BBA - Biomembranes 1862 (2020) 183191

Contents lists available at ScienceDirect

BBA - Biomembranes

journal homepage: www.elsevier.com/locate/bbamem

Solution structure and oligomeric state of the *E. coli*glycerol facilitator

Mary Hernando[a], George Orriss[a], Jacqueline Perodeau[b], Shixing Lei[a], Fraser G. Ferens[c], Trushar R. Patel[d], Jörg Stetefeld[a], Andrew J. Nieuwkoop[b], Joe D. O'Neil[a,*]

[a] *Department of Chemistry, University of Manitoba, Winnipeg, MB R3T 2N2, Canada*
[b] *Department of Chemistry and Chemical Biology, Rutgers School of Arts and Sciences, 123 Bevier Rd., Piscataway, NJ 08854, USA*
[c] *Department of Microbiology, University of Manitoba, Winnipeg, MB R3T 2N2, Canada*
[d] *Alberta RNA Research & Training Institute, Department of Chemistry & Biochemistry, Lethbridge University, 4401 University Drive, Lethbridge, Alberta T1K 3M4, Canada*

ARTICLE INFO

Keywords:
Glycerol facilitator
Membrane protein
Negative stain transmission electron microscopy
Size-exclusion chromatography multi-angle light scattering
Small angle X-ray scattering
Solid-state NMR spectroscopy

ABSTRACT

Protein dynamics at atomic resolution can provide deep insights into the biological activities of proteins and enzymes but they can also make structure and dynamics studies challenging. Despite their well-known biological and pharmaceutical importance, integral membrane protein structure and dynamics studies lag behind those of water-soluble proteins mainly owing to solubility problems that result upon their removal from the membrane. *Escherichia coli* glycerol facilitator (GF) is a member of the aquaglyceroporin family that allows for the highly selective passive diffusion of its substrate glycerol across the inner membrane of the bacterium. Previous molecular dynamics simulations and hydrogen-deuterium exchange studies suggested that protein dynamics play an important role in the passage of glycerol through the protein pore. With the aim of studying GF dynamics by solution and solid-state nuclear magnetic resonance (NMR) spectroscopy we optimized the expression of isotope-labelled GF and explored various solubilizing agents including detergents, osmolytes, amphipols, random heteropolymers, lipid nanodiscs, bicelles and other buffer additives to optimize the solubility and polydispersity of the protein. The GF protein is most stable and soluble in lauryl maltose neopentyl glycol (LMNG), where it exists in a tetramer-octamer equilibrium. The solution structures of the GF tetramer and octamer were determined by negative-stain transmission electron microscopy (TEM), size-exclusion chromatography small-angle X-ray scattering (SEC-SAXS) and solid-state magic-angle spinning NMR spectroscopy. Although NMR sample preparation still needs optimization for full structure and dynamics studies, negative stain TEM and SEC-SAXS revealed low-resolution structures of the detergent-solubilized tetramer and octamer particles. The non-native octamer appears to form from the association of the cytoplasmic faces of two tetramers, the interaction apparently mediated by their disordered N- and C-termini. This information may be useful in future studies directed at reducing the heterogeneity and self-association of the protein.

Abbreviations: A(0), X-ray scattering amplitude; CHS, cholesteryl hemisuccinate; CP, cross-polarization; DARR, dipolar assisted rotational resonance; DDM, dodecyl-β-D-maltoside; DHPC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; DLS, dynamic light scattering; DMNG, decyl maltose neopentyl glycol; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DNase, deoxyribonuclease I; p(r), pairwise distance distribution function; DSC, differential scanning calorimetry; D_{max}, maximum molecule length; DTAB, dodecyltrimethylammonium bromide; EDTA, ethylenediaminetetraacetic acid; GF, glycerol facilitator; HPLC, high performance liquid chromatography; HDX, hydrogen-deuterium exchange; IMAC, immobilized metal affinity chromatography; IMP, integral membrane protein; IPTG, isopropyl β-D-1-thiogalactopyranoside; I(s), X-ray scattering intensity; LDAO, lauryldimethylamine-N-oxide; LMNG, lauryl maltose neopentyl glycol; LS, light scattering; LMPC, lysomyristoylphosphatidylcholine; LB, lysogeny broth; MALS, multi-angle light scattering; MAS, magic angle spinning; MBP, maltose binding protein; MSP, membrane scaffold protein; MD, molecular dynamics simulations; MWCO, molecular weight cut off; NMR, nuclear magnetic resonance; OG, octyl-β-D-glucopyranoside; OD, optical density; PDB, protein data bank; PMSF, phenylmethyl-sulfonyl fluoride; PCA, protein conjugate analysis; R_g, radius of gyration; RMS, root mean squared; $R_{1\rho}$, rotating frame longitudinal relaxation; RF, radio frequency; RHP, random heteropolymer; RD, relaxation dispersion; RNase, ribonuclease A; s, SAXS scattering vector; SAXS, small-angle X-ray scattering; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SMALP, styrene maleic anhydride lipid particles; SEC, size-exclusion chromatography; SUV, small unilamellar lipid vesicle; TB, terrific broth; TDM, tetradecyl-β-D-maltoside; TEM, transmission electron microscopy; TEV, tobacco etch virus; TM, transmembrane; UV, ultraviolet; V_p, protein excluded volume

[*] Corresponding author.

E-mail address: joe.oneil@umanitoba.ca (J.D. O'Neil).

https://doi.org/10.1016/j.bbamem.2020.183191
Received 3 October 2019; Received in revised form 5 December 2019; Accepted 10 January 2020

1. Introduction

Protein dynamics at atomic resolution can provide deep insights into the biological activities of proteins [1–4]. For example, calmodulin adopts at least 40 different conformations permitting more than 300 eukaryotic proteins to bind different conformers in response to varying cellular calcium concentrations in the calcium second messenger system. This allows a single protein to translate a complex, time-varying signal into multiple possible outcomes [5,6]. Another well-studied example is the conformational landscape of apodihydrofolate reductase that appears to encompass all of the intermediates on the catalytic pathway suggesting that enzyme catalysis proceeds by way of substrate-induced perturbations of the conformational landscapes of enzymes [7]. However, the dynamic flexibility of proteins can also impede structural analysis. Conformational flexibility frequently prevents crystallization and structure determination by X-ray diffraction. Protein dynamics can also broaden Nuclear Magnetic Resonance (NMR) lines and impede resonance assignment needed for structure and dynamics measurements. And while cryo-electron microscopy may be able to provide structural information on a wider range of proteins than X-ray diffraction it does so at the cost of freezing the biologically important atomic dynamics by making measurements on frozen samples at −182 °C [8].

Approximately 50% of small molecule drugs target Integral Membrane Proteins (IMP) [9] yet only 1% of the unique structures deposited in the Protein Data Bank (PDB) are IMPs compared to an estimated 23% of the eukaryotic proteome that they constitute [10]. Structure and dynamics studies of IMPs lag behind those of water-soluble proteins because of low expression levels, toxicity upon overexpression [11], difficulties in the formation of well-ordered crystals, and solubilization problems owing to their large hydrophobic surfaces. Hundreds of different detergents are available to solubilize IMPs [12] but finding the best detergent for any individual IMP is still an empirical process [9]. Some studies suggest that although some IMPs are properly folded when solubilized by detergent micelles, they may not retain their full *in vivo* biological activity [13]. Detergent solubilization of IMPs frequently strips away natively interacting lipids, compromising the native structure and function owing to changes in lateral bilayer pressure and hydrophobic mismatch [14–19]. In addition to detergents, other agents available for solubilization of IMPs include styrene maleic anhydride lipid particles (SMALPs) [20], amphipathic polymers or amphipols [21], random heteropolymers [22], lipid nanodiscs [23], bicelles [24], lipid vesicles [25] and protein fusions [26].

The *E. coli* glycerol facilitator (GF) is a member of the Major Intrinsic Protein superfamily and the aquaglyceroporin family [27]. The atomic resolution crystal structure of GF at 2.2 Å [28] illuminated the physical and chemical features of the GF pore that account for the high selectivity of its substrate glycerol. GF is a tetramer with four independently working pores [28]. Each monomer is a small helix bundle that consists of six transmembrane (TM) helices and two half-helices [27]. Each half-helix contains a highly conserved Asn-Pro-Ala (NPA) motif on its N-terminus that plays an important role in the selectivity of the pore [29]. Molecular Dynamics (MD) simulations of GF in the presence and absence of glycerol in the pore suggested that the glycerol channel widens by about 0.5–1 Å in response to interactions with glycerol [30]. Interestingly, TM7, the half-helix closest to the periplasm, exhibits hydrogen-deuterium exchange (HDX) characteristics similar to the flexible loops in the protein suggesting high backbone flexibility in this region [31]. One of the conserved NPA motifs is found in the loop connecting it to the helix which suggests that TM7 dynamics may play a role in the diffusion of glycerol through the GF pore. Glycerol facilitator is also of interest as there are several known AQP-related human disease states, the aquaporinopathies [32–42].

Solid-state NMR spectroscopy has recently made important advances that permit the measurement of atomic dynamics in many non-crystallizable proteins including membrane proteins [43–46]. These advances are facilitating high-resolution structural information from a variety of non-crystalline biological and non-biological materials and, for example, led to the first high-resolution structures of amyloid fibrils [47]. The development of faster Magic Angle Spinning (MAS) rotors, up to 110 kHz, has permitted the detection of backbone and side-chain proton resonances in multidimensional experiments [48] increasing sensitivity and eliminating the need for line-narrowing by protein deuteration [45]. A suite of multidimensional pulse sequences has been developed that now permit rapid assignment of backbone and side-chain ^{1}H, ^{13}C and ^{15}N resonances [45,48–50] as well as the measurement of dynamics over timescales ranging from ps–ms [51,52]. Two recent examples include the use of solid-state NMR to discover the proton conduction and gating mechanism in the influenza M2 proton channel [53] and the conformational landscape of the prokaryotic KcsA K^{+} channel and how it is affected by pH and potassium ion concentrations [46,54]. One of the remaining challenges for the application of solid-state NMR to IMPs is in sample preparation. Solid-state NMR has been successfully applied to IMPs dissolved with the help of detergents, nanodiscs [55] and lipid vesicles [56]. In the case of lipid vesicles, finding the correct lipid:protein ratio to yield high quality ssNMR spectra is still a trial-and-error process. It appears that the lipid:protein ratio subtly affects the static protein conformational heterogeneity that gives rise to line-broadening in spectra [48,57].

With the goal of studying the glycerol facilitator structure and dynamics in a membrane-mimetic environment by solid-state NMR spectroscopy, in this contribution we report the overexpression of isotope-labelled GF using Studier's autoinduction system [58] and examine the behavior of GF in a wide variety of solubilizing agents by size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS). We also report the characterization of tetramer and octamer GF assemblies by SEC-MALS, negative-stain transmission electron microscopy (TEM), SEC-SAXS (small angle X-ray scattering) and solid-state NMR spectroscopy.

2. Materials and methods

A full list of materials and additional experimental details are provided in the Supplementary material.

2.1. Glycerol facilitator expression constructs and protein purification

For the present work, a new codon-optimized gene was created (GenScript, Nanjing, China) adding an N-terminal His_6 purification tag and TEV cleavage site to the GF gene. The gene was ligated into a pET28b(+) vector that was transformed into *E. coli* BL21(DE3) cells. The M_r of the expressed protein is 32,523 Da.

