The lipid environment of Escherichia coli Aquaporin Z

Victoria Schmidt, Marlon Sidore, Cherine Bechara, Jean-Pierre Duneau, James N. Sturgis

The lipid environment of *Escherichia coli* Aquaporin Z

Victoria Schmidt[a], Marlon Sidore[a], Cherine Bechara[b], Jean-Pierre Duneau[a], James N. Sturgis[a,*]

[a]*LISM UMR 7255, CNRS and Aix-Marseille University, 31 Chemin Joseph Aiguier, 13402 Marseille cedex 20, France*
[b]*IGF, CNRS, INSERM, University of Montpellier, Montpellier, France*

Abstract

In this study, we have investigated the lipids surrounding AqpZ, and the effects of a destabilizing mutation $W_{14}A$ (Schmidt and Sturgis 2017) on lipid protein interactions. In a first approach, we used Styrene Maleic Acid copolymer to prepare AqpZ containing nanodiscs, and these were analyzed for their lipid content, investigating both the lipid head-group and acyl-chain compositions. These results were complemented by native mass spectrometry of purified AqpZ in the presence of lipids, to give insights of variations in lipid binding at the surface of AqpZ. In an effort to gain molecular insights, to aid interpretation of these results, we performed a series of coarse grained molecular dynamics simulations of AqpZ, in mixed lipid membranes, and correlated our observations with the experimental measurements. These various results are then integrated to give a clearer picture of the lipid environment of AqpZ, both in the native membrane, and in lipid nanodiscs. We conclude that AqpZ contains a lipid binding-site, at the interface between the monomers of the tetramer, that is specific for cardiolipin. Almost all the cardiolipin, in AqpZ containing nanodiscs, is probably associated with this site. The SMA 3:1 nanodiscs we obtained contain a rather high proportion of lipid, and in the case of nanodiscs containing AqpZ cardiolipin is depleted. This is possibly because, in the membrane, there is little cardiolipin not associated with binding sites on the surface of the different membrane proteins. Surprisingly, we see no evidence for lipid sorting based on acyl chain length, even in the presence of a large hydrophobic mismatch, suggesting that conformational restrictions are energetically less costly than lipid sorting.

Keywords: Styrene Maleic Acid, SMALP, Lipidomics, Cardiolipin, Membrane Protein

1. Introduction

Transfer of information, and material, between the environment and the interior of living cells, is mediated by membrane proteins embedded in the lipid bilayer. Because of this important role, they are the targets of more than 50% of drugs currently on the market [1]. However, studying transmembrane proteins remains a major challenge, because of the lipid membrane that both forms their native environment, and strongly influences their structure and function. For instance, it is known that interactions between negatively charged phospholipids and positively charged amino acid residues contribute to the orientation of membrane proteins [2, 3, 4]. Moreover, there is an abundance of structural data, particularly from X-ray crystallography, of lipids bound to membrane proteins pointing a role of specific lipids for protein function [5, 6, 7]. Thus, a major challenge in membrane protein biology is to understand how the structural determinants of membrane proteins and membrane lipids influence the folding, structure and function of transmembrane proteins, in their native environment.

Annular lipids, that form a dynamic rapidly exchanging solvation shell around membrane proteins [8], could be analyzed using lipid nanodiscs, allowing the study of the effects of local lipids on protein function [9]. Amongst the methods for forming lipid nanodiscs, Styrene Maleic Acid (SMA) copolymer is particularly interesting. This polymer is able to directly dissolve membranes, without the use of classical detergent, forming solubilized patches that can contain membrane proteins along with an annular lipid shell held in solution by a coat of SMA [10, 11]. By isolating membrane proteins associated with a lipid annulus, SMA-Lipid-Particles (SMALP's) have proved able to produce stable GPCRs suitable for biophysical characterization [12]. SMA copolymer has also been used to study the lipid environment of various membrane proteins including the potassium channel KcsA [13] and the Sec translocon [14]. In both these studies, an enrichment in negatively charged lipids, especially of cardiolipin (CL), was found in the lipid annulus, pointing to a potential importance for their function, or a heterogeneity of lipid composition.

Tightly bound structural lipids, as opposed to annular lipids, can be frequently detected in high quality membrane protein crystallographic or cryo-EM structures [15, 16, 17]. However, defining small molecules within atomic structures remains problematic, as it is often hard to dis-

*Corresponding author
Email address: james.sturgis@univ-amu.fr (James N. Sturgis)

tinguish detergent molecules from lipids, and to define the positions of dynamic hydrocarbon side-chains and buried head groups [18]. Using intact mass spectrometry (MS), it is nowadays possible to follow specific membrane protein–lipid interactions, allowing both the quantification, and fine characterization, of bound lipids [19]. Intact MS has been successfully used to identify endogenous co-purified lipids with membrane protein complexes [20, 21, 22], as well as to follow the kinetics and thermodynamics of exogenously added lipids [23, 24]. However, it is difficult to get a precise structural vision of these binding sites. Fortunately, molecular dynamic simulations, and in particular coarse grain (CG) simulations, have demonstrated their ability to reproduce many phenomenon, at time scales up to 100 μsec necessary for the description of complex systems. Despite the reduction of resolution and the simplification of the energy potential, CG models have already been used successfully to study several different types of lipid-protein interactions, giving an increasingly profound, and frequently experimentally confirmed, vision of the molecular organization of lipid membranes [25, 26, 27, 28].

