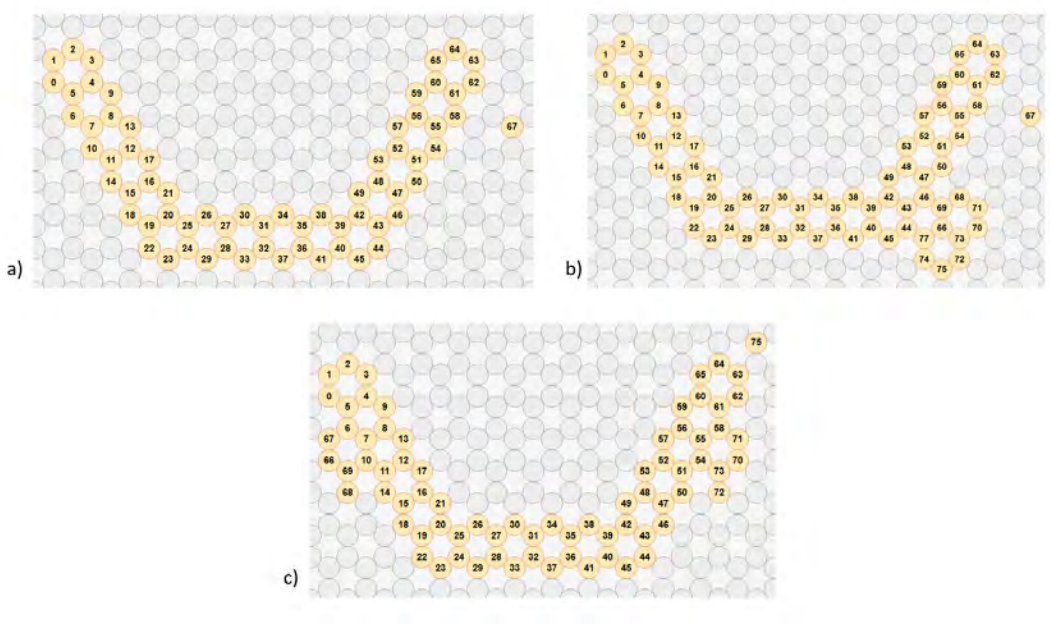

Figure 15: Slice panel view of plug barrel designs including a) the original cross-section with no additions, b) the “flag” design with the addition of an orientation marker for use in image reconstruction, and c) the cost-effective design with the addition of two small side structures to allow to re-use of staples between the plug and socket halves of the design.

Figure 15 depicts some of the different preliminary designs of the plug. Originally the plug had no additional features along the outside (Figure 15 a)). When the barrel forms it will have a six-fold symmetry along the plane of the face, making the orientation of the membrane protein impossible to determine and adding difficulty to 3D reconstruction. Figure 15 b) depicts the addition of an elucidation marker called a “flag.” The flag acts as an orientation marker by breaking the symmetry of the structure. In electron microscopy images, the orientation of open-side-up barrels can be identified and used in the reconstruction algorithm to better determine protein structures. However, the drawback of this asymmetry is the increased building cost of the structure. Because the overall goal of the membrane imaging project is to develop a method for imaging membrane proteins, the structure will be used repeatedly, and the cost of manufacture has to be considered. The asymmetry of the second design requires the plug and socket half of the barrel to have unique staples. Even though the two designs are geometrically the same except for the socket and plug

edges, no building blocks can be reused. In essence, this nearly doubles the cost. Figure 15 c) shows the addition of two small side structures that allow the plug and the socket halves to use the same amount of scaffold in each part of the design and allow similar parts of the design to be exactly identical. Thus, fewer unique staples are required to build the structure, reducing the cost by nearly 40%. Because of the significantly reduced cost, the third version (15 c) of the plug half of the barrel was chosen for further work. However, the "flag" design may become useful in future membrane protein imaging studies that may benefit from an orientational marker during image reconstruction.

2.1.3 Scaffold Routing

The design process required several iterations of the scaffold routing and staple crossovers at each stage. In this phase of the project, no experimental testing could be performed, so design considerations were chosen based on past successful designs in the field as well as simulation testing. The first step is to determine the best route for the scaffold strand. This is the long single-stranded DNA that gets twisted into the desired shape and is held in place by staple strands. Standard scaffold strands are sourced from bacteriophages and have known sequences. The scaffold strand used in this design is called p8064 and is 8064 bases long. This scaffold needs to be carefully routed in such a way as to increase the stability of the origami structure. Unstable routings could cause the structure to misfold or fall apart, making it unusable. In this section, two published scaffold routing methods are reviewed.

Raster routing with Sheets

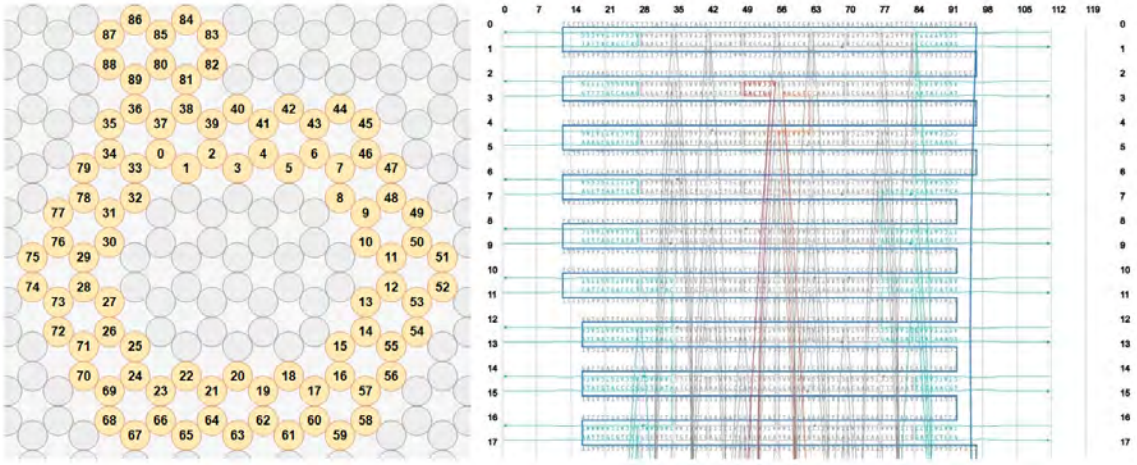

Figure 16: CaDNAno design schematic of the cryo-electron imaging transcription protein support structure from (Martin et al., 2016). The design employs raster-style scaffold routing (RIGHT) and numbers the helices in a linear sheet fashion (LEFT).

The first method of scaffold routing was used for the development of a DNA support structure that was a 26.2 nm tall closed hexagonal prism with a 13.6 nm diameter cavity (Martin et al., 2016). The scaffold routing technique as shown in Figure 16 had two main characteristics. First, as seen on the left, the scaffold was routed linearly through the helices, essentially creating sheets of connected helices that would wrap around themselves to form the structure (note how helices 0-33 form a sheet that helices 34-79 wrap around). This requires the structure to fold into a plane first before wrapping around itself to form the 3D structure. However, this may not be favorable because DNA has to interact over large distances. Second, as seen on the right, the scaffold moved through each helix for its entire length before moving down and back through the next helix in a raster-style pattern (right-down-left-down-right, etc.). In the study, the structure was folded and purified with a yield between 50% and 90% (Martin et al., 2016).

Mid-seam routing with Bundles

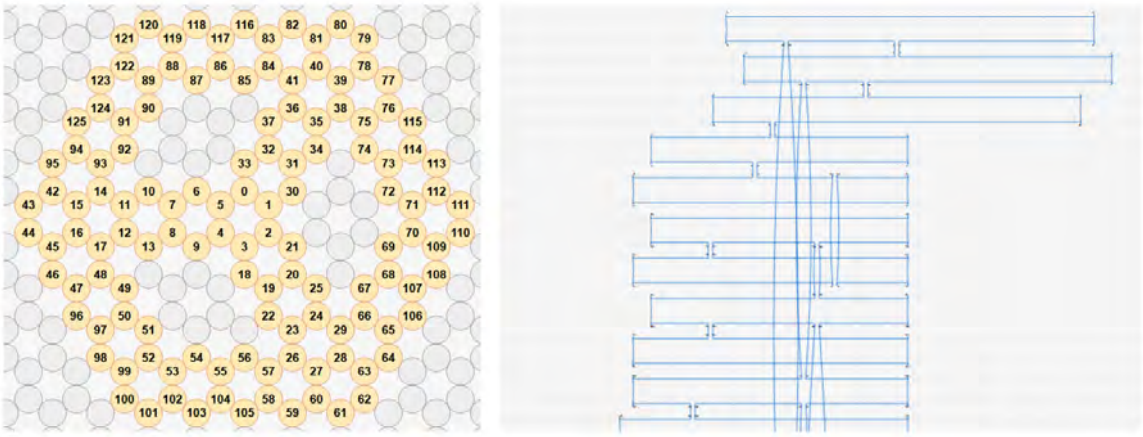

Figure 17: Modified caDNAno schematics for "Wheel pore" design for nanopore studies from (Krishnan et al., 2016) where the staple strands are removed for clarity. The design employs mid-seam style scaffold routing (RIGHT) and numbers the helices in groups (see helix 0-5, 6-9, 10-13, etc. LEFT).

A second successful routing technique was developed for a DNA origami structure called the "wheel pore" (Krishnan et al., 2016). This structure was 28 nm tall with a 20 nm inner diameter filled with additional DNA helices that functioned as nanopores. Unlike the previous technique, the helices were numbered closer together, forming bundles of helices before moving on to the rest of the structure (note the grouping of 0-5, 6-9, 10-13, etc.). In addition, the scaffold within each helix was routed such that the scaffold traveled only halfway through the helix before crossing over creating what resembles a "seam" down the center. This method is called a "mid-seam design" and is utilized in many published structures such as the 2D shapes created by (Rothemund) and in the 6-helix bundles (Gür et al). While the yields for the wheel pore are not reported, the yields for the 6-helix bundles by Gür are around 90% and the yields for the successful 2D shapes were at least 70% (Gür et al., 2016) (Rothemund, 2006).