All culture media were adjusted to pH 7 before sterilization. 25 mL overnight *E. coli* pre-cultures were grown in lysogeny broth (LB) or terrific broth (TB) at 37 °C with shaking at 300 rpm. Final cultures (LS-5052 for preparations of ^{15}N and ^{13}C uniformly labelled GF or C-501750 for optimized incorporation of ^{13}C for uniform labelling [59]) were inoculated with 5 mL of overnight pre-culture per litre of final culture and shaken at 300 rpm for 48 h at 20 °C. Cells were harvested by centrifugation at 5000 ×*g* for 15 min at 4 °C [60]. Harvested cells were lysed by 3 rounds of freeze-thaw cycles and incubated with lysozyme (150 mg per litre of culture) for 30 min, and DNase (Deoxyribonuclease I) and RNase (Ribonuclease A) (1 mg each per litre of cells) for 10 min, with gentle stirring. The pH of the lysate was monitored between freeze-thaw cycles and maintained at approximately pH 7. Membrane fractions were prepared by layering the lysate over 70% w/v sucrose and ultracentrifuging at 22,000 rpm for 60 min at 4 °C. The supernatant was discarded, and the membrane fractions were stored on top of the sucrose layer and frozen at −80 °C. GF was extracted from frozen membrane fractions in 25 mM sodium phosphate buffer pH 7.4, containing 150 mM NaCl, 0.04% w/v sodium azide (Buffer A), and one of 30 mM DDM, 150 mM SDS, 20 mM DTAB, 20 mM OG, or 0.5 mM LMNG and 0.1 mM CHS for 3 h at room temperature with gentle stirring. Insoluble matter was removed by centrifugation at 12,000 rpm for 1 h at 18 °C. Detergent-solubilized GF was purified using immobilized metal (nickel) chelate chromatography (IMAC) [61]. For studies carried out on DDM-solubilized GF, the resin was first washed with 30 mL of Buffer A containing 0.5 mM DDM (Buffer 5–30) and then washed with the same buffer containing 60 mM imidazole (Buffer 10–30) until the A_{280} was below 0.01. The protein was eluted from the IMAC resin using Buffer 5–30 containing 250 mM imidazole (Buffer 5–250). Fractions of 1.5 mL were collected until the A_{280} was below 0.01.

SEC-MALS, negative stain TEM, and SAXS studies were done on GF extracted from the membrane with DDM as described above and on LMNG-solubilized protein. Exchange of LMNG for DDM was done on the IMAC resin, where buffers 5–30, 10–30, and 5–250 contained 0.5 mM LMNG instead of DDM. Buffer 5–30 was used to wash the resin and one resin volume of buffer was left in the column for 1 h at room temperature to allow for detergent exchange. In screening other detergents and solubilization agents except for LMNG, detergent exchange was done at the final elution step in the IMAC column by washing the resin with one column volume of detergent-free Buffer A, equilibrating the column with buffer 5–250 containing the appropriate detergent for 1 h followed by elution. For all GF purifications, protein yield was determined using absorbance at 280 nm and $\varepsilon = 38{,}305\ M^{-1}\ cm^{-1}$.

2.2. Size exclusion chromatography coupled with multi-angle light scattering

SEC-MALS measurements were carried out on the fraction eluted from the IMAC column containing the largest amount of protein. Samples were centrifuged for 5 min at 15,000 rpm to remove large aggregates. For SEC-MALS measurements samples were gel filtered using a Superdex 200 Increase 10/300 column on an AKTA Pure 25 chromatography system. Gel filtration was done with buffer 5–30 prefiltered through a 0.22 μm filter. Following elution from the column, the samples were analyzed in line by the UV absorbance detector of the AKTA Pure 25 chromatography system followed by the DAWN Heleos-II light scattering (LS) and OptiLab T-rEX refractive index detectors in series. Protein conjugate analysis (PCA) was performed using the ASTRA software (Wyatt Technology) to determine the protein mass of the protein-detergent complexes.

2.3. Fluorescence spectroscopy

Intrinsic GF fluorescence spectra were measured on a Jasco J810 spectropolarimeter/fluorometer. Sample temperature was maintained at 20 °C with a built-in Peltier device. GF samples were placed in a 1 cm path length quartz fluorescence cuvette and spectra were acquired from 300 to 450 nm using a 1 s response time, a 1 nm data pitch, a scanning speed of 10 nm/min, a bandwidth of 5 nm and an excitation wavelength of 280 nm. Samples were equilibrated at room temperature at each pH for 10 min before measuring the emission.

2.4. Negative stain transmission electron microscopy

DDM- and LMNG-solubilized GF samples were eluted from the IMAC column, pooled and dialysed overnight using a 3500 MWCO dialysis tubing (Fisher Scientific) to remove imidazole and then concentrated using an Amicon Ultra centrifugal filter (100 kDa MWCO, Millipore Sigma). Samples were concentrated to approximately 300 μL and gel filtered on the SEC-MALS system as described above with detergent-containing buffer to remove aggregates and assess the quality of the protein.

The negative stain electron microscopy was done at the Boulder Electron Microscopy Services in the Department of Molecular, Cellular, and Developmental Biology at the University of Colorado. Approximately 4 μL of GF diluted with SEC-MALS buffer 1:100 were adsorbed to a glow discharged continuous carbon copper grid (Electron Microscopy Sciences, Hatfield, PA). The sample was allowed to adhere for 30 s, blotted off, washed twice with water, then stained with 2% aqueous uranyl acetate for 30 s before blotting dry.

Image data were collected on an FEI Tecnai F20 FEG transmission electron microscope (FEI-Company, Eindhoven, The Netherlands) operating at 200 kV. Images were acquired at a magnification of 62,000 × and a defocus of −0.6 μm using a total dose of approximately 15 to 30 electrons/$Å^2$. Images were recorded binned by two on a 4 K × 4 K Gatan Ultrascan 895 CCD camera (Gatan Inc., Pleasanton, CA). With this camera at a microscope magnification of 62,000 ×, the resulting pixel size corresponds to 7.3 Å on the specimen. SerialEM software [62] was used to automate the data acquisition. Images were processed and 2D averages were computed using Relion [63].

2.5. Size-exclusion chromatography coupled with small-angle X-ray scattering

LMNG-solubilized GF samples were prepared similarly to DDM-solubilized GF negative stain TEM samples. After dialysis, samples were concentrated with an Amicon Ultra centrifugal filter (100 kDa MWCO, Millipore Sigma) and again gel filtered. After the second gel filtration step, fractions containing GF tetramers and octamers were collected based on the peak elution position, pooled, and concentrated using an Amicon Ultra centrifugal filter (100 kDa MWCO, Millipore Sigma) to approximately 5 mg/mL.

Several in-line SEC-SAXS data sets were collected at Beamline 21, Diamond Light Source (Didcot, Oxfordshire, UK). A 50 μL aliquot containing 5 mg/mL of LMNG-solubilized GF was passed over a 2.4 mL Superdex 200 Increase column equilibrated with buffer 5–30 containing 0.1 mM LMNG using an Agilent 1200 high performance liquid chromatography system. The eluate from the column flowed directly through the X-ray beam allowing for the serial collection of SAXS data of protein-detergent complexes separated by size. Further information about the system can be found here: https://www.diamond.ac.uk/Instruments/Soft-Condensed-Matter/small-angle/B21/description.html and in the Supplementary material.

Scatter [64] software was used to process the SAXS data; models of the GF homotetramer surrounded by the LMNG detergent shell were constructed using the GF crystal structure (PDB: 1FX8) [28] and the Memprot algorithm [65]. The GF octamer model was created using Pymol (https://www.schrodinger.com/pymol) to generate symmetry mates and LMNG was added using the CHARM-GUI [66] Micelle Builder [67]. The SAXS data were processed using the ATSAS [68] suite including Crysol [69], Gnom [70], and Sasref [71]. Details about how the LMNG-solubilized octamer was built are provided in the Supplementary material.

2.6. Solid-state NMR spectroscopy of LMNG-solubilized GF

After elution from the IMAC column, samples containing ^{15}N and ^{13}C isotope-labelled GF were pooled and dialysed overnight against a buffer containing 25 mM sodium phosphate pH 7.5, 150 mM sodium chloride, 0.04% sodium azide, and 0.1 mM LMNG to remove imidazole and reduce the concentration of LMNG. After dialysis, the pooled samples were concentrated to approximately 75 μL using a centrifugal concentrator and then dialysed against the same buffer containing 50% w/v PEG 2000. Dialysis was stopped at an approximate sample volume of 20 μL and a protein concentration of approximately 32 mg/mL.

Approximately 8 mg of ^{13}C-, ^{15}N-labelled LMNG-solubilized GF were packed into a 1.6 mm ssNMR rotor. All NMR spectra were acquired on a 600 MHz Bruker Avance III spectrometer (Bruker Biospin Ltd.) equipped with a 1.6 mm 40 kHz MAS (Magic Angle Spinning) probe (Phoenix NMR) at 0 °C. For all experiments, the ^{1}H 90 degree pulse was 160 kHz at 77.5 W, the ^{13}C 90 degree pulse was 125 kHz at 315 W, and the ^{15}N 90 degree pulse was 95 kHz at 400 W. The MAS rate was 13.333 kHz. The ^1H transmitter offset was 2 kHz. Spinal-64 ^1H decoupling at 110 kHz was used during all acquisition periods. 1D hN and hC Cross-Polarization (CP)-MAS spectra were obtained with 256 scans and processed with 60 Hz line broadening. The CP contact time for the hN experiment was 1.5 ms. The ^{15}N transmitter offset was 120 ppm and the SW was 39.7 kHz. The ^1H pulse was rectangular at 87 kHz. The ^{15}N pulse was a 90%–100% linear ramp about 71 kHz. The CP conditions for the hC experiment were a contact time of 1 ms, a rectangular 99 kHz ^1H pulse, and a ^{13}C 90%–100% tangent ramp centered at 100 kHz. The transmitter offset was 100 ppm and the SW was 81.5 kHz. The 2D DARR (Dipolar Assisted Rotational Resonance) [72] with a 10 ms mixing time were acquired with States-TPPI (Time-Proportional Phase Incrementation) in the indirect dimension, over 4.5 days. The transmitter frequency offset was 100 ppm in both directions; the spectra shown are a summation of four blocks of the experiment with 32 scans and 768 rows. The same hC CP conditions stated above were used. The SW was 81.5 kHz in t2 and 80 kHz in t1. The DARR field on the ^1H channel was 13.333 kHz. Acquisition times were 24.6 ms in t2 and 4.6 ms in t1. NMR spectra were processed using NMRPipe [73] with an exponential apodization function of 80 Hz and a 150 Hz Gaussian; the data were zero-filled in the direct dimension to 8192 points and in the indirect dimension to 4096 points. The first three points were corrected with linear prediction using 12 coefficients. In the DARR experiment, magnetization is transferred from protons to 1-bond-coupled ^{13}C and then transferred through space to nearby ^{13}C. At short mixing times nearby intra-residue connections are observed but for long mixing times inter-residue connections can be observed and used for resonance assignment. T_2 measurements were done with a pseudo-2D spin-echo experiment. The ^{13}C transmitter offset was centered in the Cα region, at 52 ppm. The refocusing pulse was a 637 μs Rsnob to decouple the neighboring COs and Cβs. There was a total of 8 rows acquired, each with 16 scans. The direct dimension was processed using the same parameters as the hC spectrum. T_2 relaxation times were fitted using Topspin Dynamics Center (Bruker Biopsin Ltd.) with an integral region of 23 to 81 ppm, and a calculated T_2 of 1.8 ms.