The major intrinsic proteins form an important family of membrane proteins. With members in all kingdoms of life, they form channels allowing passive transport across biological membranes. AquaporinZ (AqpZ), the *E. coli* water channel, is responsible for the gated passive water entry into cells [29], allowing homeostasic regulation of cell volume and osmotic pressure. In common with other members of the family, its structure contains six transmembrane alpha helices that form the pore and two supplementary half-helices with the characteristic NPA motif entering into the channel forming an hourglass-like structure [30]. In their native membrane environment, they form homo-tetramers, resistant to SDS and trypsin denaturation [31]. In a previous study, we have examined the stability of these tetramers [32], and shown that the tetrameric structure can be considerably destabilized by creating stearic clashes in the hydrophobic interface. The function of *E. coli* AqpZ has been shown to be dependent of lipid composition [33]. AqpZ reconstituted in liposomes, made of lipids from a CL deficient *E. coli* strain, exhibits a striking reduction in water transport, compared to the same protein reconstituted in lipids from the wild-type strain. In agreement, native MS demonstrated that while neither phosphatidylethanolamine (PE) nor phosphatidylglycerol (PG) significantly stabilized AqpZ structure, CL strongly bound AqpZ and resulted in a substantial increase in protein stability, in the gas phase [33]. Moreover, measurements on the protein, reconstituted into Membrane Protein Scaffold (MSP) nanodiscs, showed a very strong association between AqpZ and the bilayer, resulting in tetramer dissociation being favored over lipid removal [34].

In this study, we have investigated lipids surrounding AqpZ, both tightly bound and annular, and the effects of a destabilizing mutation $W_{14}A$ [32] on these lipids. As a first approach, we used SMA to prepare AqpZ containing nanodiscs, and these were analyzed for their lipid content, investigating both the lipid head-group and acyl-chain compositions. Intact MS was then used to compare lipid binding affinities between mutant AqpZ proteins. In an effort to gain molecular insights into these results, we performed a series of CG molecular dynamics simulations, of AqpZ in mixed lipid membranes, and correlated our observations with the experimental measurements. These various results are then integrated to give a clearer picture of the lipid environment of AqpZ in the native membrane, and in lipid nanodiscs.

2. Materials and Methods

2.1. Materials

All lipids were from Avanti polar lipids (Coger SAS, Paris, France) , including MS internal standards, *E.coli* lipids POPE (1-palmitoyl-2-oleyl-sn-glycero-3-phosphoethanolamine), POPG (1-palmitoyl-2-oleyl-sn-glycero-3-phospho-(1-rac-glycerol) and CL (cardiolipin).

Deuterated fatty acids were from CDN isotopes (Cluzeau Info Lab, Sainte Foy la Grande, France) and Cayman (Bertin Pharma, Montigny le Bretonneux, France).

LC-MS/MS quality solvents were purchased from Fischer Scientific (Illkirch, France). All other chemicals of the highest grade available were purchased from Sigma Aldrich (Saint-Quentin Fallavier, France).

Hydrolyzed styrene-maleic acid copolymer (SMA) with a 3:1 styrene to maleic acid ratio in NaOH solution at 25% w/v was provided by Polyscope (Xiran SL25010 S25).

2.2. Protein expression and purification

BL21(DE3)pLys strain was transformed with the pET-6His-TeV plasmid into which was cloned the AqpZ-$C_{20}S$ mutant gene [31]. For some control experiments the W3110 and W3110 Δcardiolipin strain [35] were used. The $W_{14}A$ mutant was derived from this as described previously [32]. Bacterial cultures were grown in Luria-Bertani broth, containing 25 μg/ml ampicillin, and incubated for 13 to 16 hours at 37°C, diluted 100-fold into fresh broth and grown to an A_{600} of 1.5. 1 mM Isopropyl-β-D-1-thiogalactopyranoside (IPTG) was then added and the culture was incubated for a further 3 hours at 30°C, before harvesting by centrifugation, 20 minutes at 9000×g. Harvested cells were resuspended in 1/200 of the original culture volume of ice cold lysis buffer (20 mM Tris-HCl, 1 mM phenylmethylsulfonylfluoride (PMSF), 0.1 mg/ml deoxyribonuclease 1 pH 8.0). Cells were broken with 3 passages through an Emulsiflex-C5 (Avestin) cell disrupter. Unbroken cells, the cell wall, most of the outer membrane and other debris, were separated from the cell lysate by a 30 min centrifugation at 10,000×g. The supernatant, containing membrane fragments, was centrifuged 60 minutes at 140,000×g and the pellet resuspended in 20 mM Tris-HCl, 200 mM NaCl pH 8.0.

Membranes were solublized by adding 3g of SMA, previously dialysed against Tris-HCl 20 mM pH 8.0, per gram

of membrane and incubated for one hour at room temperature with agitation. Insoluble material was pelleted by a 60 minutes centrifugation at 14,000×g, and the soluble fraction was kept at 4°C until purification. Protein purification was performed by affinity chromatography, on an His-Trap HP column, connected to the Akta system (both from GE Healthcare). The column was equilibrated, with 10 column volumes (CV) of wash buffer (20 mM Tris-HCl, 200 mM NaCl, 50 mM Imidazole pH 8.0), the solubilized membrane fraction was charged and the column, and washed with 5 CV of wash buffer, to remove non-specifically bound material. Elution was performed with a 10 CV gradient, running from 0% to 100% elution buffer (20 mM Tris-HCl, 200 mM NaCl, 500 mM Imidazole pH 8.0). Fractions, containing AqpZ in nanodiscs, were pooled and further purified, by size exclusion chromatography.