This second method of bundling neighboring and using mid-seam routing was chosen for this project for its diverse record of success. In addition, because our nanodisc capturing structure is so large compared to other structures, the mid-seam routing might promote better folding. This is because folding is dominated by short-range interactions where the staples are more likely to connect spatially close domains (Dunn et al., 2015). In the mid-seam design, there are more scaffold crossovers between the neighboring helices increasing the probability that staples will make

desired connections. The one disadvantage of the method is that the seam in the middle could be less stable. However, this could be mediated by careful staple connections that close the seam (Selnihhin & Andersen, 2015).

2.1.4 Staples

After finalizing the scaffolding routing for the two halves of the barrel, the next step is designing the staple connections. The caDNAno software includes an "Auto-Staple" tool that populates a given design with staples according to the 21 base pair rotation crossover rules (Douglas, Marblestone, et al., 2009). These staples need to be manually adjusted to meet a set of requirements known to promote better folding. This section describes these requirements as well as three additional features, including "sticky-ends" interactions to support the joining of the barrel, symmetrical staples between the two halves, and poly-thymine (poly-T) extensions to reduce aggregation between barrel structures.

Requirements

After populating the design file with staple paths, it is necessary to optimize the length and annealing regions of each staple. To promote self-assembly staples are recommended to have lengths between 18-49 bases with a mean of 30-42 bases (Douglas, Dietz, et al., 2009). For our structure, the maximum length of the staples was extended to 60 base pairs. In addition, studies on the folding pathways of DNA origami determine that staples with 14 base annealing regions, or uninterrupted stretches connected to the scaffold, can dramatically increase self-assembly success (Ke, Bellot, Voigt, Fradkov, & Shih, 2012). As such, care was taken to observe each of the 455 staples to ensure that there was a region of at least 14 bases to optimize the attachment of the staple during self-assembly. As seen in the following table, the staples in the barrel structure meet design requirements in minimum and maximum staple length and mean. The standard distribution and the number of staples that are identical between designs, called "cores," are also reported.

Condition	Socket	Plug
# < 18	0	0
# > 49	27	28
Max (bp)	60	60
Mean (bp)	40.8	39.7
Median (bp)	41	41
Mode (bp)	41	41
Standard Deviation	7.5	8.6
Core Count	159	159
Total #	230	225
% core	0.69	0.71

Table 1: Staples Distribution Check. Analysis of staples distribution in the socket and plug barrel design. The acronym (bp) stands for base pairs or the number of nucleotides in the length of the staple strand. Core staples refer to staples that are identical between the plug and socket design.

Features

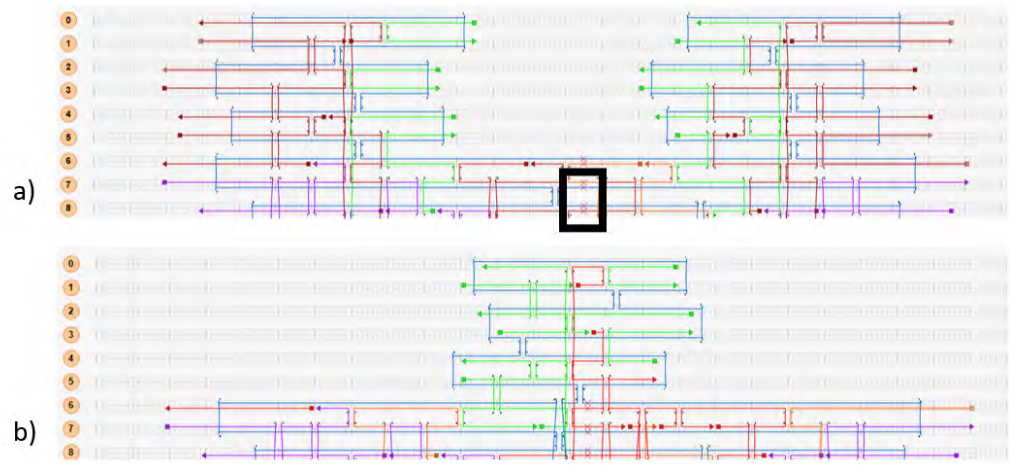

Figure 18: caDNAno design files for the first 15 helices in the socket (a) and plug (b) designs. The continuous blue line indicates the scaffold strands. The red, orange, green, and purple lines indicate staple strands. Green strands include sticky-end extensions or recessions. Purple strands indicate that the staple is identical between designs and red/orange strands indicate the staple is unique between designs. Single-stranded extensions are visible on both sides of each structure and 10 base extensions of thymine.

One helpful feature of the staple design is single-stranded poly-T extensions. At the top and bottom edge of the structure, single-stranded staples extend out by 10 bases (all thymine). These staples can be seen extending out on the left and right of the blue scaffold in Figure 18. Poly-T extensions are commonly used in origami designs to prevent multiple structures from sticking together. The negatively charged backbone of the extensions will repel other structures preventing unwanted aggregation. For cryo-electron microscopy, these overhangs also serve an additional purpose to orient the support structure. Martin et al. found that the exposed aromatic bases like those of poly-T extensions interact with the hydrophobic air-water interface during freezing, creating a larger distribution of structures orienting face up as desired for protein imaging applications (Martin et al., 2016).

Another feature is the identical staples between the plug and socket designs. These symmetrical staples are colored purple in Figure 18. Once the scaffold routing was matched up between the plug and socket with the side extensions in the plug design as described previously (see image c in Figure 15), the staples could be configured by matching start and end points. 159 staples were able to be matched, reducing the synthesis cost for the structure.

The last feature is the addition of sticky ends at the plug-socket interface. Originally, the only mechanism that dimerizes (holds together) the two halves of the barrel is the stacking interaction where DNA stabilizes with nucleobase stacking bonds (Gerling, Wagenbauer, Neuner, & Dietz, 2015). This occurs with DNA structures that fit together geometrically, such as the socket and plug. The nucleobase stacking bonds compete with electrostatic repulsion from the negatively charged backbone of DNA (Gerling et al., 2015). With multilayered shape complementarity, these stacking interactions can stabilize the structure but are very sensitive to cation concentrations in the solution (Gerling et al., 2015). To further strengthen the binding of the plug and socket halves, sticky ends were added. These are four nucleotide extensions and complementary retractions at the ends of the DNA helices that participate in stacking interactions (see Figure 14). This allows the extended bases to bind to the exposed scaffold in the other half of the design, thus dimerization occurs both by stacking and base-pair interactions, closing the barrel.

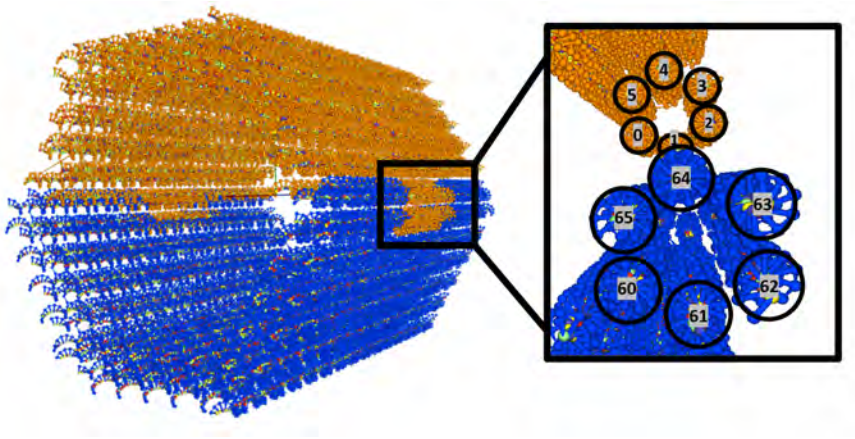

Figure 19: 3D render of the closed barrel design with the plug half (orange) and the socket half (blue). Numbers on the helices correspond to the numbering in the caDNAno files and indicate which helices of the plug align with which helices of the socket when forming dimers determining complimentary sticky-end sequences.

Figure 19 shows the geometry of the stacking interactions. The orange plug joins with the blue socket, and helix 1 will align with helix 61. The sticky ends are thus designed such that the extensions of helix 1 will be complementary to the exposed scaffold in helix 61. When the plug fits into the socket the sticky ends will hybridize the two half barrels together.

Analysis and Checks

Because of the large number of staples and manual changes implemented during the design process several checks in Excel were utilized to ensure that staples remained within length parameters and that core staples were indeed identical in both designs.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Start	End	Sequence S4	Length	Start/End =	Seq =	Length =	Start	End	Sequence P4	Length	
2	31[140]	36[136]	TTTAATTTAAAGGAATTGC	32	TRUE	TRUE	TRUE	31[140]	36[136]	TTTAATTTAAAGGAATTGCC	32	#aa00ff
3	36[114]	39[125]	TTGTGCTGGTAATATCCCT	31	TRUE	TRUE	TRUE	36[114]	39[125]	TTGTGCTGGTAATATCCCTA	31	#aa00ff
4	36[135]	41[161]	TTATATTACCGCCAGCCAT	32	TRUE	TRUE	TRUE	36[135]	41[161]	TTATATTACCGCCAGCCATT	32	#aa00ff
5	40[86]	33[97]	TTAGAGTAGATGATTAGTA	44	TRUE	TRUE	TRUE	40[86]	33[97]	TTAGAGTAGATGATTAGTAC	44	#aa00ff
6	46[94]	35[90]	TTAGACTTTACATCGCCCA	45	TRUE	TRUE	TRUE	46[94]	35[90]	TTAGACTTTACATCGCCCAG	45	#aa00ff
7	18[111]	23[97]	TTACGAGGCGTGGTGCCAT	38	TRUE	TRUE	TRUE	18[111]	23[97]	TTACGAGGCGTGGTGCCAT	38	#aa00ff
8	9[138]	9[165]	TGTTGGGAAGGGCGATCG	28	TRUE	TRUE	TRUE	9[138]	9[165]	TGTTGGGAAGGGCGATCG?	28	#aa00ff
9	29[47]	30[53]	TGTTACTTAGCCGGAAACA	40	TRUE	TRUE	TRUE	29[47]	30[53]	TGTTACTTAGCCGGAAACAA	40	#aa00ff
10	23[85]	19[85]	TGGTTGGCCAGGGTGGTTT	29	TRUE	TRUE	TRUE	23[85]	19[85]	TGGTTGGCCAGGGTGGTTT	29	#aa00ff
11	17[119]	10[126]	TGGCAGCCTCCGGCAACGA	35	TRUE	TRUE	TRUE	17[119]	10[126]	TGGCAGCCTCCGGCAACGA	35	#aa00ff
12	47[53]	35[44]	TGCCTTTGCCCGAACGTAA	48	TRUE	TRUE	TRUE	47[53]	35[44]	TGCCTTTGCCCGAACGTAAT	48	#aa00ff
13	19[46]	16[20]	TGAGGTCATTGCAGGCGCT	46	TRUE	TRUE	TRUE	19[46]	16[20]	TGAGGTCATTGCAGGCGCT	46	#aa00ff
14	26[76]	20[68]	TGAACCATCACCCAGGGCC	23	TRUE	TRUE	TRUE	26[76]	20[68]	TGAACCATCACCCAGGGCG	23	#aa00ff
15	36[108]	32[98]	TCTTCACTGAGGTTGCTTTC	25	TRUE	TRUE	TRUE	36[108]	32[98]	TCTTCACTGAGGTTGCTTTG	25	#aa00ff
16	35[91]	39[98]	TCGTAAATGAATTTTTACCC	33	TRUE	TRUE	TRUE	35[91]	39[98]	TCGTAAATGAATTTTTACCG	33	#aa00ff
17	20[67]	16[56]	TCCCTGCCCTGCGGAGGTG	40	TRUE	TRUE	TRUE	20[67]	16[56]	TCCCTGCCCTGCGGAGGTG	40	#aa00ff
18	8[139]	13[165]	TATTACGGATTAAGTTGGG	40	TRUE	TRUE	TRUE	8[139]	13[165]	TATTACGGATTAAGTTGGG	40	#aa00ff
19	57[92]	48[94]	TATCCGAGTATGTTAGCAA	34	TRUE	TRUE	TRUE	57[92]	48[94]	TATCCGAGTATGTTAGCAA	34	#aa00ff