3. Results and discussion

3.1. Detergent-screening and optimization of GF solubility

The codon-optimized GF construct yielded approximately 20 mg/L of GF when cells were grown in LB or TB and extracted with DDM, and 10 mg/L when grown in M9 minimal medium, compared to approximately 5 mg/L and 1 mg/L respectively, for our previous construct [74]. Further increases in protein production were achieved by growing cells in Studier's autoinduction media (LS-5052 and C-750701) which reproducibly yielded 30 to 40 mg of GF per litre of cell culture when extracted with DDM. The utility of DDM for structural studies of IMPs is supported by the observation that 30% of the unique IMP structures deposited into the PDB during 2016–18 were of proteins extracted from membranes using DDM [10]. We also extracted *E. coli* membranes with 20 mM octyl-β-D-glucoside (OG) [28], 20 mM sodium dodecyl sulfate, 20 mM dodecyl trimethyl ammonium bromide, Styrene Maleic Acid 2 K co-polymer [20] or 0.5 mM LMNG and 0.1 mM CHS, but none of these worked as well as DDM.

A SEC-MALS elution profile of DDM-solubilized GF (Fig. 1S Supplementary material) showed that after elution from the IMAC column the protein exists as a mixture of about equal amounts of the native tetramer and an octameric species, with smaller amounts of some larger aggregates. Optimization of the NaCl concentration, pH and buffer concentration resulted in an elution profile with a single peak (Fig. 1A) that protein conjugate analysis (PCA) estimated to have a M_r of 150 kDa, reasonably close to the GF tetramer M_r of 130 kDa. Unfortunately, the protein slowly aggregates and 24 h later, the SEC-MALS elution profile of the same sample stored at room temperature showed two peaks corresponding to an octamer (calculated PCA M_r approximately 290 kDa) and the tetramer (Fig. 1B). We also explored the effects of storage temperature on GF aggregation and found that the protein exhibited a lower tendency to aggregate when stored at 4 °C compared to storage at room temperature and − 80 °C. In addition, diluting the protein 1/10 with buffer 5–30 without imidazole and 0.5 mM DDM and storage at 4 °C prolonged the period of time over which GF remained soluble. We also tested a variety of osmolytes and hydrotropes including 5% w/v glycerol [75], 0.5 M arginine [76], 5% and 10% v/v ethanol [77], 12 mM ATP [78], and 1 mM EDTA to determine if they could reduce GF self-association and prevent precipitation. None of these were effective as described in the Supplementary material.

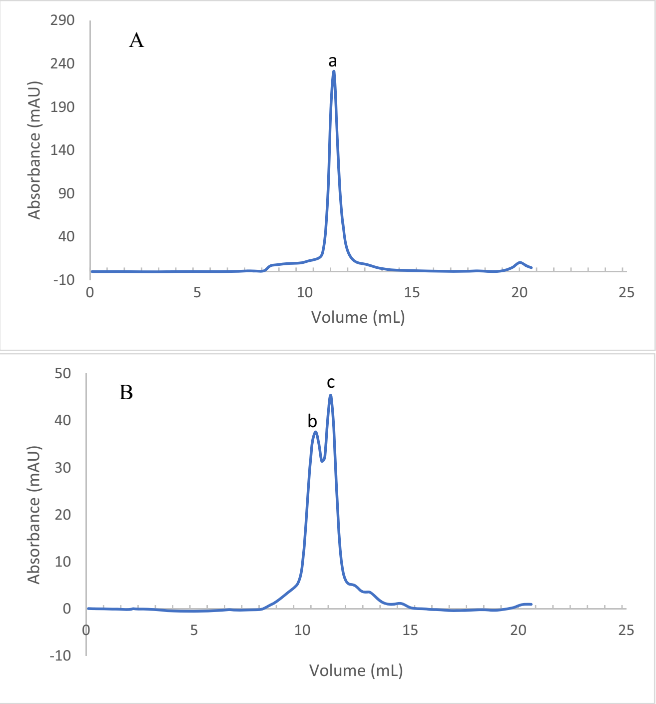

Fig. 1. Optimizing the solubility and polydispersity of DDM-solubilized GF. A) A SEC-MALS elution profile showing a pure homotetramer of DDM-solubilized GF where peak (a) corresponds to the GF homotetramer. B) The sample from (a) stored for 24 h at room temperature shows two peaks corresponding to the octamer (peak b) and tetramer (peak c). Protein oligomer sizes were determined by protein conjugate analysis.

Since GF is soluble at high concentrations in DDM but eventually precipitates, we explored exchanging the protein into other solubilizing agents following membrane extraction with DDM. Exchange was done while the protein was immobilized on the IMAC column and included the detergents lauryl dimethyl amine-N-oxide (LDAO), lysomyristoylphosphatidylcholine (LMPC), and decyl maltose neopentyl glycol (DMNG). As reported in the Supplementary material, none of these media provided environments in which GF was soluble for extended periods of time. Exchange of DDM-solubilized GF into LMNG while bound to the IMAC resin resulted in protein that elutes from a SEC-MALS column predominantly as a tetramer with a smaller octamer peak and a small amount of higher M_r aggregates (see Fig. 4S in Supplementary materials and Fig. 6A). LMNG-solubilized protein stored for 9 days at room temperature shows no visible precipitate, and the SEC-MALS elution profiles are remarkably similar to those of freshly purified GF showing a small increase in the fraction of octamer (Fig. 4S in Supplementary materials). LMNG-solubilized protein was by far the most stable of all our GF preparations and these results were highly reproducible over many repetitions. Differential scanning calorimetry showed that LMNG-solubilized GF has enhanced thermal stability with a melting point of 84 °C compared to a melting temperature of 76 °C in

DDM (Fig. 5S Supplementary material). Addition of CHS to the LMNG-solubilized GF resulted in increased polydispersity of the sample according to SEC-MALS (Fig. 6S, Supplementary materials).

Attempts were made to reconstitute DDM-solubilized GF into small, unilamellar lipid vesicles (SUVs) composed of native *E. coli* lipids, empty preformed DMPC/DHPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine/1,2-dihexanoyl-*sn*-glycero-3-phosphocholine) bicelles with q ratios of 4 and 0.5 [79], Amphipol A8-35 [80], MSP1E3D1 nanodiscs [81] and 36 kDa average molecular weight RHPs. RHPs are random polymers of four methacrylate-based monomers designed to match the surfaces of proteins and recently were shown to solubilize aquaporin produced by cell-free synthesis [22]. As described in more detail in the Supplementary material, none of these media in our hands provided environments in which GF was monodisperse and soluble for extended periods of time.

3.2. *Effect of pH on the oligomeric structure of GF*

Previous pH titrations of GF followed by SDS-PAGE appeared to show that the protein dissociates into monomers, dimers, and trimers at low pH but that a significant fraction of the tetramer could be refolded when the pH was restored [82]. However, in those experiments the oligomeric state of GF was monitored *via* SDS-PAGE and the high concentrations of SDS might have contributed to the observed oligomeric species. Here, we used intrinsic protein fluorescence to monitor the effects of pH on GF tertiary structure as well as the folding and refolding of the tetramer in the absence of denaturing SDS.

GF monomers have 5 tryptophan residues (Fig. 9S, Supplementary materials) that can be used to monitor GF structure. Although the protein was excited at 280 nm where both tyrosine and tryptophan absorb, tyrosine emission is most likely emitted by nearby tryptophans owing to resonance energy transfer. The Trp residues reside in various locations including near the cytoplasmic and periplasmic faces, buried in the interior of the protein, on the lipid-interacting surface and at the tetramer protein-protein interface. Fig. 2 shows the fluorescence intensity of DDM-solubilized GF at the starting, low, and ending pH's of the titration (pH 7.4, 3.7, and 7.5 respectively). At pH 7.4, the tryptophan fluorescence maximum is at 335 nm and contains shoulders at 319 nm and 309 nm. At the lowest pH (3.7), the shoulders have disappeared, the fluorescence intensity is enhanced, and the main peak has shifted to 341 nm. The red shift suggests that there are more tryptophan residues exposed to water possibly because the GF tetramer has dissociated at low pH. This interpretation agrees with the SDS-PAGE results that showed the protein tetramer dissociating below pH 6 [82].

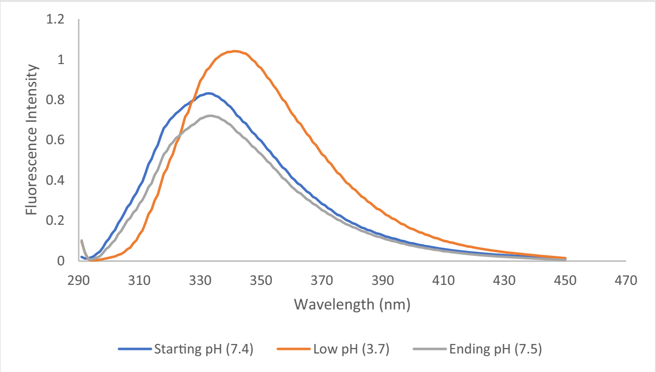

Fig. 2. A pH titration of GF tryptophan fluorescence with excitation at 280 nm. Spectra were measured at pH 7.4 (blue), pH 3.7 (orange) and following restoration to pH 7.5 (grey). Substantial reversibility of the titration upon restoration of the pH to 7.5 (grey), is indicated by the shoulders at 319 nm and 309 nm that reappear, and the main peak that returns to 335 nm. However, the intensity is quenched more than what was observed prior to acidification and increased Rayleigh scattering suggest the formation of aggregates.

When the sample is titrated back to the starting pH (7.5), the shoulders at 309 and 319 nm are restored and the main peak shifts back to 335 nm, but the fluorescence is slightly more quenched than it was at the start of the titration. This may be owing to incomplete restoration of the folded protein in which tyrosine fluorescence resonance energy transfer is reduced compared to the neutral pH folded state. These results support the earlier SDS-PAGE results [82] suggesting that much of the protein can be restored to its neutral pH tetrameric state following acidification and this observation might permit studies of the tetramer assembly process. Unfortunately, the protein is not very stable at acidic pH, slowly forming irreversible precipitates so that NMR studies on the putative monomer at pH 3.7 were not possible. During the titration the amount of Rayleigh scattering at the excitation wavelength increases suggesting increasing amounts of aggregated protein that could also explain the lowered fluorescence yield after restoration to neutral pH.

3.3. *Structural characterization of the GF octamer*

Under many conditions including in LMNG, GF exists as a mixture of octamers and tetramers so we next sought to characterize the octamer structure to identify the cause of the aggregation. Knowing which parts of the protein are interacting to form the octamer might enable the introduction of mutations that would eliminate the self-association. Examination of the GF crystal structure symmetry mates [28] reveals that the cytoplasm-facing surfaces of the tetramers face each other suggesting that the interface between the tetramers is occupied by the disordered N- and C-termini (Fig. 3). Although the disordered termini are not shown in the crystal structure, it is possible that they may be interacting to promote octamer formation. Negative stain Transmission Electron Microscopy (TEM), Small Angle X-ray Scattering (SAXS) and solid-state MAS NMR spectroscopy were employed to elucidate the solution structure of the octamer.