The Superose 6 10/300 GL (GE Healthcare) column, connected to the Akta system, was equilibrated with 2 CV of Tris-HCL 20 mM, NaCl 200 mM buffer, and the sample separated at an elution rate of 0.25 mL/min. Fractions containing AqpZ in nanodiscs were collected. Particles sizes were measured by dynamic light scattering (DLS) on a Zetasizer Nano-series Nano-S instrument (Malvern). Samples were equilibrated for 300 s at 25°C. Default software settings were used for optimizing measurement settings and duration. Multiple narrow modes analysis of the correlation data was used to obtain volume-based particle size distributions using the manufacturers software.

SDS-PAGE was performed using standard laboratory procedures.Briefly, samples were prepared by denaturing 3 parts protein solution with 1 part 4x sample buffer (Tris-HCl pH 6.8 400mM, glycerol 40%, SDS 4%, β-mercaptoethanol 2.5%, Bromophenol Blue 0.04%) without heating. Solubilized samples were loaded onto 12.5% SDSpolyacrylamide gels with a 5% stacking buffer and separated using a Mini-PROTEAN electrophoresis system (Bio-Rad). Gels were stained with Coommassie Blue (40% ethanol, 20% acetic acid, 0.25% Coommassie blue) at room temperature, and then destained (40% ethanol, 20% acetic acid).

For electron microscopy, carbon-coated copper grids (Electron Microscopy Sciences) were glow-discharged for 30 sec. A droplet of sample, 5 μL, was placed on the grid and removed after 60 sec by blotting with filter paper (Whatman). The grid was then washed with water, and excess water was removed by blotting with filter paper prior to staining with 5 μL of 2% uranyl acetate for 60 sec. Excess of stain was removed by blotting with filter paper and the grid was dried at room temperature. Images were recorded on a FEI Tecnai 200 kV electron microscope operating at a voltage of 200 kV and a defocus of 2.5 μm, using a Eagle-CCD camera (FEI).

2.3. Lipid Extraction and Thin Layer Chromatography

Lipids were extracted and analysed as follows from membranes, SMA solubilised membranes, or purified proteins, prepared as described above, following the procedure of Bligh and Dyer[36]. Briefly, 200 μl of lipid-containing samples were added to 1.5 mL chloroform:methanol, 20 μl HCl 1 M and mixed over 1 hour. Phase separation was achieved by a 3 min centrifugation at 9,000×g and the lipid-containing organic phase was collected. Debris were washed by adding 500 μl chloroform, again centrifuged and the organic phases pooled. Four drops of benzene were added to the organic fraction and well mixed. Lipids were dried under nitrogen before resuspending in 200μl methanol:chloroform 2:1 (vol/vol). Lipid extracts were stored under nitrogen at -20°C.

For TLC analysis, silica gel on TLC plates (thickness 250 μm Sigma (Fluka)) were washed with chloroform:-methanol:acetic acid 25:65:5 (vol/vol/vol) and dried for 1 hour. Lipid samples were deposited with a syringe on the plate and this was developed with the same solvent. Phospholipids were visualized by spraying the plates with Molybdenum blue 1.3% reagent (Sigma) to reveal phosphate containing molecules[37]. Phosphate were quantified from images of the TLC plates using ImageJ [38, 39] software to integrate the peaks, see figure 2B for the standard curve.

To facilitate comparison lipid concentrations are reported in wt%. The desitometric measurement allowed a direct determination on the basis of phosphate content, this was converted to wt% based on the average molecular mass calculated using the fatty acid composition determined in the membranes, see below, and the phosphorous content. The average molecular masses per phosphorous were respectively: 646.17, 622.95, and 565.01; for PE, PG and CL.

2.4. Lipid chain length characterization

For targeted analysis each lipid extract (20 μl) was spiked with 20 μL of an internal standard mix containing (17:0/ 17:0)-PE (2 μg), (17:0/ 17:0)-PG (2μg), (14:0/ 14:0/ 14:0/ 14:0)-CL (0.1 μg). Total fatty acid analysis was conducted using 25 μL of lipid extract spiked with myristic acid-d3 (260 ng), palmitic acid-d3 (1128 ng), stearic acid-d3 (840 ng), linoleic acid d4 (650 ng), arachidic acid-d3 (1.04 ng), arachidonic acid-d8 (432 ng), behenic acid-d3 (1.08 ng), DHA-d5 (108 ng), Lignoceric-d4 (0.52 ng) and cerotic acid-d4 (0.4 ng).