Figure 20: A screenshot of an excel file comparing staples strands from the socket design (S4) with the plug design (P4). Formulas in columns E through G check if the start/end position, sequence, and length are identical between the two designs. This serves to verify the sequences of the oligonucleotides to be ordered.

Along with the staple analysis in Table 1 that checks the lengths and distribution of the staples, Figure 20 demonstrates how the core staples were processed. The file of sequences generated by the caDNAno software is used to check that the length, start-end location, and sequence were identical in the socket (labeled S4) and the plug (labeled P4) designs. Once the staples met the design requirements and the sticky-end sequences were determined the design of the staples was finished.

2.1.5 Computer Simulation

In this study, two types of computer simulation were used to assess the designs, called CanDo and oxDNA. As discussed in the introduction, CanDo models the DNA structure as elastic rods that can stretch, twist, and bend and oxDNA is a coarse-grained model Monte Carlo simulation and represents the thermodynamic and mechanical properties of DNA. The results of the CanDo and oxDNA simulations on the barrel design are presented below.

CanDo

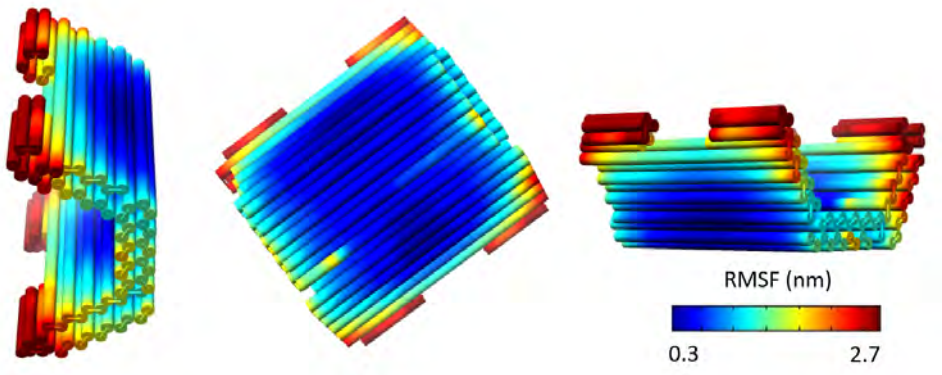

Figure 21: CanDo simulation of the Socket Barrel viewed from three orientations. The DNA double-helices are modeled as elastic rods and flexibility is represented as root-mean-square fluctuations. Red represents the highest fluctuations and blue represents the smallest.

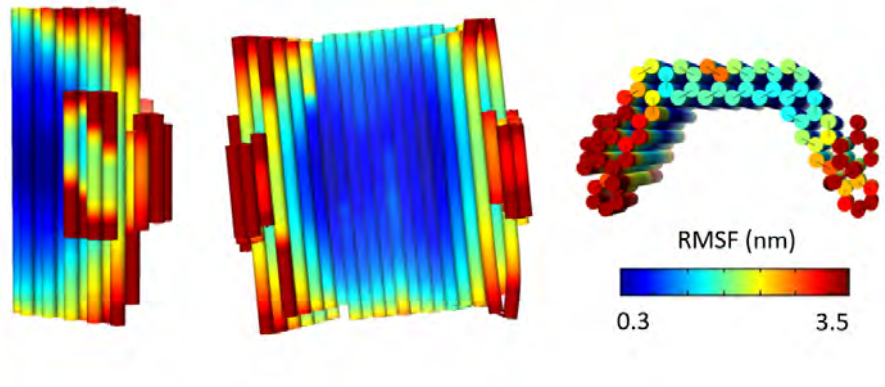

Figure 22: CanDo simulation of the Plug Barrel viewed from three orientations. The DNA double-helices are modeled as elastic rods and flexibility is represented as root-mean-square fluctuations. Red represents the highest fluctuations and blue represents the smallest.

Figures 21 and 22 show the results of the CanDo simulation on the socket and plug halves of the barrel, respectively. The flexibility of a helix is represented as a root-mean-squares heat map, with red being the most flexible and blue being the least (Castro et al., 2011). In both designs, the center of the structure is the most stable with the groups of helices and crossovers promoting desired stiffness. The socket and plug extensions of the design are the most flexible because the helices lack stability from close neighbors, and the design constraints also called for fewer crossovers in these portions. Noting this, some crossovers were added to decrease the flexibility of the plug-and-socket interface, but because of the nature of the design they are still quite flexible. It is predicted that once the dimers form by stacking and sticky-end interactions, the flexibility will drop dramatically, and the structure will be adequately stiff for our purposes.

oxDNA

Next, the designs were simulated using a more powerful coarse-grain model, oxDNA. This model can account for more features of the DNA such as major/minor grooving, electrostatic effects including in the solution, coplanar base stacking, and sequence-dependent interactions (Sengar, Ouldridge, Henrich, Rovigatti, & Šulc, 2021). The program is also much more computationally heavy than CanDo. Compared to the half-hour computation times, oxDNA models run of these barrels took upwards of 12 hours on the CPU and 8 hours on GPU. Long-term simulations (one millisecond) took three days of computation time.

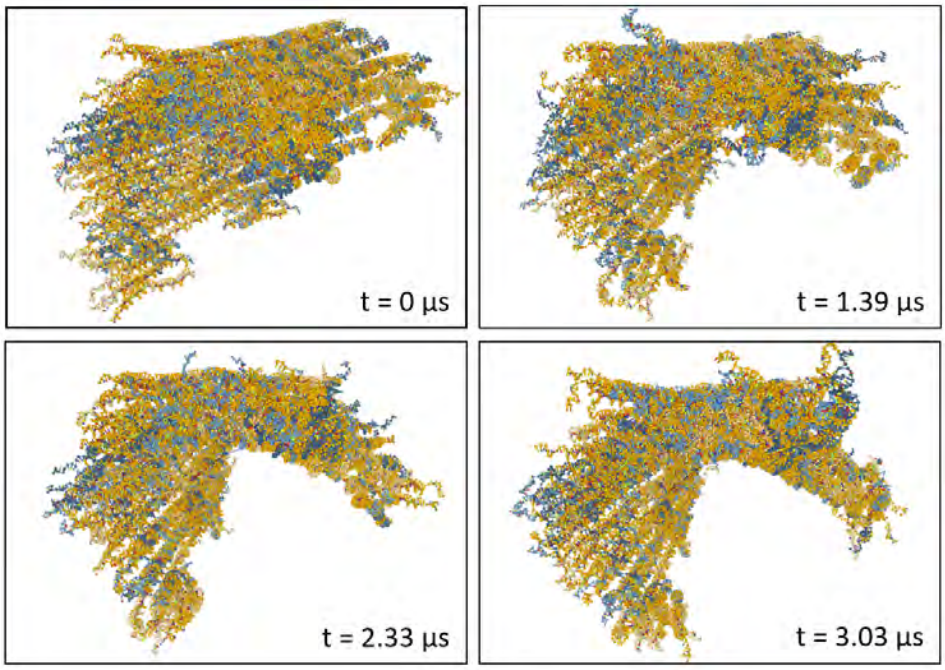

Figure 23: oxDNA simulation of the socket half of the barrel with a large structural twist.

The first of two design problems revealed by oxDNA is shown in Figure 23. The results from an older version of the socket barrel show the development of a large twist throughout the structure. This occurs because caDNAno assumes the DNA has exactly 10.5 base pairs per turn which very slightly overestimates the actual geometry of DNA in the structure. When a structure is hundreds of base pairs long this can result in torque forces that create a global twist (Dietz, Douglas, & Shih, 2009). To combat this right-handed twist deformation, a counter-left-handed twist is created by base pair deletion. This base pair deletion is shown as a red X in the middle of the structure in Figure 18.

An additional problem encountered when simulating with oxDNA is topological in nature. There are single-stranded loops of scaffold at the end of the designs which serve to use up the excess scaffold and allow for crossovers that create the desired geometry. When the loops are converted from the caDNAno design file to the oxDNA configuration file the loops can load physically within double-stranded helices within the design, making it impossible for the simulation to untangle them. This topological problem causes nonphysical forces that rip the structure apart, ruining the simulation. More careful positioning of the single-stranded loops in both the caDNAno and oxDNA

configuration files solves this. Once the twist and topological problems were solved the resulting simulation video from the oxDNA analysis can be used to assess the structure. On the most basic level, the simulation serves to verify that the structures do not fall apart under their own strain. From progressive snapshots at various times in the simulation videos (24), the structure fluctuates as expected due to internal forces and Brownian motion while successfully staying together.