3.3.1. *Negative stain transmission electron microscopy*

Negative stain TEM was conducted on tetrameric DDM-solubilized GF and on tetrameric and octameric LMNG-solubilized GF. The micrographs collected for DDM-solubilized GF tetramers were of poor quality (Fig. 10S, Supplementary materials) and could not be used to calculate 2D class averages. LMNG-solubilized GF tetramers and octamers were separated by SEC-MALS and octamer and tetramer fractions were collected based on their elution positions. During SEC, the LMNG concentration was lowered from 0.5 mM to 0.01 mM to reduce to a

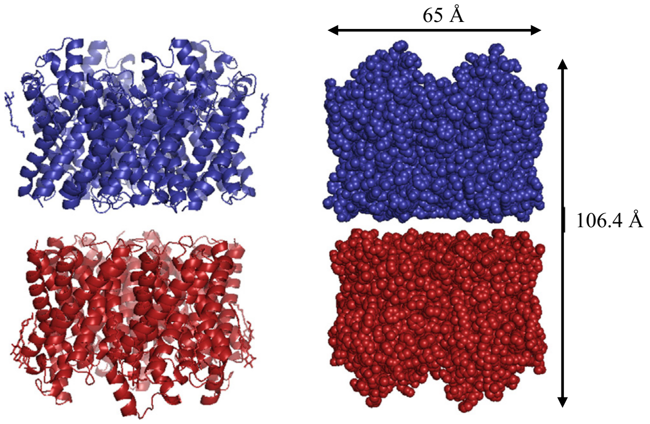

Fig. 3. The crystal structure of GF symmetry mates reveals the potential structure of the GF octamer. The cytoplasm-facing surfaces of the tetramers are at the interface. The disordered N- and C-termini (not seen in the crystal structure) protrude from the cytoplasm-facing surfaces and may occupy the interface, possibly causing the observed self-association of the tetramers. GF is shown in ribbon (left) and space-filling (right) models; blue and red denote the different tetramers (PDB: 1FX8 [28]).

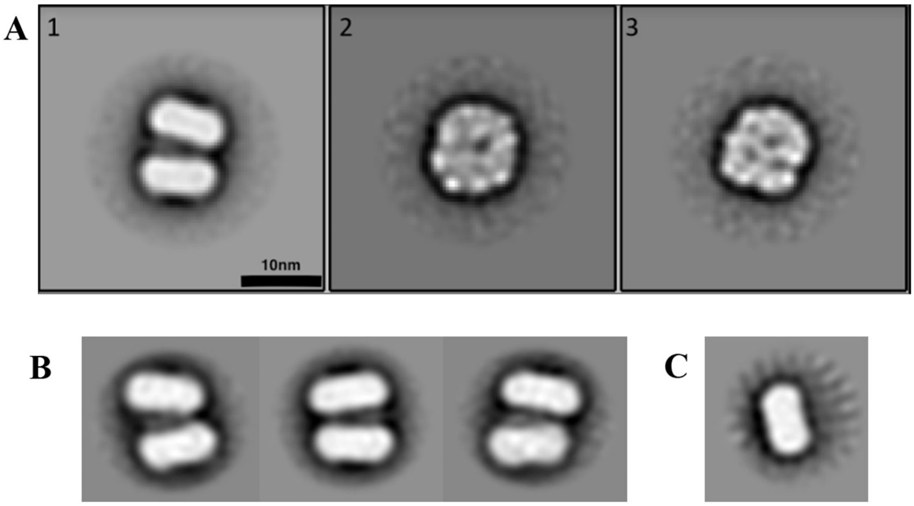

Fig. 4. 2D class averages of negative-stain TEM images of GF. A) 2D class averages of LMNG-solubilized GF octamer. The three most populated class averages are shown where (1) used 5899 particles to form the image, (2) used 5079 particles, and in (3), 4618 particles were used. The scale bar = 100 Å. B) Some less populated 2D class averages of octamers. C) A 2D class average image of a tetramer from the octamer preparation.

minimum the number of empty micelles. LMNG has a very low CMC and it is challenging to remove empty micelles [83]. Fig. 11S (Supplementary material) shows one of the 400 electron micrographs taken of the LMNG-solubilized GF octamer preparation. Both the LMNG-solubilized GF octamer and tetramer preparations contained a variety of structures that might reflect the presence of LMNG spherical micelles and worm-like structures that have been observed previously [83]. Empty LMNG micelles have been reported to have diameters of about 5 nm by negative stain TEM [83]. However, both sets of images also contain an abundance of particles that appear to be octamer and tetramer. This suggests that, following their separation by SEC, the octamers and tetramers re-equilibrate to form a mixture.

Fig. 4 shows the three largest classes of particle based on 2D averages from the octamer sample. Care should be taken in interpreting negative stain TEM images of membrane proteins in detergent owing to artifacts that can arise from images of the detergent even below the CMC [84]. For example, the "*halo*" that appears to surround each image in Fig. 4 is an artefact of image averaging that was used to create the class averages and is commonly observed [83]. However, when a crystal structure is known, the proteins can be confidently identified when the dimensions of the crystal structure agree with the observed dimensions of the negatively-stained TEM particles. Keeping in mind that negative stain TEM images the surfaces of particles, the largest 2D class image (Fig. 4A-1) shows what appears to be a GF octamer in which the tetramers are stacked in a manner that appears very similar to what is observed in the crystal structure (Fig. 3) with the main difference being that the orientations of the two tetramers are tilted with respect to each other. Fig. 4B shows three more sets of less abundant octamers obtained from the same data set. The longest distance across the octamer in the crystal structure is about 106 Å as shown in Fig. 3 which is in good agreement with the TEM images shown in Fig. 4. Also in good agreement are the widths of each tetramer, which in the crystal structure [28] is about 65 Å. Between the two tetramers, electron density is observed that likely arises from the disordered N- and C-termini. In the crystal structure [28], the two tetramers are separated by about 10 Å at their closest approach. However, the cytoplasmic surfaces contain a number of cavities that could also accommodate the disordered termini.

One difference between the X-ray and TEM structures is that the two tetramers are not parallel to one another in the TEM octamer images (Fig. 4A-1 and B). A possible explanation for the tilt of the structures is that the N-terminus of our GF construct contains an additional 20 residues from the His$_6$-tag. There is no mention of whether or not a purification tag was present on the GF protein that produced the X-ray diffraction structure in Fig. 3 [28]. Fig. 5A highlights the N- and C-terminal residues of each monomer in the crystal structure. In each case, the termini emerge near the periphery of the proteins, suggesting the possibility that they could escape the tetramer interface altogether if crystal packing forces require this. In contrast, the images in Fig. 4A-1 and B suggest that the lowest free energy state of the termini in the TEM structures is occupying the space between the tetramers and perhaps their interaction explains the octamer formation.

The second and third most populous 2D classes appear to be views of the octamer or tetramer looking down on the periplasmic or cytoplasmic surfaces. Fig. 5B presents a model of the GF tetramer surrounded by an LMNG detergent belt build as described in the Materials and methods section and discussed in more detail below. The diameter of the detergent belt circumscribing each tetramer is about 103 Å and this agrees with the dimensions of the images shown in Fig. 4A-2 and A-3. It is difficult to determine if the image surfaces are periplasmic or cytoplasmic. On the cytoplasmic surface one might expect to observe 4 "*towers*" protruding from the tetramer that project above the membrane. These are shown most clearly in the space-filling model in Fig. 3. The highest point of each tower is comprised by the C-terminus of helix-4, the N-terminus of helix-5, and a connecting loop. On the other hand, if the surfaces are cytoplasmic, one might expect to observe 4 amino and 4 carboxy termini projecting from the surfaces (Fig. 5). Bear in mind, however, that the resolution of the images is on the order of 17 Å and that the termini are expected to be disordered. Finally, that the SEC-purified octamer preparation contains tetrameric GF is evidenced by the 2D class average images shown in Fig. 4C. Both dimensions of this particle are in good agreement with the X-ray diffraction structure shown in Fig. 3 and the model in Fig. 5B.

3.3.2. Size-exclusion chromatography coupled with small angle X-ray scattering

SEC-SAXS experiments were performed on LMNG-solubilized GF to investigate the solution conformation of the self-associated octamers and to further explore the nature of the octameric assembly observed *via* negative stain EM. GF octamers and tetramers were separated before the SAXS measurements *via* in-line SEC. Fig. 6A shows the elution profile from the SEC column. Integrated scattering intensity (blue) is plotted against time intervals over which SAXS measurements were made (Frame Number). As an indication of particle size, the radii of gyration (R_g) calculated from the SAXS measurements of the eluted fractions are plotted in red. The elution profile shows high M_r aggregates, followed by octamer and tetramer peaks. The R_g values measured for the octamer and tetramer agree well with the particle sizes identified in the SEC-MALS measurements above.

A SAXS scattering curve averaged over several frames for the tetramer is shown in Fig. 6B. Fig. 6C shows a Guinier plot of the scattering data, $\ln[I(s)]$ *vs.* s^2, that can be used to provide a model-free estimate of the R_g of the protein from the slope at low angles. The Guinier

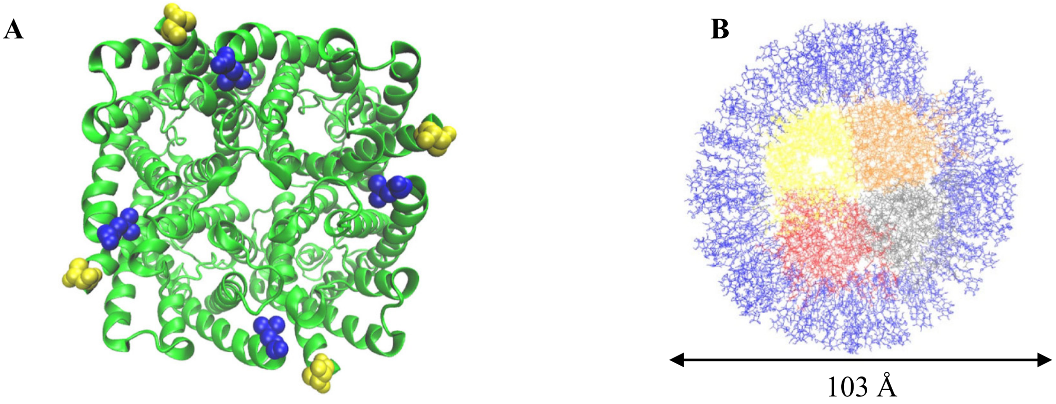

Fig. 5. A) The crystal structure of the GF tetramer (green ribbon) looking down on the periplasmic face with the N- and C-terminal residues of each monomer shown as yellow and blue spheres, respectively. The location of the terminal residues in the monomers suggests that the termini could escape the interface of the tetramers in the octamer. (PDB ID: 1FX8). B) A model of the GF tetramer surrounded by an LMNG detergent belt built as described in the Materials and methods section. The LMNG molecules are shown in blue and the monomers in the GF tetramer are individually coloured yellow, orange, red, and grey.

approximation assumes that for low scattering angles, $\left(s < \frac{1}{R_g}\right)$, $I(s) \sim I(0)\exp\left[-\frac{(sR_g)^2}{3}\right]$ where A(0) is the forward scattering amplitude at zero angle [85]. The R_g is the average root mean squared (rms) distance of the electrons to the center of density in the protein. For the GF tetramer, the scattering data gave an R_g of 49.9 Å. The intercept of the Guinier plot can also be used to obtain an estimate of the M_r of the protein from the known concentration and I(0); this was determined from Fig. 6C to be 368.2 kDa for tetrameric GF [86,87]. Subtracting the mass of the protein tetramer from this estimate and dividing by the molecular weight of an LMNG molecule suggests that about 236 LMNG molecules help solvate the GF tetramer.

The distance distribution function p(r) reports on the paired differences in electron separation in a molecule, is useful for detecting conformational changes and is also used to determine the maximum length of the molecule, D_{max}, and the R_g [85]. The D_{max} is the point on the distance distribution graph where the curve reaches the horizontal axis. For GF, the p(r) plot presented in Fig. 12S-A, (Supplementary materials) yielded a D_{max} of 97 Å and R_g of 55.6 Å. Integrating the scattering intensity provides an estimate of the protein excluded volume, V_p. For GF, the volume determined is 271,294 $Å^3$. By comparison, the volume of the GF tetramer crystal structure calculated using the Vadar 1.8 server [88] is 145,146.5 $Å^3$. This does not include the approximately 31,226 $Å^3$ that the disordered N- and C- termini occupy leaving 94,921 $Å^3$ or 35% of the total volume occupied by the detergent molecules.