PE's and CL's were analysed as was previously described [40]. Total fatty acids were analyzed as described by Baarine et al [41]. Targeted analysis of PG's was conducted by LC-MS/MS as follows. Lipids were separated by high performance liquid chromatography (HPLC), on an Agilent 1200 equipped with an auto sampler, a binary pump and a column oven. During analysis, the reverse phase chromatography column (ZorBAX EclipsePlus C18 - 2.1 mm x 100 mm, 1.8 μm, Agilent Technologies) was maintained at 45°C. The mobile phases consisted of acetonitrile/water (90/10 v/v) containing 0.2% ammonium hydroxide and 0.5% acetic acid (A) and of propan-2-ol/methanol (90/10) containing 0.2% ammonium hydroxide and 0.5% acetic acid (B). PG's were separated by a linear gradient as follows: 30% B up to 50% B in 12 min then up

to 100% B in 1 min and maintained at 100% B for 1 min. The column was equilibrated with 30% B for 4 min before each injection. The flow rate was maintained at 0.25 mL/min. The auto sampler temperature was set at 6°C. Two microliters of sample were injected. The LC system was coupled to an Agilent 6460 QqQ triple quadrupole mass spectrometer equipped of a Jet Stream electrospray ionization source (ESI) set up as follows: source temperature 300°C, nebulizer nitrogen gas flow rate 10 L/min at 20 psi, nitrogen sheath gas flow 11 L/min, temperature 290°C, capillary voltage 3500 V, nozzle voltage 750 V. Relative quantitation of PG's was performed by calculating the response ratio of the considered phospholipid to 17:0/17:0-PG internal standard.

2.5. Lipid titration by native mass spectrometry

Prior to MS analysis, $W_{14}A$ and $C_{20}S$ AqpZ mutants were buffer exchanged into 200 mM ammonium acetate buffer pH 8.0 (Sigma), supplemented with either 0.02% DDM or 0.5% C_8E_5 detergent, using Bio-Spin microcentrifuge columns (Bio-Rad Laboratories). Proteins were quantified by the Bradford method[42], after buffer exchange (concentrations ranged between 5 and 15 μM). All lipids were dissolved in $CHCl_3$ in a glass vial, lipid films were generated by solvent evaporation under a stream of N_2 and dried in a vacuum chamber over-night. Dry lipid films were dissolved in 200 mM ammonium acetate, 0.5 % C_8E_5, to a final concentration in the range of 2-5 mg/mL, and sonicated in a water bath for 1 hour. The resulting lipid stock solution concentration was determined by phosphorus assay [43].

A serial dilution for each lipid was made to reach different ligand concentration, then 0.5 μL of lipid solution was mixed with 4.5 μL of the protein solution in C_8E_5 detergent, incubated for 15 min at room temperature, then loaded into a borosilicate emitter (Thermo Scientific). Intact MS spectra were recorded on a Synapt G1 HDMS instrument (Waters Corporation) modified for high mass analysis and operated in ToF mode. Samples were introduced into the ion source using borosilicate emitters (Thermo Scientific). Optimized instrument parameters were as follows: source pressure 5.3 mbar, capillary voltage 1.5 kV, sampling cone voltage 180 V, extractor cone voltage 4 V, transfer collision voltage 40 V, argon flow rate 5 mL/min and trap bias 30 V. Collision voltage in the trap was optimized depending on the detergent to the minimum activation required to strip the detergent micelle. Spectra were deconvoluted using the software UniDec [44]. Computation of individual binding site K_D values was performed using the Data Collector Module in UniDec. Extracted intensities for the *apo*-protein and each separate bound state were fit to a sequential binding model using free K_D values for each site.

2.5.1. E. coli Inner Membrane Model

A CG model of the *E. coli* inner membrane was constructed, using the previously published lipidomic analysis [45]. While the composition of the inner membrane of *E. coli* has many different types of lipids, the CG approximation, using the MARTINI force field [46, 47], results in only 9 different types of lipids, shown in supplementary table S2. The CL model was taken from [48], with the addition of a second species containing 2 unsaturations instead of 4.

The model membrane was built and solvated with a 0.15 mol/L NaCl in water using INSANE [49]; briefly, the program generates a lipid bilayer by distributing the lipids over a grid, with lipid structures derived from simple template definitions. Three systems were built: the first with only lipids, the second with a tetrameric AqpZ molecule in a lipid bilayer, and the third with an AqpZ $W_{14}A$ tetramer. Each system has an average size of 20x20x10 nm. It was found necessary to stabilize the CG tetrameric protein structure as this had a tendency to collapse in the simulations, this was achieved by filling the central void in the tetramer structure with two phospholipid molecules. It should be noted that the acyl chain length of the PE and PG phospholipids are represented by 4 beads while those of CL are 5 beads. This is the result of combining the original [46] and an updated [50] MARTINI lipid topologies, the CL topology being available only in the former.

2.5.2. Molecular Dynamics Simulation

All simulations were performed using the GROMACS simulation software (version 5.1) [51, 52]. The models were enclosed in an orthorhombic box with periodic boundary conditions. The systems were first equilibrated with a 1 ns simulation in the NVT ensemble followed by a 1 ns simulation in the NPT ensemble. The equilibrated systems were then used as starting configurations for 15 μs molecular dynamic simulations, of which only the final 12 μs were analyzed. The parameters used in the simulations correspond to those previously reported [53]. Briefly, the simulations use a 20 fs time-step, the neighbor list is updated using the Verlet neighbor search algorithm every 20 steps with the neighbor list length being automatically determined. Leonard–Jones and Coulomb potentials and forces are cut off at 1.1 nm, with the potentials shifted to zero at the cut off. Velocity rescale [54] and Parinello–Rahman coupling schemes [55] are used with coupling parameters of 1.0 and 12.0 ps^{-1}. The leapfrog integrator is used with temperature and pressure both kept constant at 325 K and 1 bar respectively.