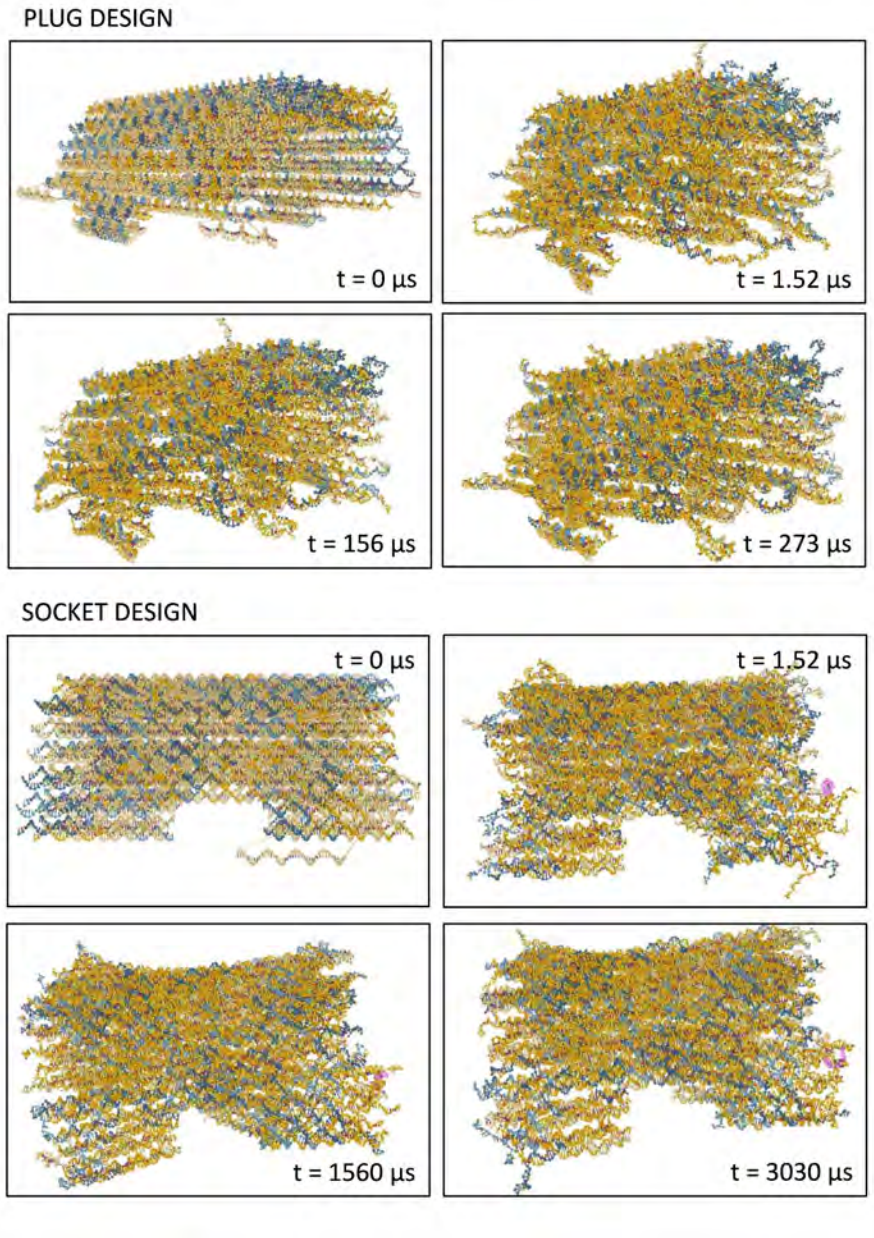

Figure 24: oxDNA Simulation of the final socket and plug barrel design. The structures exhibit no global twist, and the single-stranded scaffolds do not block the plug-socket interfaces.

We also wanted to check if anything (i.e., the single-stranded loops of the scaffold or the two extensions in the plug design) ever moves and blocks the plug-socket binding areas as such an event could prevent the desired formation of dimers. In Figure 24 at t = 0 the unrelaxed single-stranded loop is shown at the bottom of the structure. In the subsequent snapshots, this loop is highlighted in pink and shows that it moves toward the outside of the structure and does not linger near the socket interface. Similarly, the snapshots from the plug design reveal no interference of the features with the plug interface.

2.2 Modifying Nemesis Design

During the design of the original plug and socket barrel, collaborators Jing Huang and Barbra Saccà from the University of Duisburg-Essen shared a similar structure that had already been developed for another purpose. This design could be easily modified for nanodisc attachment and since it is already experimentally tested, we know that it will self-assemble. Unlike the original design which has not yet been experimentally verified, it would not need time-consuming optimization of folding temperature. As such the project shifted from the original design to the modifications of this design called the "Nemesis Cage." The following sections introduce the original "Nemesis Cage" as well as the modifications that were done to optimize it for this study.

2.2.1 Original Echo- Narcissus Design

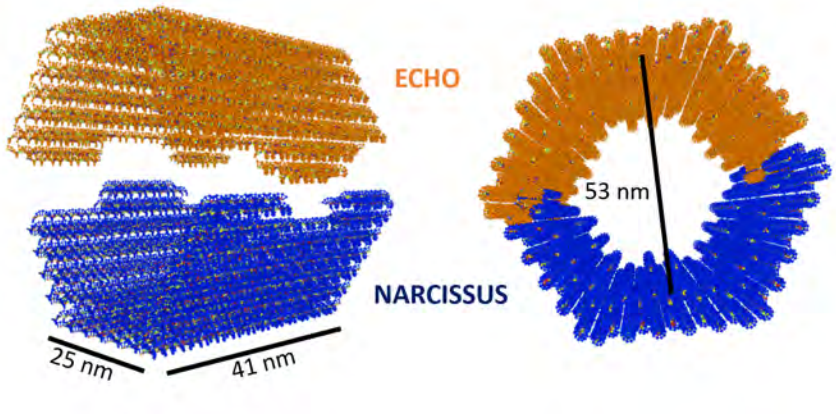

Figure 25: 3D render of the DNA nanostructure created by Huang and Saccà, referred to as the "Nemesis Cage." The orange half of the structure is called "Echo" and the blue half is called "Narcissus." The two halves are geometrically identical but contain different staple crossovers and sequences.

The original Nemesis cage consists of two different halves of a hexagonal barrel each called "Echo" and "Narcissus." They are quite similar in dimension to the plug-socket design; however, one key difference is the socket and plug interface. As seen in Figure 25, both the blue and orange halves of the barrel have a plug extension on one side and a socket cutout on the other. Echo and Narcissus have the same geometric design but are turned by 180°around the central axis of the hexagonal prism. In essence, they are self-complementary shapes. The structures can form homodimers (Echo-Echo or Narcissus-Narcissus) or heterodimers (Echo-Narcissus). Heterodimers are formed with the use of hybridization staples at the plug-socket interface.

2.2.2 Modification: Self-Complementary Dimers

For protein studies, homodimers (Echo-Echo) are more cost-efficient than heterodimers (Echo-Narcissus). However, it is important to control when the two halves dimerize. The nanodisc needs to be attached to one half of the barrel first before the second half of the barrel can close the structure. The goal is to modify the Nemesis cage to reduce the number of unique staples needed, but still control when dimerization occurs.

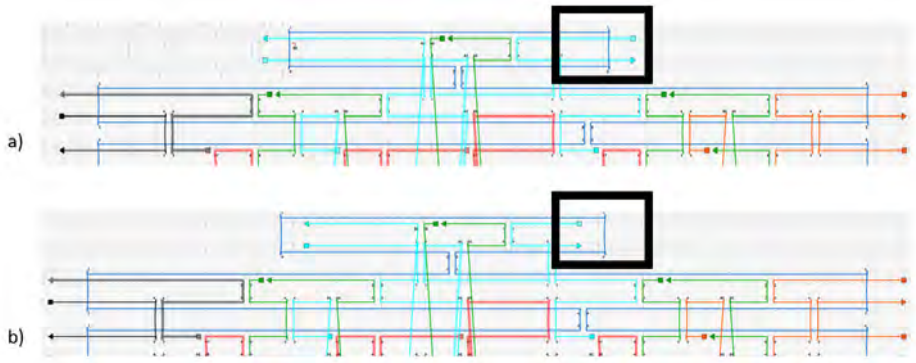

Figure 26: Top portion of the caDNAno schematics of the modified Echo designs. The overhang design (a) includes four base long extensions on the staples at the plug and socket interfaces. The recessed design (b) leaves four bases of the single-stranded scaffold at the plug and socket interfaces. The two designs are identical except for the four base pair changes.

The Echo half was modified into two different structures called the "Overhang" and "Recessed" designs. The overhang design has longer staples at the interface (indicated by rectangles in Figure 26) which should repel other overhang designs preventing dimerization. The recessed design has areas of single-stranded scaffold at the interfaces which should both affect the geometry and the ability to have stacking interactions as well as act like a "thermal brush" and repel other recessed

designs. The idea is that the two designs cannot form dimers until they are mixed together, allowing the nanodisc to be attached to one half before closing the hexagonal prism. Once the overhang and recessed designs are mixed, they should dimerize based on the same sticky-end interactions described in previous sections.

2.2.3 Nanodisc Attachment Design

Once the design of the structure was optimized, the next step was to design the attachment of the nanodisc inside. The nanodisc consists of a ring of DNA that is 15 nm in diameter containing chemical modifications that allow it to hold a patch of lipids where the membrane protein can reside. The outside ring of DNA has single-stranded extensions of known sequences that can be used to attach to complementary sequences like a Velcro strip. Adding these single-stranded complementary sequences as overhangs in the barrel structure, the nanodisc can be attached inside. This is easily done by extending select staples in positions inside the barrel.