Kratky plots are graphs of $s^2I(s)$ *vs.* s and are helpful in diagnosing disorder in a protein [85]. A distinct peak in a Kratky plot indicates a folded conformation and a plateau is diagnostic of a disordered chain. Normalized or dimensionless Kratky plots (normalized for protein

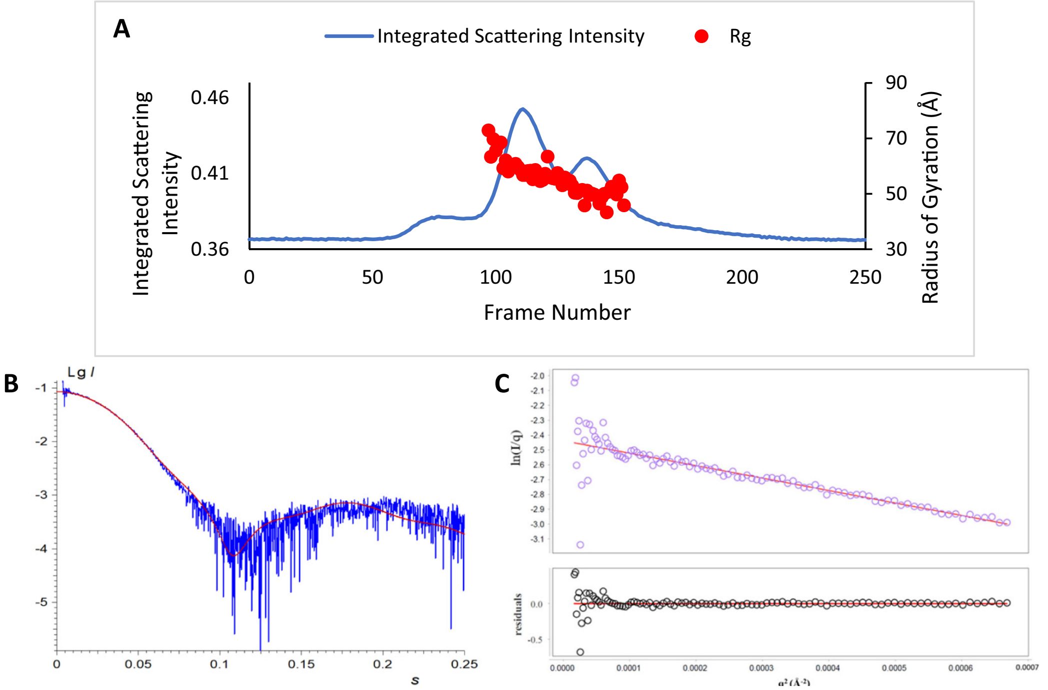

Fig. 6. LMNG-solubilized GF tetramer SAXS scattering data. A) The SEC elution profile of LMNG-solubilized GF at a concentration of 8 mg/mL showing peaks of tetramer, octamer, and a small amount of high M_r oligomers. B) The baseline-subtracted averaged tetramer SAXS scattering curve (blue) and the fit of the model calculated using the Memprot algorithm [65] (red). The y-axis is the SAXS scattering intensity I on a logarithmic scale. The x-axis, s in units of nm^{-1}, is the scattering vector where $s = 4\pi \sin\theta/\lambda$. C) A Guinier plot of the data in B.

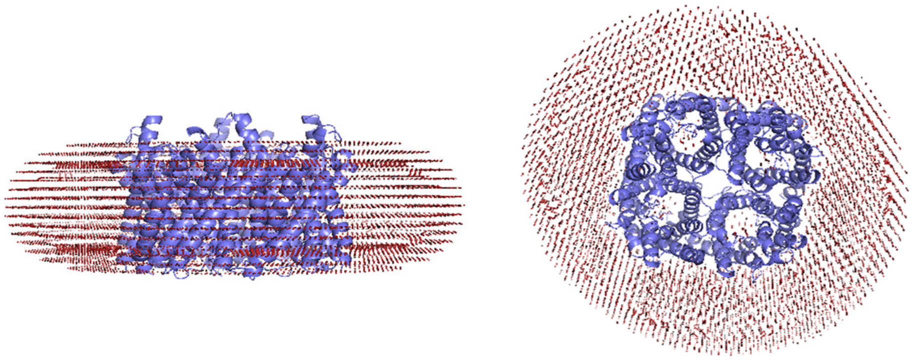

Fig. 7. A 3D reconstruction of DDM-solubilized GF using SEC-SAXS data for LMNG-solubilized protein. The GF homotetramer is shown in ribbon representation (blue) surrounded by a shell of detergent molecules (red) using the electron density for DDM. Modelled with the Memprot algorithm [65] and the GF crystal structure (PDB: 1FX8) [28].

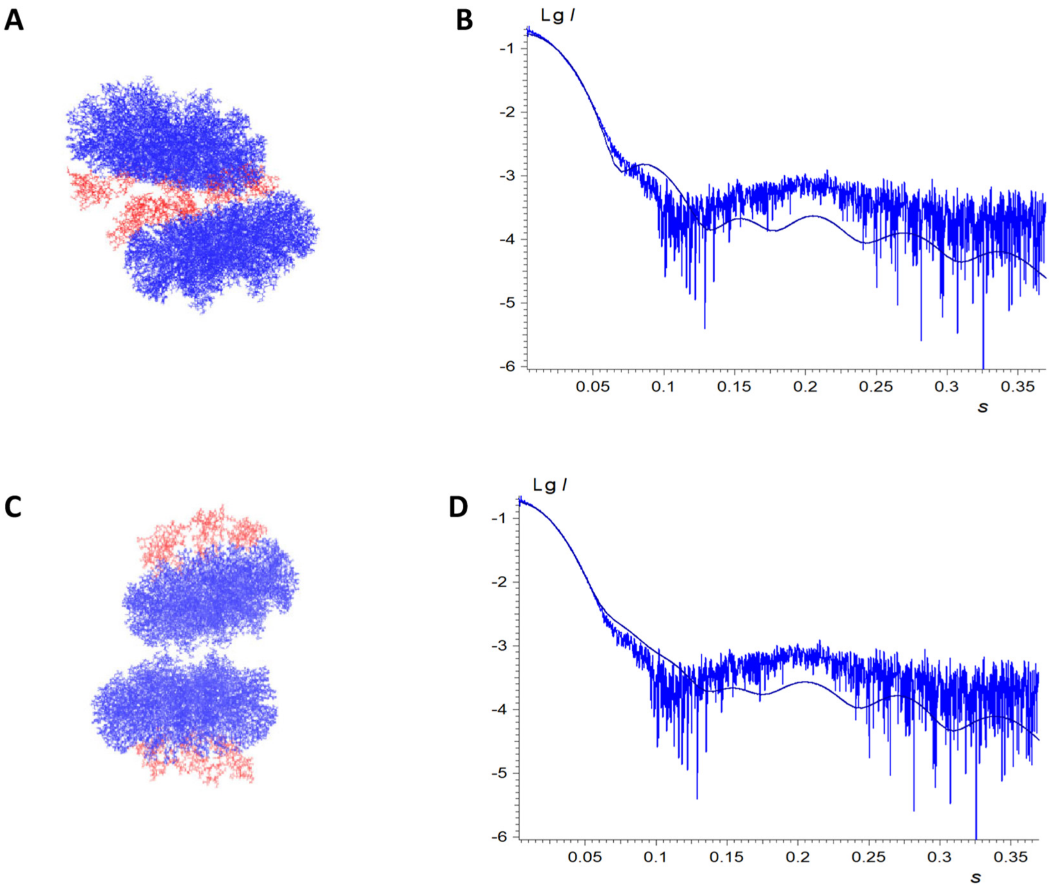

Fig. 8. GF octamer models and their SAXS scattering curves. A) A structure of the GF octamer with LMNG detergent belts around each tetramer (blue) with periplasmic surfaces facing each other and N- and C-termini (red) modelled as described in the Materials and methods section. B) The SAXS scattering curve (dark blue) of the structure in (A) generated with CRYSOL [69] and compared to the measured octamer SAXS curve (light blue). C) A model of the GF octamer with LMNG detergent belts around each tetramer with cytoplasmic surfaces facing each other and N- and C-termini (red) modelled as described in the Materials and methods section. D) The SAXS scattering curve (dark blue) of the structure in (C) generated with CRYSOL [69] and compared to the measured octamer SAXS curve (light blue).

concentration, volume and R_g) are graphs of $\frac{(sR_g)^2 I(s)}{I(0)}$ *vs.* sR_g and are used to detect changes in conformation or oligomerization. For globular proteins a peak should appear at about √3 [85]. Deviations from this value suggest flexibility in the biomolecules or asymmetry. The Kratky and normalized Kratky plots shown in Figs. 12S-B and S-C (Supplementary material) have peaks at 1.70 and confirm that the tetramer is globular and not very flexible.

The scattering data in Fig. 6B and the X-ray diffraction structure of the GF tetramer [28] were used to prepare a low-resolution, *ab initio* 3D reconstruction of the tetramer surrounded by a DDM detergent shell using the Memprot algorithm [65]. Memprot could not fit the protein to an LMNG shell owing to the lack of electron density information for LMNG [89]. The simulated SAXS scattering curve in Fig. 6B fits the measured data very well, the χ^2 difference between the two curves being 1.49. The good fit of the model to the scattering data is convincing evidence that the measured structure is a tetramer solubilized by a detergent shell. However, the model in Fig. 7 suffers some deficiencies. The particle is comprised of 482 DDM molecules (equivalent to 241 LMNG molecules) and is 130 Å in diameter, significantly larger than could be accounted for by a single layer of DDM (or LMNG) monomers 12–13 Å in length [90] as shown in Fig. 5B. Furthermore, the model in Fig. 7 is 30 Å wider than the octamer particles observed by TEM shown in Fig. 4. It's possible that additional detergent is present in the model owing to the missing N- and C-termini from the X-ray diffraction structure [28].