2.5.3. Simulation analyses

The lipid densities were computed using VMD's Volmap plugin [56]. All the scripts used in this work are freely available on request. The lipid enrichment analysis consists of comparing the lipid composition at various distances around the protein. To calculate lipid dissociation constants the average residency times in lipid binding sites were calculated.

3. Results

3.1. Characterization of AqpZ Nanodiscs

His-tagged AqpZ nanodiscs were purified from *E. coli* cytoplasmic membranes. In the first step SMA solubilized membranes were separated by chromatography on a Ni-NTA column, figure 1A. After this one step affinity purification the protein nanodiscs are reasonably pure, figure 1A inset. In the SDS PAGE AqpZ-C_{20}S is mainly tetrameric, appearing at 80 kDa, with a small amount of the monomeric form visible at 28 kDa, irrespective of the purification conditions. The C_{20}S mutant of AqpZ is slightly destabilized and was originally described by Borgnia [31]. The tetrameric assembly of this protein is resistant to SDS denaturation. However, in SMA, some higher oligomers can be seen, this may be due to the very high protein concentration resulting in crowding or SDS limitation. In contrast, and as previously reported [32], the W_{14}A mutant shows mainly monomers on SDS-PAGE. The same AqpZ in dodecyl maltoside (DDM) detergent solution elutes at 300 mM imidazole, however in SMA nanodiscs 150 mM imidazole is sufficient to elute the protein. This shift of elution to lower imidazole concentrations has been reported previously [57], and rationalized in terms of the impaired binding of the His-tag due to electrostatic interactions with the strongly anionic SMA.

Size exclusion chromatography (SEC) after affinity purification, figure 1B, shows a major peak around 15 ml. There is very little sign of aggregation though some material elutes from the column slightly before of the main AqpZ peak. In view of the protein purity after affinity chromatography, figure 1 A insert, this heterogeneity would appear to be associated with the polymer, and possibly the associated lipids. The main peak corresponds to a hydrodynamic volume of 293 $\pm$ 11 nm^3 and 298 $\pm$ 5 nm^3 for the C_{20}S and W_{14}A mutants respectively. Interestingly these AqpZ nanodiscs, or SMA-lipid particles (SMALP), have very similar sizes, this contrasts markedly from the situation in DDM detergent where AqpZ-W_{14}A detergent protein complexes had a hydrodynamic volume of 76 nm^3 while C_{20}S complexes were about twice as large [32]. For characterization only the main peak after SEC was used.

Purified AqpZ containing nanodiscs were analyzed by dynamic light scattering (DLS), figure 1C, and electron microscopy, figure 1D. Both methods showed a sample that was very homogeneous, with an average diameter of 12.50 nm $\pm$ 0.70. There was little or no sign of sample aggregation or evolution after preparation. In transmission electron micrographs a typical central black spot in the middle of AqpZ tetramers was seen, colored with negative stain, while particles appear in white. Assuming a discoidal shape for the SMA-nanodiscs, we can estimate a volume of 245 nm^3 based on the observed diameter. Hydrodynamic volume of both proteins, AqpZ-C_{20}S and W_{14}A, obtained by SEC is very close to this theoretical value. These results strongly suggest that in SMALP's the W_{14}A mutant is tetrameric.

We also performed compositional analysis of purified SMA-nanodiscs containing AqpZ, using Bradford and UV absorption protein assays, molybedenum blue phosphate assay and LC/MS lipid quantification, and UV absorption SMA quantification[58], (supplementary table S1 and S2). These analyses gave a composition by weight of 35.0 $\pm$ 1.0 % protein; 48.3 $\pm$ 1.5 % lipid; 16.3 $\pm$ 2.5 % SMA. This means that there are a considerable number of lipid molecules associated with each protein in a SMALP, about 140, representing multiple solvation layers.

3.2. Nanodisc lipid composition

To investigate the nature of the lipids around AqpZ within SMA nanodiscs we extracted the lipids from purified SMALP's. The isolated lipids were then analyzed both in terms of head group composition, figure 2, and the fatty acid content, figure 3.

TLC analysis of membrane phospholipid composition based on polar headgroups showed a distribution of 53.2 $\pm$ 4.6% PE, 26.7 $\pm$ 2.8% PG and 20.1 $\pm$ 1.9% CL (wt/wt). This membrane composition is in agreement with the published lipid composition for *E. coli* [59]. The lipid composition after solubilization by SMA was not modified, figure 2.

However, AqpZ containing nanodiscs consistently showed a significant depletion of CL (to 6.2 $\pm$ 2.1 % wt/wt) at the expence of PE. This observation was independent of the mutant protein examined, since little variation in lipid head-group composition was observed between nanodiscs containing the C_{20}S protein and those containing the W_{14}A protein. Our observation contrasts with previous studies on SMA nanodiscs containing KcsA [13] and Sec [14] in which a CL enrichment was observed. This indicates that the observed lipid composition of SMA nanodiscs depends on the protein they contain.