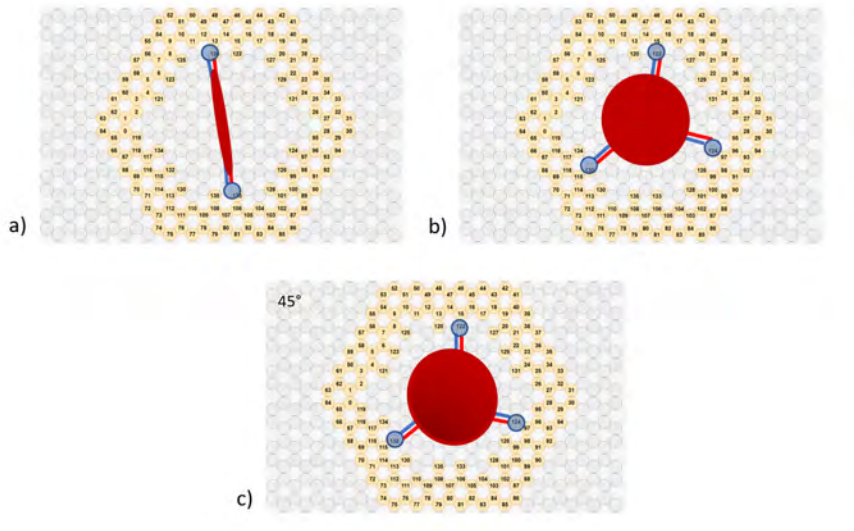

Figure 27: caDNAno slide panel schematics of the modified Echo-Echo design with red circles representing the lipid-nano disc in different orientations, vertical, horizontal, and 45 degrees.

Because the end goal of this larger project is the 3D image reconstructions of membrane proteins, it is also important to add the ability to attach the nanodisc in three different orientations: vertical, horizontal, and 45 degrees (Figure 27 a-c respectively). The design by Huang and Saccà already has 16 attachment points shown as helices 120 through 134 in Figure 27. Utilizing these, several different attachment configurations for experimental testing can be added. Figure 27 shows attachment

configurations using three staple strands (two at different points along the axis into the page in the vertical case). The design allows for flexibility in the choices of which staple positions are used to attach the nanodisc. This allows future experimental optimization studies to determine the best attachment position and methods.

CHAPTER 3

EXPERIMENTAL RESULTS

After completion of the design and simulation of the barrel, experimental testing was carried out. The main focus was on optimizing the salt and temperature conditions for robust and stable folding of the barrel halves. In the near future, the conditions of dimerization for the two halves will be optimized. Though the computer simulations indicate a stable design once it is folded, the simulations cannot guarantee that the design can successfully self-assemble. This is because the simulations do not consider the dynamics of the staples' thermodynamic motion and how each staple finds and binds to the correct part of the scaffold. The exact mechanics of the hundreds of staples are unknown and the conditions for proper folding are determined through experimental trials. Salt conditions are particularly important. This is because the phosphate backbone of DNA is negatively charged, and in the absence of counter ions like sodium (Na^+) and magnesium (Mg^{2+}), the electrostatic repulsion could be enough to destabilize the structure (Hong et al., 2017). Experimental determination of the best concentrations of counter ions is done by a technique called a "salt scan." In this technique, the same structure is folded under the exact same conditions but in varying concentrations of salt solutions. Characterization of the results reveals the most suitable concentrations that give well-folded and stable structures.

The main characterization technique for this study is agarose gel electrophoresis. In this method, DNA samples are placed in wells at the top of a porous gel and are separated into lanes. An electric field is pulled over the gel. This electric field pulls on the negatively charged DNA structures, drawing them toward the bottom of the gel. Because the gel is porous, smaller DNA strands can move quickly, while large DNA nanostructures are slowed down by the gel. This method allows DNA samples to be separated by size. Based on the location of sample bands in the gel, the completeness of the origami folding can be qualitatively analyzed.

3.1 The Recessed-Overhang Design

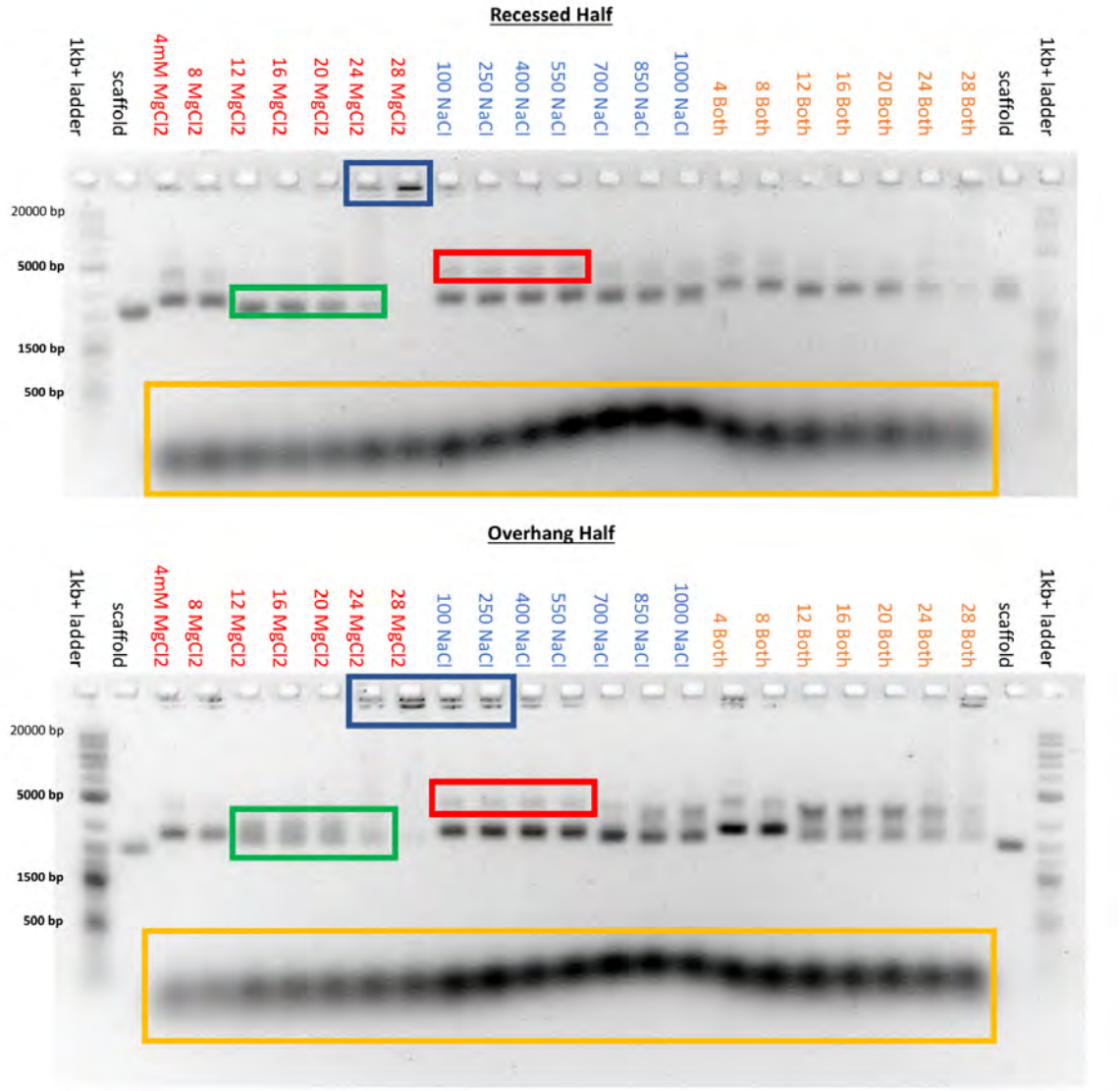

Figure 28: Agarose gel results of the Recessed and Overhang barrel salt scan. The gel depicts three groups of salt concentrations in units of millimolar: $MgCl_2$, NaCl, and both. Both indicate the concentration of $MgCl_2$ with an additional 10 mM of NaCl. The excess staple bands are marked in yellow; the potential origami bands are marked in green; the potential misfolded structures or dimers are marked in red, and hanging lane aggregates are marked in blue.

The first test of the design included the folding of each half-barrel named Overhang and Recessed (Figure 26). The two halves are identical except for the sticky ends in the plug and socket interface, with the Overhang barrel having all four base-pair extensions and the Recessed missing the symmetrical four base pairs at the interface. The designs are folded separately in large salt

scans. The following gels depict three groups of salt concentrations 4 to 28 mM of $MgCl_2$, 100 to 1000 mM of NaCl, and "both" which indicates 10 mM of NaCl with 4 to 28 mM of MgCl $_2$. Both ranges of monovalent and divalent counter ions are tested because of the differences in stabilization interactions.

Figure 28 depicts the results of the salt scan for the Recessed and Overhang barrels. On either side of the gels are control lanes. The outside lanes are called ladders which are samples with known bands that can be used to measure the migration of the samples. The second lane is the 7650 bp scaffold used to build the modified Echo design. Marked in the yellow box in both gels are the excess single-stranded staples. The design is folded with 10 times more staples than the scaffold to encourage the probability of a staple binding to the complementary portion of the scaffold. The black smear in each, called bands, does not show much information about the results of the folding except that the excess staples are present. One interesting feature is the wave-like pattern of the staple bands across the gel. Here we see that the staples with a high concentration of NaCl lanes moved slower (are higher up in the gel) than the rest. This is because the sodium cations shielded the electric field used in gel electrophoresis.

The other bands are identified by how far they migrated. The correctly folded design should appear as a band that has migrated close to but slightly less than the distance of the control scaffold band. This is because the design uses the same 7650 bp scaffold but should be slightly more compact. Example bands that are likely origamis are marked in green. If the design is well-folded it should present as a single dark and thin band. More than one band in a lane, like the bands marked in red, indicates that the structures may be aggregating together, misfolding, or forming dimers. The migration of these bands is slower than expected, indicating that the structure is much bigger than the theoretical origami. Strangely folded shapes, aggregated origamis, and dimers all could be possibilities. Extreme aggregation presents as hanging lanes (some examples are marked in blue). The presence of hanging lanes indicates that so much DNA was packed together that it could not move at all through the gel. These preliminary salt scans indicate that the most promising salt concentrations are 12 to 28 mM of $MgCl_2$ as well as 700 mM NaCl for the Overhang barrel.

Comparing the Recessed and Overhang gels, the Overhang design shows blurry origami bands. They are faint and smeared out over a larger distance. This indicates potential misfolding of the structures because correctly folded structures should be uniform in size and thus travel uniformly

within the gel. In addition, this gel also has more hanging lanes which indicate a lot more aggregation. All of this indicates that the Overhang barrel does not fold as well as the Recessed barrel. This is interesting because out of the hundreds of staples of DNA, only 24 staples are changed by extending or retracting four bases. It appears that these four base extensions greatly affect the self-assembly of the structure.