Analysis of the SEC-SAXS data from the octamer was challenging as

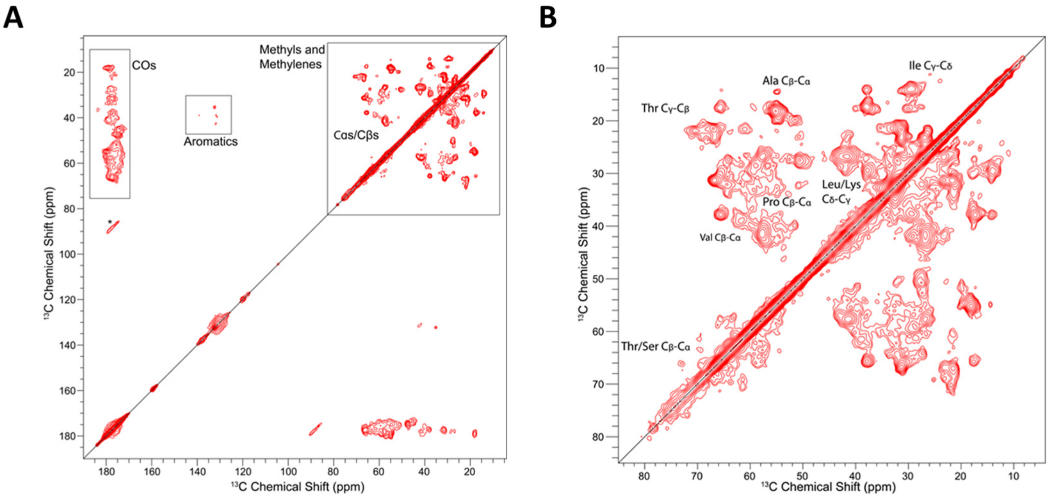

Fig. 9. 2D DARR spectrum of LMNG-solubilized GF. A) The full 2D DARR spectrum is shown with the ^{13}C-methyl and methylene, Cα and Cβ, and carbonyl regions of the spectrum identified. B) The high-field region of the spectrum shown in (A). Several crosspeaks in the spectrum are identified based on known unique chemical shift ranges.

the Memprot algorithm [65] is not configured to construct a detergent belt around each of the tetramers in the octamer. We therefore constructed models of the octamer as described in the Materials and methods section. A SAXS curve (plot of I(s) *vs.* s) of the octameric crystal structure [28] (Fig. 3) calculated using CRYSOL [69] deviated from the experimentally measured octamer SAXS curve with a χ^2 value of 115. Superimposing the DDM-tetramer model generated with Memprot (Fig. 7) on the octamer shown in Fig. 3 resulted in a model with an improved χ^2 goodness of fit of 24.7. A further improvement was made by building a model of the GF octamer with LMNG detergent belts around each tetramer and superimposing it on the high-resolution X-ray diffraction structure shown in Fig. 3. The SAXS scattering curve (Fig. 13S-B, Supplementary materials) generated from this model shows an improved fit with a χ^2 value of 9.1. Optimization of the orientations of each tetramer and its affiliated detergent belt using SASREF [91] resulted in a model (Fig. 13S-C Supplementary material) with an improved fit to the measured scattering curve (Fig. 13S-D, Supplementary material) with a χ^2 of 6.8 and resulting in a structure in which the tetramers and their detergent belts are tilted with respect to each other in a manner that is similar to that observed in the negative-stained TEM images shown in Fig. 4A and B. Note however, that the periplasmic faces of the tetramers are facing each other, not the cytoplasmic faces as in the crystal structure.

Further improvement in the fitting of the octamer SAXS data was achieved by adding models of the amino and carboxyl termini (see the Materials and methods section) to the best-fit octamer structures oriented either with their cytoplasmic surfaces interacting (Fig. 8A) or their periplasmic surfaces interacting (Fig. 8C). These structures gave the closest agreement with the measured SAXS scattering curves with χ^2 values of 3.9 and 2.4, respectively. Although the structure with the tetramers interacting through their cytoplasmic surfaces has the closest fit to the measured SAXS curve, the difference in the χ^2 values between the models in Fig. 8A and C are too small to allow for the discrimination between the models. The number of LMNG molecules surrounding each GF tetramer is 100, the maximum permitted by the Micelle Builder algorithm [67]. This compares to the 236 LMNG molecules that were estimated to associate with each tetramer based on the tetramer SAXS data above and the 482 DDMs (241 LMNG) fitted into the model in Fig. 7. Observation of the LMNG-belted octamer models (Fig. 5B) shows that some regions of the presumed membrane-interacting sites on the octamer do not interact with detergent, suggesting that the models could be improved further. However, overall there is good agreement between the experimentally-collected SAXS data and the SAXS curves generated from the octamer structures, suggesting that they are a valid depiction of the structure of the octamer in solution. Furthermore, they are in good agreement with the TEM images shown in Fig. 4.

3.3.3. *C-terminus-truncated GF*

To determine whether the disordered GF N- and/or C-termini contribute to the observed self-association of GF, a truncated GF construct was created with the last 19 amino acids, corresponding to the C-terminus, removed. Following the purification steps used for all our preparations, no truncated GF could be detected. This may indicate that the C-terminus is needed for proper insertion into the membrane or for proper folding of GF, or that truncation leads to proteolytic degradation [92,93].

3.3.4. *Solid-state NMR spectroscopy of LMNG-solubilized GF*

Optimized Studier medium C-750501 [94] yielded 30–40 mg of ^{15}N and ^{13}C-labelled GF per litre of cell culture. To prepare a sample for solid-state NMR spectroscopy the IMAC-eluted protein was gel filtered to remove high M_r soluble aggregates, empty micelles and free monomers. Fig. 14S in the Supplementary materials shows an elution profile of LMNG-solubilized GF concentrated to approximately 25 mg/mL prior to gel filtration, showing the presence of high levels of soluble aggregate. For the ssNMR experiments, the fractions containing octamer and tetramer (Fig. 14S Supplementary materials, peaks b and c) were pooled and concentrated using Amicon spin filters, achieving concentrations of 32 mg/mL in approximately 80 µL. Microdialysis with the use of PEG-2000 resulted in an NMR sample of approximately 13 mg of ^{13}C/^{15}N-labelled GF in 20 µL (650 mg/mL). For comparison, the SAXS samples were concentrated to approximately 8 mg/mL.

One-dimensional ^{15}N- (Fig. 15S-A Supplementary material) and ^{13}C-MAS NMR spectra (Fig. 15S-B, Supplementary material) of the above-described sample gave high signal intensity indicating that adequate amounts of ^{13}C and ^{15}N-labelled protein were present in the rotor. The observed chemical shift range shows that the protein is folded and predominantly helical.

Fig. 9A and B shows a 2D DARR (Dipolar Assisted Rotational Resonance) spectrum of LMNG-solubilized GF. Fig. 9A shows the full spectrum from the methyl region to the carbonyl region whereas Fig. 9B shows an expansion of the high-field region. Several resonances have been identified in Fig. 9B based on known unique ^{13}C chemical shift ranges in folded proteins [95]. For example, Ser and Thr Cβ resonances are always the most downfield shifted in proteins and Ala Cβ are always the most upfield shifted. Other unique shifts giving rise to identifiable peaks are the Pro-Cα-Cβ and Ile- Cγ-Cδ. The large number of poorly resolved Ala Cα-Cβ crosspeaks near (56 ppm, 18 ppm) is a clear indication that the protein is predominantly α-helical. Ser and Thr in α-helical structure exhibit highly similar Cα and Cβ resonances resulting in cross-peaks on or near the diagonal and this observation also confirms that the GF NMR sample is a folded helix bundle.

The average ^{13}C T_2 relaxation time of the upfield region of the spectrum from 23 to 81 ppm was measured by 1D NMR as described in the Materials and methods section yielding a T_2 of 1.79 ms with a standard deviation of 0.04. The relaxation decay is shown in Fig. 16S of the Supplementary materials. This rapid transverse relaxation precluded the measurement of 3D spectra necessary for observation and assignment of all the resonances in the protein.

The short T_2's suggest the resonances are exchange-broadened and this might arise from several sources. Although a tetramer-octamer equilibrium could broaden many of the resonances at the interface between the tetramers, it seems unlikely that this would affect the entire protein. Another possibility is that the protein is undergoing a conformational exchange in the ms-μs timescale that significantly changes the chemical shifts of most of the resonances in the protein. This would be somewhat surprising because no such conformational change has been reported in GF or in any other aquaporins. Furthermore, it is unlikely that the protein would retain its exquisite substrate specificity if it were to undergo very large conformational fluctuations, because the substrate-specificity clearly depends, to a large extent, on the pore maintaining its size within very narrow limits [28]. However, it is possible that the detergent environment has led to conformational flexibility in the protein that is not present in the native bacterial membrane. In addition, it's also possible that detergent exchange could lead to line-broadening of many resonances in the protein, especially those near the surface and interacting with the protein. Besides all these possible sources of dynamic exchange, it's also possible that the NMR resonances in the protein dissolved in LMNG at high concentrations are broadened owing to static heterogeneity. In bilayer-dissolved IMP's the lipid:protein ratio can affect the static protein conformational heterogeneity and lead to line-broadened spectra [48,57]. This problem has been overcome in some cases by careful exploration of lipid:protein ratios to discover the right combination that gives rise to one predominant conformation. Perhaps membrane proteins, especially those that do not bind a prosthetic group, are more conformationally flexible than we realize.

4. Conclusions

We report an expression and purification system using Studier's media [58] that enables the purification of 30–40 mg of ^{15}N, ^{13}C-GF per litre of labelling medium, more than adequate for pursuing solid-state NMR studies. Although dodecyl maltoside is by far the most effective agent for extracting GF from the *E. coli* inner membrane the protein tends to self-associate in DDM and in all other solubilizing media tested. Our work also shows that LMNG provides the most stable and least polydisperse preparation of GF. In addition, amphipol A8–35 [21] and the recently designed RHP's [22] look very promising as they keep the protein soluble, though polydisperse, for more than 1 week without any visible precipitation.

A major problem for studying GF by NMR spectroscopy is that the protein aggregates at high concentrations. Negative stain TEM images of GF in LMNG show that native tetramers and non-native octamers are the predominant species present at intermediate concentrations and that the tetramers self-associate through their disordered N- and C-termini and are tilted with respect to each other, not parallel as in the crystal structure. Models of LMNG-solubilized octamers in which the tetramers are tilted provide the best fits to the SEC-SAXS octamer data supporting the interpretation of the TEM images. It has not been possible to remove the N-terminal His_6 purification tag or express the protein without its C-terminus to test whether octamer formation is mediated by the disordered termini. Cleavage of affinity tags on membrane proteins often fail when the tag is close to the membrane-spanning domain of the protein [96]. But even if octamer formation could be prevented this might not prevent the formation of the higher molecular weight aggregates that form at very high protein concentrations. The large aggregates must form using other protein surfaces than those provided by the termini. This suggests that for the preparation of highly concentrated GF protein for ssNMR spectroscopy studies that lipid bilayer preparations such as bicelles [79] and vesicles [48,57] are the most likely to provide a native-like environment that reduces protein aggregation and polydispersity, and improves the quality of ssNMR spectra. Such preparations would open up the possibility of determining the role of atomic dynamics in the biological activity of GF.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank The Natural Sciences and Engineering Research Council of Canada and the Canada Foundation for Innovation for funding this work. TRP and JS acknowledge the Canada Research Chairs program. The authors are grateful also to Professor Andy Hoenger, Cynthia Lee Page and Garry P. Morgan of the Department of Molecular, Cellular, and Developmental Biology at the University of Colorado, Boulder, CO for assisting with the electron microscopy work, Vu To of the Department of Chemistry at the University of Manitoba for assistance with solution NMR measurements and Ewan McRae and Matthew McDougall of the Department of Chemistry, University of Manitoba for assistance with the SEC-SAXS measurements. We also thank the Diamond Light Source for allowing access to the B21BioSAXS beamline and Professor Ting Xu and Zhiyuan Ruan of the Department of Material Sciences and Engineering, University of California, Berkeley for providing us with samples of their RHP.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbamem.2020.183191.

References

[1] K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins, Nature. 450 (2007) 964–972, https://doi.org/10.1038/nature06522.
[2] M. Orozco, A theoretical view of protein dynamics, Chem. Soc. Rev. 43 (2014) 5051–5066, https://doi.org/10.1039/c3cs60474h.
[3] S.V. Buldyrev, S.-H. Chen, C. Corsaro, P. Baglioni, F. Mallamace, D. Mallamace, C. Vasi, S. Vasi, H.E. Stanley, Energy landscape in protein folding and unfolding, Proc. Natl. Acad. Sci. 113 (2016) 3159–3163, https://doi.org/10.1073/pnas.1524864113.
[4] G. Krainer, S. Keller, M. Schlierf, Structural dynamics of membrane-protein folding from single-molecule FRET, Curr. Opin. Struct. Biol. (2019), https://doi.org/10.1016/j.sbi.2019.05.025.
[5] S. Vetter, E. Leclerc, Novel aspects of calmodulin target recognition and activation, Eur. J. Biochem. (2003) 404–414.
[6] R. Liddington, A molecular full nelson, Nature. (2002) 373–374.
[7] K. Arora, C.L. Brooks, Multiple intermediates, diverse conformations, and cooperative conformational changes underlie the catalytic hydride transfer reaction of dihydrofolate reductase, Top. Curr. Chem. (2013), https://doi.org/10.1007/128_2012_408.
[8] T.R.D. Costa, A. Ignatiou, E.V. Orlova, Structural analysis of protein complexes by cryo electron microscopy, Methods Mol. Biol. (2017), https://doi.org/10.1007/978-1-4939-7033-9_28.
[9] H. Yin, A.D. Flynn, Drugging membrane protein interactions, Annu. Rev. Biomed. Eng. 18 (2016) 51–76, https://doi.org/10.1146/annurev-bioeng-092115-025322.
[10] S. White, Membrane proteins of known structure., (n.d.). http://blanco.biomol.uci.edu/mpstruc/.