The reduction in CL content appears at first sight to be at odds with the observation that CL is necessary for AqpZ function and binds strongly to the protein [33]. However it must be remembered that the nanodiscs contain mutiple lipids (supplementary table S1) in several layers, these include specific bound lipids and the dynamic lipid annulus. The bound lipids thus represent a minor component of the SMALP. Using the composition the AqpZ containing SMALP's we have measured, the average nanodisc is expected to contain about 4.75 CL molecules. This is clearly sufficient for each monomer to bind a CL molecule in a specific high-affinity binding site but leaves the dynamic annulus strongly depleted in CL. Thus, though there is sufficient CL in the SMALP's to bind to AqpZ high affinity sites ,there is hardly any excess.

LC/MS was used to analyze lipid acyl chains. As *E. coli* is known to form cyclo-propyl lipids [60, 61] it is not possible to distinguish using simple LC/MS if a mass difference of 2.0 is due to the formation of a double bond or a cyclo-propyl ring. Therefore, acyl chains analysis was attributed to a total acyl carbon chain length and number of double bond equivalents. The overall composition is

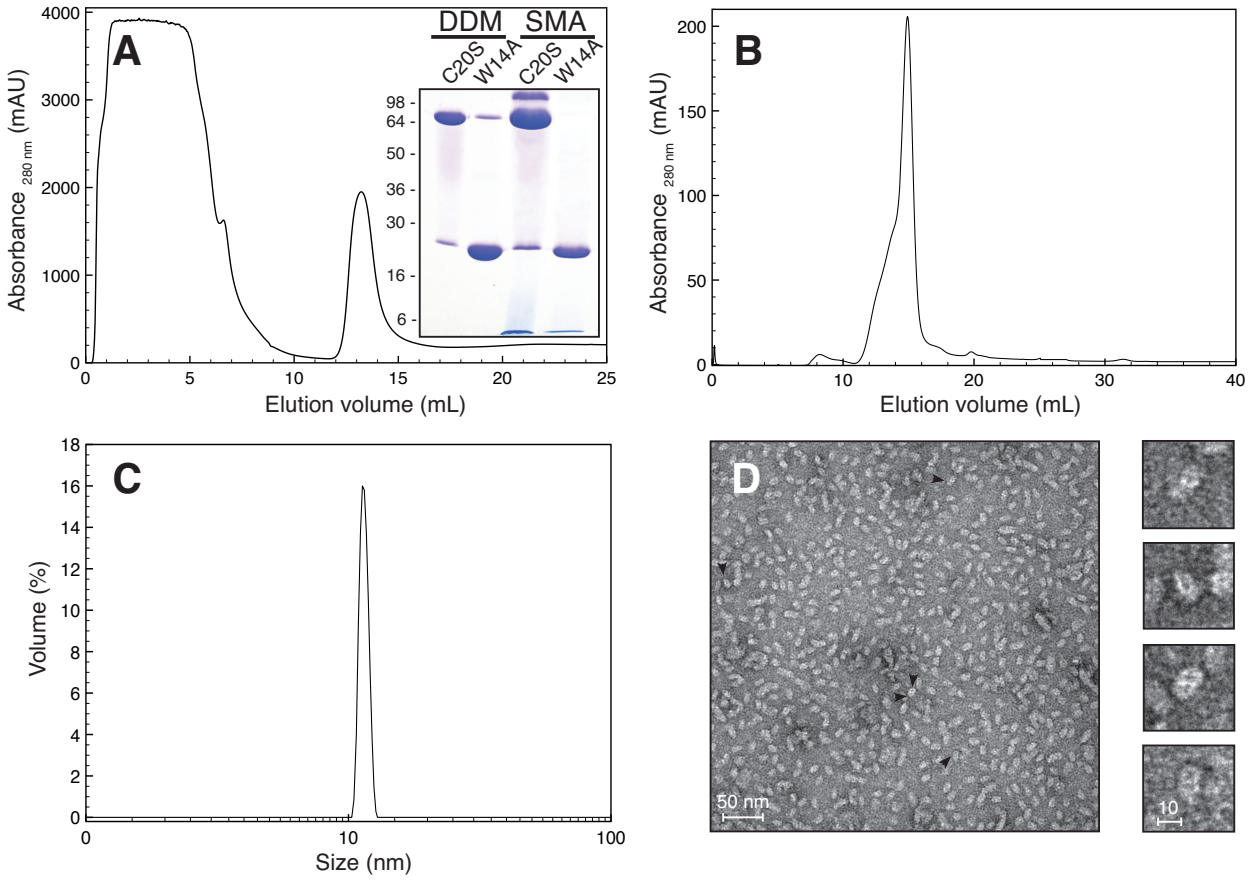

Figure 1: Purification and analysis of AqpZ nanodiscs. A, Chromatogram showing elution profile of His-tagged AqpZ-$C_{20}S$ from a Ni-NTA column. The sample was loaded over the first 5 ml, and then washed with 5 mL of wash buffer before starting elution with a linear imidazole gradient. The peak at 13.5 ml represents AqpZ protein. Inset: SDS-PAGE of purified AqpZ-$C_{20}S$ and $W_{14}A$ mutants in DDM and in SMA stained with Coomassie blue. B, Size exclusion chromatogram of SMA nanodiscs containing AqpZ after affinity purification. C, Dynamic Light Scattering size estimation of AqpZ containing nanodiscs after size exclusion chromatography. D, Negative stained transmission electron micrograph of purified nanodiscs, inset several representative individual particles.