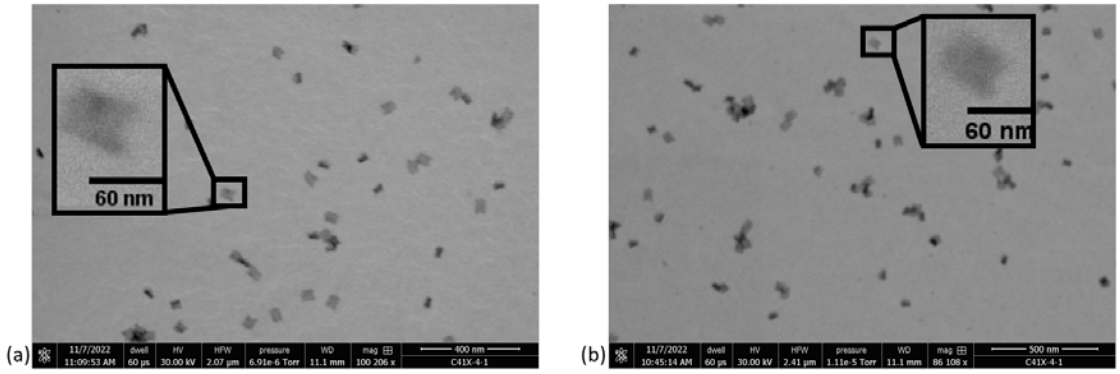

Figure 29: Electron microscopy images of the DNA nanostructures, (a) the Overhang half-barrel and (b) the Recessed half-barrel. Images indicate that partially folded structures of the correct size are formed but are not well-defined and are aggregated. All microscopy images were acquired with the help of Praneetha Sundar Prakash.

Because the agarose gel is just a qualitative measure that can only indicate the size of the sample, to confirm that the correct structure folded at these concentrations it is necessary to take images of the samples. Figure 29 shows preliminary electron microscopy images (Quanta SEM in transmission mode) of the most promising lane of the Overhang (left) and Recessed (right) structures. The images show rectangular-like shapes of the expected 60 nm size for the structure; however, the edges of the structure are not well defined and many of the structures have aggregated together. This indicates that the structures may not have fully self-assembled, and more design optimization is likely required.

3.2 The Right-Left (LR) Design

Because of the extreme difference in folding between the barrel half with all 24 overhangs and the barrel half with all 24 recessions, a logical next step is to redesign halves with 12 overhangs and 12 recessions. To prevent costly reordering and time-consuming redesigning of the structure, it was possible to swap which staples belong to the Recessed and Overhang mixes and create a new

design. This new design is called the Right-Left barrel design. In it, the two otherwise identical half-barrels have one side of the plug and socket interface extended, and one side recessed. The structure is named after the side that is extended as shown in the following figure.

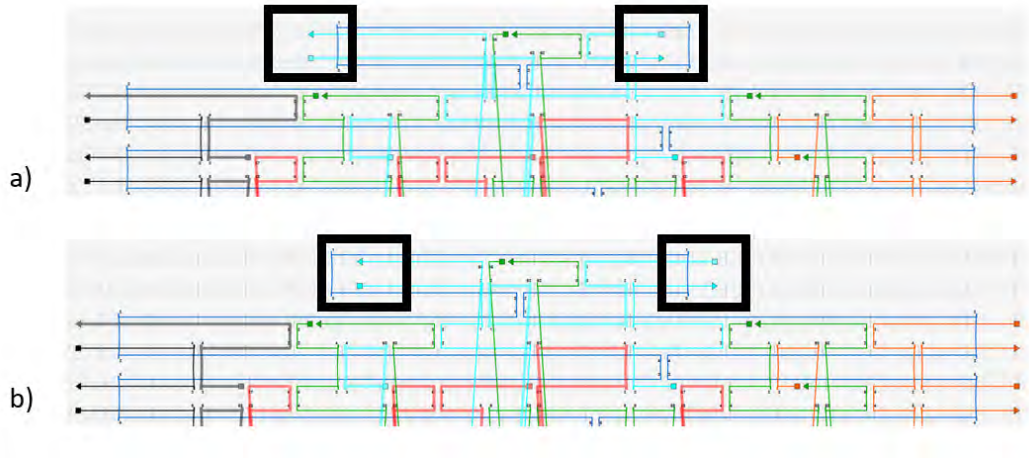

Figure 30: Top portion of the caDNAno schematics of the modified design. The Left design (a) includes four-base long extensions on the staples at the plug and socket interfaces on the left side and leaves the scaffold exposed on the right. The Right design (b) includes four bases extension on the Right with recessions on the left. The two designs are identical except for the four base pair changes at the plug-socket interface.

Because of the previous results with the Recessed-Overhang design, it seems that the number of sticky-end base-pair extensions decreases the ability of the design to properly self-assemble. Thus, it is hypothesized that reducing the number of staples with extensions from 24 to 12 will increase the robustness of the folding, while not jeopardizing the robustness of the other half-barrel that was changed from 0 to 12 extended staples. Conducting a salt scan in the 10 to 16 mM $MgCl_2$ range, the agarose gel shows the following results.

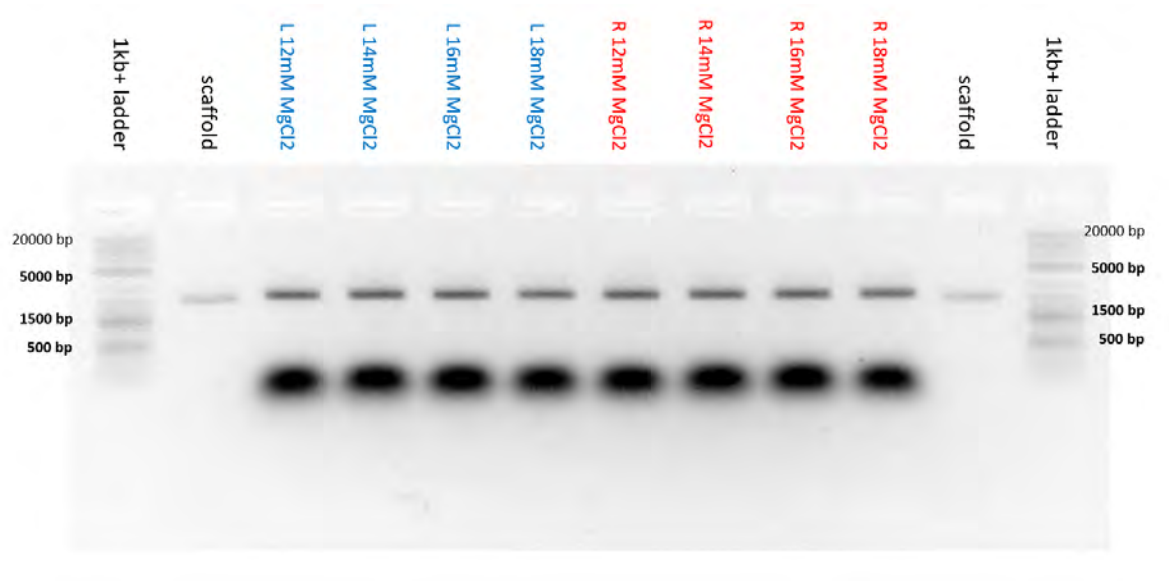

Figure 31: Agarose gel results of the $MgCl_2$ salt scan for the Right-Left design. The results show robust folding across the entire salt range.

The structure bands across all lanes in Figure 31 are sharp, dark, and travel just behind the scaffold lane. No aggregation or other bands are visible in the gel, indicating the structures are uniform in size. In addition, the quality of the bands does not vary much across the different salt concentrations, indicating that the self-assembly is robust. These results are very promising and are much improved from the Recessed-Overhang design. The structures need to be verified by electron microscopy images to prove that this design is well folded.

Figure 32 shows a TEM image of the Left barrel design extracted from the 14 mM $MgCl_2$ band. The visible structures are rectangular and approximately 50 nm long. These dark structures are suspected to be the two-dimensional projections of the target origami design. Three of the structures are shown in close-up views along with a schematic of the suspected orientation of the folded origami. The large grain size of the stain used for these images adds difficulty in determining if the structures are correctly folded. Future optimization of the stain technique can help mediate this noise and provide a better characterization of the structures.

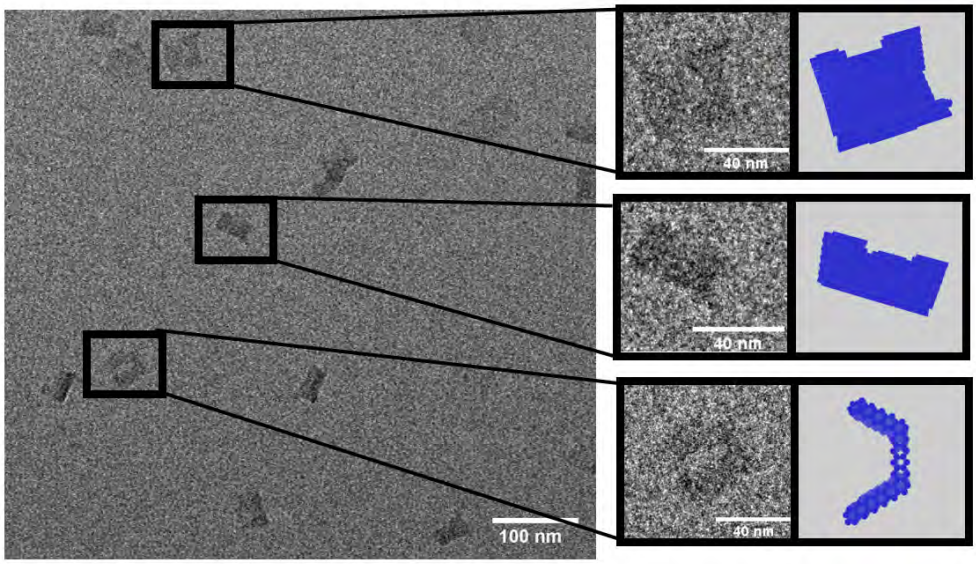

Figure 32: Transmission electron microscopy images for the Left barrel design. The small scale bars represent 40 nm. Close-ups are compared to a suspected orientation of the folded structure.