[11] S. Wagner, L. Baars, A.J. Ytterberg, A. Klussmeier, C.S. Wagner, O. Nord, P.-Å. Nygren, K.J. van Wijk, J.-W. de Gier, Consequences of membrane protein overexpression in Escherichia coli, Mol. Cell. Proteomics 6 (2007) 1527–1550, https://doi.org/10.1074/mcp.m600431-mcp200.

[12] J.-L. Popot, Membrane proteins in aqueous solutions, 2018. doi:https://doi.org/10.1007/978-3-319-73148-3.

[13] T. Caldwell, S. Baoukina, A. Brock, R. Oliver, K. Root, J. Kreuger, K. Glover, D. Tieleman, L. Columbus, Low-q bicelles are mixed micelles, J. Phys. Chem. Lett. 9 (2018) 4469–4473.

[14] A. Stetsenko, A. Guskov, An overview of the top ten detergents used for membrane protein crystallization, Crystals. 7 (2017) 197, https://doi.org/10.3390/cryst7070197.

[15] J. Kern, A. Guskov, Lipids in photosystem II: multifunctional cofactors, J. Photochem. Photobiol. B Biol. 104 (2011) 19–34, https://doi.org/10.1016/j.jphotobiol.2011.02.025.

[16] C. Hunte, S. Richers, Lipids and membrane protein structures, Curr. Opin. Struct. Biol. 18 (2008) 406–411, https://doi.org/10.1016/j.sbi.2008.03.008.

[17] M.J. Lemieux, J. Song, M.J. Kim, Y. Huang, A. Villa, M. Auer, X.-D. Li, D.-N. Wang, Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily, Protein Sci. 12 (2003) 2748–2756, https://doi.org/10.1110/ps.03276603.

[18] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G.F. Rasmussen, F.S. Thian, T.S. Kobilka, H.-J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, R.C. Stevens, High-resolution crystal structure of an engineered human $β$$_2$-adrenergic G protein–coupled receptor, Science (80-.) 318 (2007) 1258–1265, https://doi.org/10.1126/science.1150577.

[19] A. Singh, V. Upadhyay, A. Upadhyay, S. Singh, A. Panda, Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process, Microb. Cell Factories 14 (2015).

[20] S.C. Lee, T.J. Knowles, V.L.G. Postis, M. Jamshad, R.A. Parslow, Y.P. Lin, A. Goldman, P. Sridhar, M. Overduin, S.P. Muench, T.R. Dafforn, A method for detergent-free isolation of membrane proteins in their local lipid environment, Nat. Protoc. 11 (2016) 1149–1162, https://doi.org/10.1038/nprot.2016.070.

[21] S. Elter, T. Raschle, S. Arens, A. Viegas, V. Gelev, M. Etzkorn, G. Wagner, The use of amphipols for NMR structural characterization of 7-TM proteins, J. Membr. Biol. 247 (2014) 957–964, https://doi.org/10.1007/s00232-014-9669-5.

[22] B. Panganiban, B. Qiao, T. Jiang, C. Delre, M.M. Obadia, T.D. Nguyen, A.A.A. Smith, A. Hall, I. Sit, M.G. Crosby, P.B. Dennis, E. Drockenmuller, M. Olvera, D. Cruz, T. Xu, Random heteropolymers preserve protein function in foreign environments, Science (80-.) 1243 (2018) 1239–1243, https://doi.org/10.1126/science.aao0335.

[23] T.K. Ritchie, Y.V. Grinkova, T.H. Bayburt, I.G. Denisov, J.K. Zolnerciks, W.M. Atkins, S.G. Sligar, Chapter 11 Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs, Elsevier Masson SAS, 2009, https://doi.org/10.1016/S0076-6879(09)64011-8.

[24] C.R. Sanders, G.C. Landis, Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies, Biochemistry. 34 (1995) 4030–4040, https://doi.org/10.1021/bi00012a022.

[25] A. Akbarzadeh, R. Rezaei-Sadabady, S. Davaran, S.W. Joo, N. Zarghami, Y. Hanifehpour, M. Samiei, M. Kouhi, K. Nejati-Koshki, Liposome: classification, preparation, and applications, Nanoscale Res. Lett. 8 (2013) 1–8, https://doi.org/10.1186/1556-276X-8-102.

[26] D. Mizrachi, M. Robinson, G. Ren, N. Ke, M. Berkmen, A water-soluble DsbB variant that catalyzes disulfide bond formation in vivo, Nat. Chem. Biol. 13 (2017) 1022–1028, https://doi.org/10.1038/nchembio.2409.A.

[27] N. Klein, J. Neumann, J.D. O'Neil, D. Schneider, Folding and stability of the aquaglyceroporin GlpF: Implications for human Aqua (glycero) porin diseases, Biochim. Biophys. Acta (BBA)-Biomembranes 1848 (2014) 622–633 http://www.sciencedirect.com/science/article/pii/S0005273614004131.

[28] D. Fu, A. Libson, L.J. Miercke, C. Weitzman, P. Nollert, J. Krucinski, R.M. Stroud, Structure of a glycerol-conducting channel and the basis for its selectivity, Science (80-.) 290 (2000) 481–486 (doi:8914 [pii]).

[29] J. Hub, H. Grubmuller, B. de Groot, Dynamic and energetics of permeation through aquaporins. What do we learn from molecular dynamics simulations? Handb. Exp. Pharmacol. 190 (2009) 57–76.

[30] M. Jensen, S. Park, E. Tajkhorshid, K. Schulten, Energetics of glycerol conduction through aquaglyceroporin GlpF, PNAS. 99 (2002) 6731–6736.

[31] Y. Pan, H. Piyadasa, J.D.O. Neil, L. Konermann, E. Al, Conformational dynamics of a membrane transport protein probed by H/D exchange and covalent labeling: the glycerol facilitator, J. Mol. Biol. 416 (2012) 400–413, https://doi.org/10.1016/j.jmb.2011.12.052.

[32] T. Ma, M. Hara, R. Sougrat, J. Verbavatz, A. Verkman, Impaired stratum corneum hydration in mice lacking epidermal water channel aquaporin-3, J. Biol. Chem. 277 (2002) 17147–17153.

[33] D. Ng, B. Poulsen, C. Deber, Membrane protein misassembly in disease, Biochim. Biophys. Acta 1818 (2012) 1115–1122.

[34] J. Molnar, G. Szakacs, G. Tusnady, Characterization of disease-associated mutations in human transmembrane proteins, PLoS One 11 (2016).

[35] Y. Noda, E. Sohara, E. Ohta, S. Sasaki, Aquaporins in kidney pathophysiology, Nat. Rev. Nephrol. 6 (2010) 168–178.

[36] L. King, S. Nielsen, P. Agre, R. Brown, Decreased pulmonary cascular permeability in aquaporin-1-null humans, PNAS. 99 (2002) 1059–1063.

[37] V. Berry, P. Francis, S. Kaushal, A. Moore, S. Bhattacharya, Missense mutations in MIP underlie autosomal dominant "polymorphic" and lamellar cataracts linked to 12q, Nat. Genet. 25 (2000) 15–17.

[38] A. Chepelinsky, Structural function of MIP/aquaporin 0 in the eye lens; genetic defects lead to congenital inherited cataracts, Handb. Exp. Pharmacol. 190 (2009) 265–297.

[39] D. Bichet, Hereditary polyuric disorders: new concepts and differential diagnosis, Semin. Nephrol. 26 (2006) 224–233.

[40] P. Nedvetsky, G. Tamma, S. Beulshausen, G. Valenti, W. Rosenthal, E. Klussman, Regulation of aquaporin-2 trafficking, Handb. Exp. Pharmacol. 190 (2009) 133–157.

[41] V. Lennon, T. Kryzer, S. Pittock, A. Verkman, S. Hinson, IgG marker of optic-spinal multiple sclerosis binds to the aquaporin-4 water channel, J. Exp. Med. 202 (2005) 473–477.

[42] D. Wingerchuk, V. Lennon, C. Lucchinetti, S. Pittock, B. Weinshenker, The spectrum of neuromyelitis optica, Lancet Neurol. 6 (2007) 805–815.

[43] J.S. Retel, A.J. Nieuwkoop, M. Hiller, V.A. Higman, E. Barbet-Massin, J. Stanek, L.B. Andreas, W.T. Franks, B.J. Van Rossum, K.R. Vinothkumar, L. Handel, G.G. De Palma, B. Bardiaux, G. Pintacuda, L. Emsley, W. Kühlbrandt, H. Oschkinat, Structure of outer membrane protein G in lipid bilayers, Nat. Commun. (2017), https://doi.org/10.1038/s41467-017-02228-2.

[44] Y. Miao, H. Qin, R. Fu, M. Sharma, T.V. Can, I. Hung, S. Luca, P.L. Gor'kov, W.W. Brey, T.A. Cross, M2 proton channel structural validation from full-length protein samples in synthetic bilayers and *E. coli* membranes, Angew. Chemie Int 51 (2012) (8383–8386. doi:10.1002/anie.201204666).

[45] R. Zhang, K.H. Mroue, A. Ramamoorthy, Proton-based ultrafast magic angle spinning solid-state NMR spectroscopy, Acc. Chem. Res. 50 (2017) 1105–1113, https://doi.org/10.1021/acs.accounts.7b00082.

[46] V.S. Mandala, J.K. Williams, M. Hong, Structure and dynamics of membrane proteins from solid-state NMR, Annu. Rev. Biophys. 47 (2018) 201–222, https://doi.org/10.1146/annurev-biophys-070816-033712.

[47] A.T. Petkova, Y. Ishii, J.J. Balbach, O.N. Antzutkin, R.D. Leapman, F. Delaglio, R. Tycko, A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR, Proc. Natl. Acad. Sci. 99 (2002) 16742–16747, https://doi.org/10.1073/pnas.262663499.

[48] T. Schubeis, T. Le Marchand, L.B. Andreas, G. Pintacuda, 1H magic-angle spinning NMR evolves as a powerful new tool for membrane proteins, J. Magn. Reson. 287 (2018) 140–152, https://doi.org/10.1016/j.jmr.2017.11.014.

[49] E. Barbet-Massin, A.J. Pell, J.S. Retel, L.B. Andreas, K. Jaudzems, W.T. Franks, A.J. Nieuwkoop, M. Hiller, V. Higman, P. Guerry, A. Bertarello, M.J. Knight, M. Felletti, T. Le Marchand, S. Kotelovica, I. Akopjana, K. Tars, M. Stoppini, V. Bellotti, M. Bolognesi, S. Ricagno, J.J. Chou, R.G. Griffin, H. Oschkinat, A. Lesage, L. Emsley, T. Herrmann, G. Pintacuda, Rapid proton-detected NMR assignment for proteins with fast magic angle spinning, J. Am. Chem. Soc. (2014), https://doi.org/10.1021/ja507382j.