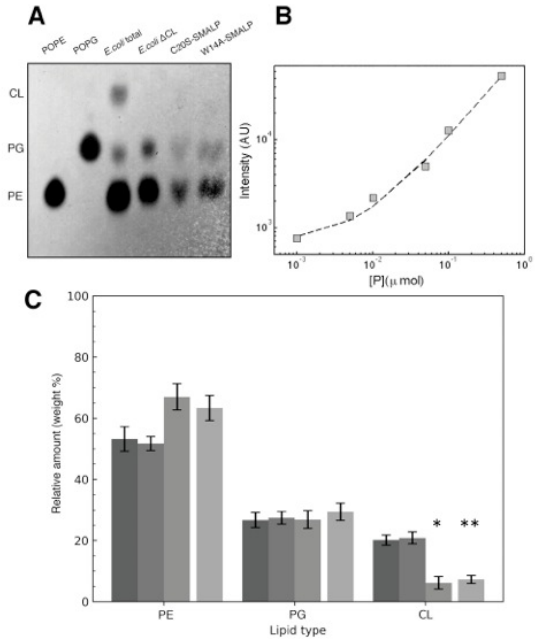

Figure 2: Polar head group composition of lipids isolated from *E. coli* inner membrane and associated with AqpZ and the destabilized $W_{14}A$ mutant in SMALP's. A, separation of polar lipids by TLC and phosphate staining. Polar lipids were isolated by standard techniques and analyzed by thin layer chromatography, phosphate staining and densitometry. B, standard curve for phopholipid determination using densitometric analysis of molybdenum blue stained TLC plates. This graph shows a $log - log$ plot to highlight the dynamic range. The dotted line represents a linear fit to the data. C, histograms to show composition of the different samples. The first three bars for each lipid concern respectively: inner membrane; SMA solubilised inner membrane; and purified protein all isolated from cells expressing AqpZ $C_{20}S$. The fourth lighter bar in each group shows results for the purified AqpZ $W_{14}A$ containing nanodiscs. Compositions shown, on the basis of phosphate content, are the average and standard error determined using three biological replicates. The CL compositions of SMALP's are significantly reduced, (*, P<0.05; **, P< 0.01) relative to the membrane composition.

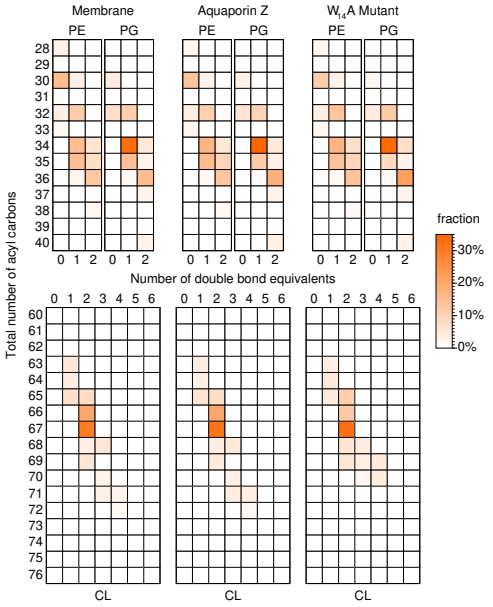

Figure 3: Profile of lipid acyl chains for different phospholipid types extracted from *E. coli* inner membrane, AqpZ-$C_{20}S$ SMA nanodiscs and AqpZ-$W_{14}A$ SMA nanodiscs. Samples were analyzed by LC/MS and the diagrams show the distribution of summed chain lengths and number of unsaturations (double bond equivalents). As CL has 4 acyl chains rather than 2 the summed chain lengths and unsaturations are larger than for PE and PG. Peak intensities were converted in orange levels, the scale bar on the right show the correspondence between intensity and fraction of the lipid class.

entirely consistent with what is known of *E. coli* phospholipids in terms of acyl chain length and unsaturations [62]. The majority of simple phospholipids (PG and PE) had acyl chains with 34 to 36 carbons in total and 1 or 2 double bond equivalents. The most common 34:1 found particularly for PG corresponding to the canonical 16:0/18:1 acyl chain composition. The acyl chain distribution of the PE was more varied than for PG, notably with the presence of significant amounts of shorter more saturated acyl chains. The distribution of CL has clearly centered about a slightly shorter average chain length than the simple phospholipids with an average chain length of 16.25 carbons as opposed to 17.25 for PG and PE. Comparison of the distributions for total inner membrane, $C_{20}S$ and $W_{14}A$ protein containing nanoparticles shows that there is no or very little lipid selection on the basis of acyl chains. Indeed the reproducibility of the measurements is remarkable. More specifically, there is no sign of selection for short lipids round the protein despite the predicted hydrophobic mismatch [25]. The predicted mismatch is indeed quite significant as the hydrophobic thickness of the protein is about 29.7Å, while that of the membrane is 34.9Å.