CHAPTER 4

CONCLUSION

In this project, two different DNA origami-based support structures are considered, the original design and the Echo-Echo. The original structure was designed from scratch using computer-aided design software, caDNAno. It was designed to capture a lipid bilayer enclosed by a nanodisc for membrane protein imaging in cryo-electron microscopy. The dimensions and ability of the design are to be built as two halves to allow for the capture of the nanodisc. The design also includes the ability to orient the nanodisc in one of three configurations (horizontal, vertical, and at forty-five degrees) to aid in the future 3D reconstruction studies of proteins. Once the design was drafted, both the scaffold routing and staple routing were optimized. Mid-seam routing with bundles was chosen as the most promising design for stability. Staples were carefully routed to allow for symmetry between the two half-barrels reducing cost by nearly 40%. Then the design was simulated using CanDo and oxDNA. Results from the simulation revealed a global twist which was eliminated by removing a base pair down the center of the design. Overall, the simulations showed promising stability.

The second support structure, originally designed by Huang and Saccà has similar capabilities including two halves and the ability to orient nanodiscs. It was first modified to allow for self-complementary dimers. To do this the original Echo-Nemesis design was converted into an Echo-Echo design, allowing the reuse of staples, and reducing cost in anticipation of continued studies in the development of protein imaging. Because this design had already been experimentally tested and optimized it was chosen to move on to the next phase of the project.

Experimental testing of the Echo-Echo design revealed difficulties in the self-assembly of the half-barrels. Even at the most promising salt concentration, 16 mM $MgCl_2$, one half of the barrel (the all-overhang half) was not folding satisfactorily. By hypothesis that the staple overhangs were affecting the folding, a slight modification changed the design from all-recessed, all-overhang to half-recessed, half-overhang barrels. This design called the Left-Right design, showed robust folding at a range of $MgCl_2$ concentrations from 12 to 18 mM.

The successful folding of the barrel is a promising result for the project. Further studies will optimize the dimerization of the two halves. The ability to build a nanoscale support structure

to capture nanodiscs is relevant to the development of an imaging method that could allow 3D reconstruction of essential membrane protein structures. Studies of membrane proteins are essential to both fundamental biology as well as the treatment of human disease.

CHAPTER 5

METHODS

5.1 Simulation Techniques

After the initial design of the barrel with caDNAno (Douglas, Marblestone, et al., 2009), computer simulation was used to predict the stability of the design.

5.1.1 CanDo

The first type of simulation used was a mechanical model of the 3D shape treating DNA helices as elastic rods with experimentally determined axial stretching, twisting, and bending stiffness (Kim et al., 2012). Invented by Do-Nyun Kim and validated by Hendrik Dietz, the software is now integrated into an online web server (http://cando-dna-origami.org/). A caDNAno file and simulation parameters were submitted to the web server. Values for average B-form DNA geometry were entered and default mechanical property values were used. The coarse resolution was used to obtain feedback for designs at each stage of development. Movies of thermal fluctuations and atomic models of the solution shape were also created by selecting these options on the online form. The results of the simulations were sent to the user by email.

5.1.2 oxDNA

The second type of simulation used was the coarse-grain model oxDNA by Oxford University (Doye et al., 2020). This simulation uses Monte Carlo and Molecular Dynamics and requires much longer simulation runtimes and GPU ability. Simulations were run both through the laptop command line and a server run by Arizona State University (oxDNA.org) (Poppleton, Romero, Mallya, Rovigatti, & Šulc, 2021). The simulation output is viewed using oxView (Bohlin et al., 2022).

For both using the server and running command line, special input files are required. caDNAno files are converted to a topology and configuration file using oxView's tacoDNA feature. The simulation parameters were indicated using a text file or through the oxDNA.org server's online form. The simulation includes two steps: relaxation and molecular dynamics. The relaxation step uses a modified backbone potential that allows the structure to relax from the strained input configuration without experiencing the large forces that should break it apart. Once the structure is relaxed into a more natural form, the molecular dynamics simulation can be run with the actual

backbone potential. The parameters used in the simulation are indicated in Table 5.1.2 and are in simulation units. A time step of 10^6 was used to do quick checks of the design while a time step of 10^9 provided a millisecond of simulation. For the final socket barrel design, relaxation took eight hours of computation time and simulation took three days.

Parameter	**Setting**
Simulation Type	DNA
Salt Concentration	1
Steps	1,000,000,000
Relax MC	100,000
Relax MD	10,000,000
Timestep	0.001
Backend Type	CUDA
Use Average Sequence Model	Yes
Print conf_interval	500,000
Print energy_every	50,000

Table 2: Simulation parameters used in oxDNA simulation of the two halves of the barrel. Values are in simulation units where 10^9 time steps correspond to a millisecond of real-time.

5.2 Folding of Origami

5.2.1 Staple Mixes

The staple oligonucleotides were ordered as well plates from Integrated DNA Technologies (IDT). The 227 staples were organized into four pools: core staples, nanodisc attachment staples, and two unique barrel half staples (ex. Recessed-only staples and Overhang-only staples). The staples were mixed into 30 μL pools with a concentration of 1 μM. The pools are stored frozen in 2 mL DNA LoBind polypropylene vials at -20°C.

5.2.2 Self-Assembly of Half-Barrels

Folding of the half-barrels was done by mixing the core, attachment, and respective half's unique staple mixes along with the p7560 scaffold. The staples were added in 10-fold molar excess with respect to the scaffold strand. The mixture was prepared in varying concentrations of $MgCl_2$,

typically 10 – 16 mM, and in a 1X TE buffer (5 mM TE, 1 mM EDTA). Then, 10 μL of the mixture was put into polymerase chain reaction (PCR) vials to be placed in a thermal cycler. The vials were incubated for 10 minutes at 65°C followed by 52°C for three hours. Samples were then stored at room temperature until further use.

5.3 Agarose Gel Electrophoresis

The gel was prepared by mixing 3 g of agarose powder, 2 mL of 1.2 Molar $MgCl_2$ solution, 20 mL of 5 X TBE buffer, and 150 mL of water in a 500 mL beaker. The beaker was then heated in the microwave until boiling. Then, 20 μL of SYBR-safe stain and water to replace any evaporation are added before the solution was poured into the gel mold. After at least 30 minutes, the gel was transferred to the electrophoresis chamber.

5.3.1 Loading and Imaging

Once the samples are ready for a gel run, the electrophoresis chamber was filled with gel running buffer (0.5 X TBE and 12 mM $MgCl_2$). Samples were mixed with Purple Gel Loading Dye from BioLabs at a one-to-six ratio and then carefully pipetted into the wells. In addition, a 1 kb plus ladder from Thermofisher Scientific and unfolded 7560 scaffolds were loaded into outside lanes as controls. The gel was run at 90 V for 3 hours at room temperature. After the run was complete the gel was imaged using a blue light transilluminator with an orange filter inside of the Fisher Scientific Ingenius 3 Manual Gel Documentation System. Images were processed using ImageJ. Samples of interest were manually excised from the gel for further analysis.

5.4 Electron Microscopy

Electron microscopy images were taken on the Quanta Scanning Electron Microscope (Quanta tSEM) in transmission mode and the 200 kV Tecnai G2 F20 transmission electron microscope (TEM) by Praneetha Sundar Prakash. The carbon-coated copper TEM grids used in both microscopes were prepared for each sample. First, the grids were cleaned using a plasma cleaner for 25 seconds. This cleaning allows the samples to stick to the surface of the grid. After cleaning, the grid was inverted onto a 10 – 20 μL drop of sample for five minutes. During this time, 2.5 μL of fresh 1 Molar sodium hydroxide was added to a 100 μL aliquot of 2% Uranyl Formate Stain to activate it. After sample incubation is complete excess liquid is gently pulled away from the grid

by placing a low-lint paper towel at the edge of the grid. Then the grid was inverted onto a 10 μL drop of Uranyl Formate for no more than 30 seconds. Again, the excess is removed using a paper towel. After this, the grid was washed with water twice by inverting it onto a 10 μL drop of Ultrapure water for one minute each wash. The grid was then allowed to air dry before storage in a grid box until imaging.

References

Aksel, T., Yu, Z., Cheng, Y., & Douglas, S. M. (2021, March). Molecular goniometers for single-particle cryo-electron microscopy of DNA-binding proteins. *Nature Biotechnology*, *39*(3), 378–386. Retrieved 2022-05-24, from https://www.nature.com/articles/s41587-020-0716-8 (Number: 3 Publisher: Nature Publishing Group) doi: 10.1038/s41587-020-0716-8

Bohlin, J., Matthies, M., Poppleton, E., Procyk, J., Mallya, A., Yan, H., & Šulc, P. (2022, August). Design and simulation of DNA, RNA and hybrid protein–nucleic acid nanostructures with oxView. *Nature Protocols*, *17*(8), 1762–1788. Retrieved 2023-03-04, from https://www.nature.com/articles/s41596-022-00688-5 (Number: 8 Publisher: Nature Publishing Group) doi: 10.1038/s41596-022-00688-5

Castro, C. E., Kilchherr, F., Kim, D. N., Shiao, E. L., Wauer, T., Wortmann, P., ... Dietz, H. (2011, January). A primer to scaffolded DNA origami. *NATURE METHODS*, *8*(3), 221–229. Retrieved 2023-01-11, from https://proxy.library.kent.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=edsbl&AN=RN286657850&site=eds-live&scope=site (Place: United States Publisher: Nature Publishing Group)

Chandrasekaran, A. R., & Zhuo, R. (2016, March). A 'tile' tale: Hierarchical self-assembly of DNA lattices. *Applied Materials Today*, *2*, 7–16. Retrieved 2023-01-09, from https://www.sciencedirect.com/science/article/pii/S2352940715300135 doi: 10.1016/j.apmt.2015.11.004

Chen, A., Majdinasab, E. J., Fiori, M. C., Liang, H., & Altenberg, G. A. (2020). Polymer-Encased Nanodiscs and Polymer Nanodiscs: New Platforms for Membrane Protein Research and Applications. *Frontiers in Bioengineering and Biotechnology*, *8*. Retrieved 2023-04-01, from https://www.frontiersin.org/articles/10.3389/fbioe.2020.598450

Dietz, H., Douglas, S. M., & Shih, W. M. (2009, August). Folding DNA into Twisted and Curved Nanoscale Shapes. *Science (New York, N.Y.)*, *325*(5941), 725–730. Retrieved 2023-01-12, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2737683/ doi: 10.1126/science