[50] V.A. Higman, Solid-state MAS NMR resonance assignment methods for proteins, Prog. Nucl. Magn. Reson. Spectrosc. 106–107 (2018) 37–65, https://doi.org/10.1016/j.pnmrs.2018.04.002.

[51] J.M. Lamley, M.J. Lougher, H.J. Sass, M. Rogowski, S. Grzesiek, J.R. Lewandowski, Unraveling the complexity of protein backbone dynamics with combined 13C and 15N solid-state NMR relaxation measurements, Phys. Chem. Chem. Phys. 17 (2015) 21997–22008, https://doi.org/10.1039/C5CP03484A.

[52] P. Rovó, R. Linser, Microsecond timescale protein dynamics: a combined solid-state NMR approach, ChemPhysChem. 19 (2018) 34–39, https://doi.org/10.1002/cphc.201701238.

[53] S. Liao, Y. Yang, D. Tietze, M. Hong, The influenza M2 cytoplasmic tail changes the proton-exchange equilibria and the backbone conformation of the transmembrane histidine residue to facilitate proton conduction, J. Am. Chem. Soc. 137 (2015) 6067–6077.

[54] M.P. Bhate, A.E. McDermott, Protonation state of E71 in KcsA and its role for channel collapse and inactivation, Proc. Natl. Acad. Sci. 109 (2012) 15265–15270, https://doi.org/10.1073/pnas.1211900109.

[55] Y. Ding, L.M. Fujimoto, Y. Yao, F.M. Marassi, Solid-state NMR of the Yersinia pestis outer membrane protein ail in lipid bilayer nanodiscs sedimented by ultracentrifugation, J. Biomol. NMR 61 (2015) 275–286, https://doi.org/10.1007/s10858-014-9893-4.

[56] C. Barnaba, A. Ramamoorthy, Picturing the membrane-assisted choreography of cytochrome P450 with lipid nanodiscs, ChemPhysChem. 19 (2018) 2603–2613, https://doi.org/10.1002/cphc.201800444.

[57] D. Lacabanne, M.-L. Fogeron, T. Wiegand, R. Cadalbert, B.H. Meier, A. Böckmann, Protein sample preparation for solid-state NMR investigations, Prog. Nucl. Magn. Reson. Spectrosc. 110 (2019) 20–33, https://doi.org/10.1016/j.pnmrs.2019.01.001.

[58] F.W. Studier, Protein production by auto-induction in high density shaking cultures, Protein Expr. Purif. 41 (2005) 207–234, https://doi.org/10.1016/j.pep.2005.01.016.

[59] F.W. Studier, Studier Recipe, (2005), p. 234.

[60] D. Manley, J.D. O'Neil, B.S. Selinsky (Ed.), Preparation of Glycerol Facilitator for Protein Structure and Folding Studies in Solution BT - Membrane Protein Protocols: Expression, Purification, and Characterization, Humana Press, Totowa, NJ, 2003, pp. 89–101, , https://doi.org/10.1385/1-59259-400-X:89.

[61] H. Block, B. Maertens, A. Spriestersbach, N. Brinker, J. Kubicek, R. Fabis, J. Labahn, F. Schäfer, Immobilized-metal affinity chromatography (IMAC). A review, Methods Enzymol. 463 (2009) 439–473, https://doi.org/10.1016/S0076-6879(09)63027-5.

[62] D.N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements, J. Struct. Biol. 152 (2005) 36–51, https://doi.org/10.1016/j.jsb.2005.07.007.

[63] S.H.W. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination, J. Struct. Biol. 180 (2012) 519–530, https://doi.org/10.1016/j.jsb.2012.09.006.

[64] S. Förster, L. Apostol, W. Bras, Scatter: software for the analysis of nano-and

mesoscale small-angle scattering, J. Appl. Crystallogr. (2010), https://doi.org/10.1107/S0021889810008289.

[65] J. Pérez, A. Koutsioubas, Memprot: a program to model the detergent corona around a membrane protein based on SEC-SAXS data, Acta Crystallogr. Sect. D: Biol. Crystallogr. 71 (2015) 86–93, https://doi.org/10.1107/S1399004714016678.

[66] S. Jo, T. Kim, V.G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM, J. Comput. Chem. 29 (2008) 1859–1865, https://doi.org/10.1002/jcc.20945.

[67] X. Cheng, S. Jo, H.S. Lee, J.B. Klauda, W. Im, CHARMM-GUI micelle builder for pure/mixed micelle and protein/micelle complex systems, J. Chem. Inf. Model. (2013), https://doi.org/10.1021/ci4002684.

[68] D. Franke, M.V. Petoukhov, P.V. Konarev, A. Panjkovich, A. Tuukkanen, H.D.T. Mertens, A.G. Kikhney, N.R. Hajizadeh, J.M. Franklin, C.M. Jeffries, D.I. Svergun, ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions, J. Appl. Crystallogr. 50 (2017) 1212–1225, https://doi.org/10.1107/S1600576717007786.

[69] D. Svergun, C. Barberato, M.H.J. Koch, CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 32 (2017) vi, https://doi.org/10.1016/s0167-8116(15)00050-6.

[70] D. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.

[71] M.V. Petoukhov, J.B. Vicente, P.B. Crowley, M.A. Carrondo, M. Teixeira, D.I. Svergun, Quaternary structure of flavorubredoxin as revealed by synchrotron radiation small-angle X-ray scattering, Structure. 16 (2008) 1428–1436, https://doi.org/10.1016/j.str.2008.06.009.

[72] K. Takegoshi, S. Nakamura, T. Terao, 13C–1H dipolar-driven 13C–13C recoupling without 13C rf irradiation in nuclear magnetic resonance of rotating solids, J. Chem. Phys. 118 (2003) 2325–2341, https://doi.org/10.1063/1.1534105.

[73] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293, https://doi.org/10.1007/BF00197809.

[74] D.M. Manley, M.E. McComb, H. Perreault, L.J. Donald, H.W. Duckworth, J.D. O'Neil, Secondary structure and oligomerization of the *E. coli* glycerol facilitator. Biochemistry. 40 (2000) 12303–12311, https://doi.org/10.1021/bi000703t.

[75] T. Street, D. Bolen, G. Rose, A molecular mechanism for osmolyte-induced protein stability, PNAS. 103 (2006) 13997–14002.

[76] B.M. Baynes, D.I.C. Wang, B.L. Trout, Role of arginine in the stabilization of proteins against aggregation, Biochemistry. 44 (2005) 4919–4925, https://doi.org/10.1021/bi047528r.

[77] D. Eliezer, P.A. Jennings, H.J. Dyson, P.E. Wright, Populating the equilibrium molten globule state of apomyoglobin under conditions suitable for structural characterization by NMR, FEBS Lett. 417 (1997) 92–96, https://doi.org/10.1016/S0014-5793(97)01256-8.

[78] A. Patel, L. Malinovska, S. Saha, J. Wang, S. Alberti, Y. Krishnan, A.A. Hyman, ATP as a biological hydrotrope, Science (80-.) 356 (2017) 753–756, https://doi.org/10.1126/science.aaf6846.

[79] A. Piai, Q. Fu, J. Dev, J.J. Chou, Optimal bicelle q for solution NMR studies of protein transmembrane partition, Chesmistry. 23 (2017) 1361–1367, https://doi.org/10.1186/s40945-017-0033-9.Using.

[80] C. Le Bon, A. Marconnet, S. Masscheleyn, J.L. Popot, M. Zoonens, Folding and stabilizing membrane proteins in amphipol A8–35, Methods 147 (2018) 95–105, https://doi.org/10.1016/j.ymeth.2018.04.012.

[81] A. Nath, W.M. Atkins, S.G. Sligar, Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins, Biochemistry. 46 (2007) 2059–2069, https://doi.org/10.1021/bi602371n.

[82] J. Galka, S. Baturin, D. Manley, A. Kehler, J.D. O'Neil, Stability of the glycerol facilitator in detergent solutions, Biochemistry. 47 (2008) 3513–3524.

[83] N.G. Aduri, H.A. Ernst, B.K. Prabhala, S. Bhatt, T. Boesen, M. Gajhede, O. Mirza, Human proton coupled folic acid transporter is a monodisperse oligomer in the lauryl maltose neopentyl glycol solubilized state, Biochem. Biophys. Res. Commun. 495 (2017) 1738–1743, https://doi.org/10.1016/j.bbrc.2017.12.008.

[84] T. Gewering, D. Januliene, A.B. Ries, A. Moeller, Know your detergents: a case study on detergent background in negative stain electron microscopy, J. Struct. Biol. 203 (2018) 242–246, https://doi.org/10.1016/j.jsb.2018.05.008.

[85] A.G. Kikhney, D.I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins, FEBS Lett. 589 (2015) 2570–2577, https://doi.org/10.1016/j.febslet.2015.08.027.

[86] H. Fischer, M. De Oliveira Neto, H.B. Napolitano, I. Polikarpov, A.F. Craievich, Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale, J. Appl. Crystallogr. 43 (2010) 101–109, https://doi.org/10.1107/S0021889809043076.

[87] V. Piiadov, E. Ares de Araújo, M. Oliveira Neto, A.F. Craievich, I. Polikarpov, SAXSMoW 2.0: online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale, Protein Sci. 28 (2019) 454–463, https://doi.org/10.1002/pro.3528.

[88] L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R.F. Boyko, B.D. Sykes, D.S. Wishart, VADAR: a web server for quantitative evaluation of protein structure quality, Nucleic Acids Res. 31 (2003) 3316–3319, https://doi.org/10.1093/nar/gkg565.

[89] D.A. Korasick, J.J. Tanner, Determination of protein oligomeric structure from small-angle X-ray scattering, Protein Sci. 27 (2018) 814–824, https://doi.org/10.1002/pro.3376.

[90] S. Abel, F.Y. Dupradeau, E.P. Raman, A.D. MacKerell, M. Marchi, Molecular simulations of dodecyl-β-maltoside micelles in water: influence of the headgroup conformation and force field parameters, J. Phys. Chem. B (2011), https://doi.org/10.1021/jp109545v.

[91] M.V. Petoukhov, D.I. Svergun, Global rigid body modeling of macromolecular complexes against small-angle scattering data, Biophys. J. 89 (2005) 1237–1250, https://doi.org/10.1529/biophysj.105.064154.

[92] B.T. Arachea, Z. Sun, N. Potente, R. Malik, D. Isailovic, R.E. Viola, Detergent selection for enhanced extraction of membrane proteins, Protein Expr. Purif. 86 (2012) 12–20, https://doi.org/10.1016/j.pep.2012.08.016.

[93] A. Varshavsky, N-degron and C-degron pathways of protein degradation, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 358–366, https://doi.org/10.1073/pnas.1816596116.

[94] A. Sivashanmugam, V. Murray, C. Cui, Y. Zhang, J. Wang, Q. Li, Practical protocols for production of very high yields of recombinant proteins using Escherichia coli, Protein Sci. 18 (2009) 936–948, https://doi.org/10.1002/pro.102.

[95] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C.F. Schulte, D.E. Tolmie, R. Kent Wenger, H. Yao, J.L. Markley, BioMagResBank, Nucleic Acids Res. 36 (2007) D402–D408, https://doi.org/10.1093/nar/gkm957.

[96] B. Dang, M. Mravic, H. Hu, N. Schmidt, B. Mensa, W.F. DeGrado, SNAC-tag for sequence-specific chemical protein cleavage, Nat. Methods 16 (2019) 319–322, https://doi.org/10.1038/s41592-019-0357-3.