.1174251

Douglas, S. M., Dietz, H., Liedl, T., Högberg, B., Graf, F., & Shih, W. M. (2009, May). Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, *459*(7245), 414–418. Retrieved 2022-05-23, from https://www.nature.com/articles/nature08016 (Number: 7245 Publisher: Nature Publishing Group) doi: 10.1038/nature08016

Douglas, S. M., Marblestone, A. H., Teerapittayanon, S., Vazquez, A., Church, G. M., & Shih, W. M. (2009, August). Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Research*, *37*(15), 5001–5006. Retrieved 2022-05-23, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2731887/ doi: 10.1093/nar/gkp436

Doye, J. P. K., Fowler, H., Prešern, D., Bohlin, J., Rovigatti, L., Romano, F., ... Snodin, B. E. K. (2020). The oxDNA coarse-grained model as a tool to simulate DNA origami. *arXiv*, 17. Retrieved from https://arxiv.org/abs/2004.05052 doi: 10.48550/ARXIV.2004.05052

Dunn, K. E., Dannenberg, F., Ouldridge, T. E., Kwiatkowska, M., Turberfield, A. J., & Bath, J. (2015, September). Guiding the folding pathway of DNA origami. *Nature*, *525*(7567), 82–86. Retrieved 2022-06-22, from https://www.nature.com/articles/nature14860 (Number: 7567 Publisher: Nature Publishing Group) doi: 10.1038/nature14860

Gerling, T., Wagenbauer, K. F., Neuner, A. M., & Dietz, H. (2015, March). Dynamic DNA devices and assemblies formed by shape-complementary, non–base pairing 3D components. *Science*, *347*(6229), 1446–1452. Retrieved 2022-05-24, from https://www.science.org/doi/10.1126/science.aaa5372 (Publisher: American Association for the Advancement of Science) doi: 10.1126/science.aaa5372

Gür, F. N., Schwarz, F. W., Ye, J., Diez, S., & Schmidt, T. L. (2016, May). Toward Self-Assembled Plasmonic Devices: High-Yield Arrangement of Gold Nanoparticles on DNA Origami Templates. *ACS Nano*, *10*(5), 5374–5382. Retrieved 2022-05-20, from https://doi.org/10.1021/acsnano.6b01537 (Publisher: American Chemical Society) doi: 10.1021/acsnano.6b01537

Hedin, L. E., Illergård, K., & Elofsson, A. (2011, August). An Introduction to Membrane Proteins. *Journal of Proteome Research*, *10*(8), 3324–3331. Retrieved 2023-03-22, from https://doi.org/10.1021/pr200145a (Publisher: American Chemical Society) doi: 10.1021/pr200145a

Hong, F., Zhang, F., Liu, Y., & Yan, H. (2017, October). DNA Origami: Scaffolds for Creating

Higher Order Structures. *Chemical Reviews*, *117*(20), 12584–12640. Retrieved 2022-05-20, from https://doi.org/10.1021/acs.chemrev.6b00825 (Publisher: American Chemical Society) doi: 10.1021/acs.chemrev.6b00825

Huang, J., Gambietz, S., & Saccà, B. (2022). Self-Assembled Artificial DNA Nanocompartments and Their Bioapplications. *Small*, *n/a*(n/a), 2202253. Retrieved 2023-03-04, from https://onlinelibrary.wiley.com/doi/abs/10.1002/smll.202202253 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smll.202202253) doi: 10.1002/smll.202202253

Iric, K., Subramanian, M., Oertel, J., Agarwal, N. P., Matthies, M., Periole, X., ... Schmidt, T. L. (2018, October). DNA-encircled lipid bilayers. *Nanoscale*, *10*(39), 18463–18467. Retrieved 2022-09-20, from https://pubs.rsc.org/en/content/articlelanding/2018/nr/c8nr06505e (Publisher: The Royal Society of Chemistry) doi: 10.1039/C8NR06505E

Joshi, F. (2017). *DNA Origami Barcodes for Immunolabelling* (Unpublished master's thesis). Biotechnologisches Zentrum (BIOTEC) Technische Universit¨at Dresden Germany.

Kallenbach, N. R., Ma, R.-I., & Seeman, N. C. (1983, October). An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, *305*(5937), 829–831. Retrieved 2023-03-11, from https://www.nature.com/articles/305829a0 (Number: 5937 Publisher: Nature Publishing Group) doi: 10.1038/305829a0

Ke, Y., Bellot, G., Voigt, N. V., Fradkov, E., & Shih, W. M. (2012, August). Two design strategies for enhancement of multilayer–DNA-origami folding: underwinding for specific intercalator rescue and staple-break positioning. *Chemical science (Royal Society of Chemistry : 2010)*, *3*(8), 2587–2597. Retrieved 2023-01-09, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3957201/ doi: 10.1039/C2SC20446K

Kermani, A. A. (2021). A guide to membrane protein X-ray crystallography. *The FEBS Journal*, *288*(20), 5788–5804. Retrieved 2023-03-22, from https://onlinelibrary.wiley.com/doi/abs/10.1111/febs.15676 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/febs.15676) doi: 10.1111/febs.15676

Kim, D.-N., Kilchherr, F., Dietz, H., & Bathe, M. (2012, April). Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures. *Nucleic Acids Research*, *40*(7), 2862–2868. Retrieved 2023-01-11, from https://www.ncbi.nlm.nih.gov/pmc/articles/

PMC3326316/ doi: 10.1093/nar/gkr1173

Krishnan, S., Ziegler, D., Arnaut, V., Martin, T. G., Kapsner, K., Henneberg, K., ... Simmel, F. C. (2016, September). Molecular transport through large-diameter DNA nanopores. *Nature Communications*, *7*(1), 12787. Retrieved 2022-06-23, from https://www.nature.com/articles/ncomms12787 (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/ncomms12787

Luan, B., & Aksimentiev, A. (2008, November). DNA Attraction in Monovalent and Divalent Electrolytes. *Journal of the American Chemical Society*, *130*(47), 15754–15755. Retrieved 2023-04-01, from https://pubs.acs.org/doi/10.1021/ja804802u doi: 10.1021/ja804802u

Marenduzzo, D. (2018). *The Physics of DNA and Chromosomes*. IOP Publishing. Retrieved 2022-06-28, from http://iopscience.iop.org/book/978-0-7503-1602-6 doi: 10.1088/978-0-7503-1602-6

Martin, T. G., Bharat, T. A. M., Joerger, A. C., Bai, X.-c., Praetorius, F., Fersht, A. R., ... Scheres, S. H. W. (2016, November). Design of a molecular support for cryo-EM structure determination. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(47), E7456–E7463. Retrieved 2022-05-24, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5127339/ doi: 10.1073/pnas.1612720113

Nygaard, R., Kim, J., & Mancia, F. (2020, October). Cryo-electron microscopy analysis of small membrane proteins. *Current Opinion in Structural Biology*, *64*, 26–33. Retrieved 2023-03-10, from https://www.sciencedirect.com/science/article/pii/S0959440X20300774 doi: 10.1016/j.sbi.2020.05.009

Ouldridge, T. E., Louis, A. A., & Doye, J. P. K. (2011, February). Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *The Journal of Chemical Physics*, *134*(8), 085101. Retrieved 2023-03-10, from https://aip.scitation.org/doi/full/10.1063/1.3552946 (Publisher: American Institute of Physics) doi: 10.1063/1.3552946

Overington, J. P., Al-Lazikani, B., & Hopkins, A. L. (2006, December). How many drug targets are there? *Nature Reviews Drug Discovery*, *5*(12), 993–996. Retrieved 2023-03-10, from https://www.nature.com/articles/nrd2199 (Number: 12 Publisher: Nature Publishing Group) doi: 10.1038/nrd2199

Poppleton, E., Romero, R., Mallya, A., Rovigatti, L., & Šulc, P. (2021, July). OxDNA.org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures. *Nucleic Acids Research*, *49*(W1), W491–W498. Retrieved 2023-03-04, from https://doi.org/10.1093/nar/gkab324 doi: 10.1093/nar/gkab324

Pray, L. (2008). *Discovery of DNA Double Helix: Watson and Crick | Learn Science at Scitable.* Retrieved 2023-03-09, from http://www.nature.com/scitable/topicpage/discovery-of-dna-structure-and-function-watson-397 (Cg_cat: Discovery of DNA Structure and Function: Watson and Crick Cg_level: ESY Cg_topic: Discovery of DNA Structure and Function: Watson and Crick)

Rothemund, P. W. K. (2006, March). Folding DNA to create nanoscale shapes and patterns. *Nature*, *440*(7082), 297–302. Retrieved 2023-01-08, from https://www.nature.com/articles/nature04586 (Number: 7082 Publisher: Nature Publishing Group) doi: 10.1038/nature04586

Sedeh, R. S., Pan, K., Adendorff, M. R., Hallatschek, O., Bathe, K.-J., & Bathe, M. (2016, January). Computing Nonequilibrium Conformational Dynamics of Structured Nucleic Acid Assemblies. *Journal of Chemical Theory and Computation*, *12*(1), 261–273. Retrieved 2023-01-12, from https://doi.org/10.1021/acs.jctc.5b00965 (Publisher: American Chemical Society) doi: 10.1021/acs.jctc.5b00965

Selnihhin, D., & Andersen, E. S. (2015). Computer-Aided Design of DNA Origami Structures. In M. A. Marchisio (Ed.), *Computational Methods in Synthetic Biology* (pp. 23–44). New York, NY: Springer. Retrieved 2023-01-08, from https://doi.org/10.1007/978-1-4939-1878-2_2 doi: 10.1007/978-1-4939-1878-2_2

Sengar, A., Ouldridge, T. E., Henrich, O., Rovigatti, L., & Šulc, P. (2021). A Primer on the oxDNA Model of DNA: When to Use it, How to Simulate it and How to Interpret the Results. *Frontiers in Molecular Biosciences*, *8*. Retrieved 2022-07-07, from https://www.frontiersin.org/articles/10.3389/fmolb.2021.693710

Sprengel, A., Lill, P., Stegemann, P., Bravo-Rodriguez, K., Schöneweiß, E.-C., Merdanovic, M., ... Saccà, B. (2017, April). Tailored protein encapsulation into a DNA host using geometrically organized supramolecular interactions. *Nature Communications*, *8*(1), 14472. Retrieved 2022-05-20, from http://www.nature.com/articles/ncomms14472 doi: 10.1038/ncomms14472