Taken together, lipid analysis reveals a preference for AqpZ in SMALP's to accumulate PE as opposed to CL, in contrast to previous studies where CL accumulation has been observed. Moreover, no or little selection of lipids with specific chain length or unsaturation was visible for

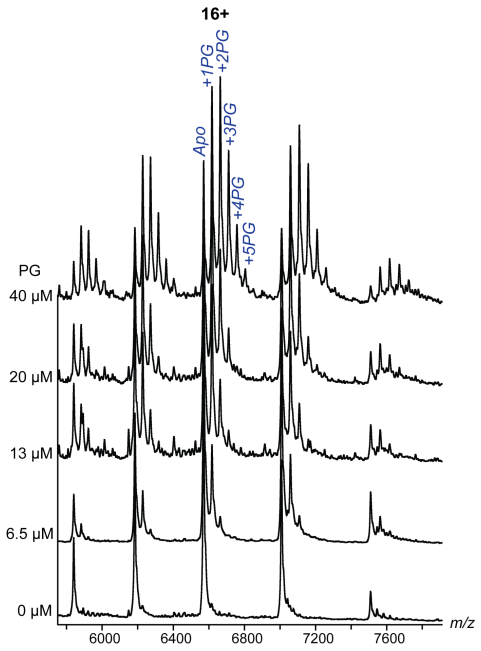

Figure 4: nanoESI spectra of Aqp $C_{20}S$ in the presence of PG showing a concentration dependent increase in the amount of lipids bound to the intact AqpZ tetramer. AqpZ concentration was 7 μM, in 200 mM ammonium acetate pH 8, 0.5% C_8E_5.

AqpZ-containing SMALP's.

3.3. Phospholipid binding and affinity of AqpZ

Binding affinities of the three major *E. coli* phospholipid types PE, PG and CL for the AqpZ $C_{20}S$ and $W_{14}A$ mutants were determined using intact MS. In this experiment different synthetic, exogenously-added, phospholipids compete with the detergent n-octyl-penta oxyethylene (C_8E_5) for binding sites on the surface of purified AqpZ tetramers. Then the number of lipid molecules that remain bound to the protein in the mass spectrometer, under activation conditions that strip the detergent molecules, can be measured and fit to titration curves.

First, intact MS of both AqpZ mutants in DDM detergent showed the presence of tetramers, supplementary figure S1. Under minimum activation energies required to strip the detergent molecules, we could still detect stripped trimers and monomers since DDM is known to be a "sticky" detergent requiring a relatively high energy in order to detach it from proteins in the gas phase. However, the $W_{14}A$ tetramers were less stable than those of $C_{20}S$ mutant since their relative intensity compared to the monomers was lower under similar solution and MS conditions. This is in agreement with previous observations indicating the destabilization of tetramers by this mutation [32].

To compare the relative affinities of both mutants for the three lipid groups tested, we performed a lipid titration under the same solution conditions of 200 mM ammonium acetate buffer pH 8 supplemented with 0.5% C_8E_5. Titration showed a concentration-dependent binding of up to 5 lipids per AqpZ tetramer (figure 4). Binding affinities were then fitted using the sequential binding model with all free K_Ds using the software UniDec [44], supplementary figure S2.

Lipid	$C_{20}S$	$W_{14}A$
POPE	17.3 ± 2.2	35.2 ± 1.6
	38.3 ± 6.8	73.9 ± 5.6
POPG	13.7 ± 2.0	19.8 ± 1.0
	26.7 ± 2.9	45.7 ± 9.7
CL	2.84 ± 0.18	8.64 ± 1.26
	6.69 ± 0.94	20.83 ± 0.58
POPE	64.2 ± 3.0	129.8 ± 7.8
POPG	43.4 ± 4.2	78.5 ± 3.5
CL	11.1 ± 0.4	31.3 ± 1.2

Table 1: K_D's in μM of phospholipids from *E. coli* AqpZ observed by MS. In the top part of the table the sequential dissociation constants obtained from fits to the MS data, supplementary figure S2. In the lower part of the table the dissociation constant calculated assuming the observations arise from 4 identical sites, one on each monomer of the tetramer.

As shown in table 1, dissociation constants reveal the same phopholipid preferences for $C_{20}S$ and $W_{14}A$ mutants, with CL binding the strongest, followed by the negatively charged PG and finally zwitterionic PE showing the lowest affinity. As the proteins are tetramers it is reasonable to expect that each protein has four equivalent binding sites, one for each monomer all with the same affinity. In such cases, it is expected that the apparent dissociation constants follow the series 0.25 (1/4), 0.67 (2/3), 1.5 (3/2), 4.0 (4/1). The observed dissociation constants in all cases follow reasonably well this expectation, and in the lower part of the table we show the estimated dissociation constants assuming 4 equivalent sites. The quality of the fit to this model is apparent in the low levels of uncertainty in the 4 equivalent site dissociation constants, all less than 10%. We therefore conclude that the observed phospholipid binding is to four specific sites on the tetrameric protein surface.

The binding can to some extent be separated in to a head group specific part and an acyl- chain part. Comparison of the values for PG and PE show that the ratio of the dissociation constants is similar (about 1.5 : 1.0) for both proteins. This would suggest that the change in the protein-protein interface does not modify the head group binding site specificity, in both proteins PG is preferred.

The stronger CL binding can be correlated with the larger number of acyl chains and so a more extended binding site, presumably displacing more detergent molecules and resulting in more contacts between the CL and protein. Because of the size difference it is hard to draw quantitative conclusions from this.

It is readily apparent that for all the phospholipids studied the affinity for the $W_{14}A$ mutant is considerably reduced. Indeed, the loss of affinity caused by the mutant, of about 2 fold for PE and PG, suggests that the perturbation affects the acyl-chain binding that is com-

mon for both lipids. The more pronounced effect for CL is consistent with this view. The location of the mutation in the structure at the protein-protein interface suggests lipid binding site is also close to this interface.

At first sight the change in affinity might seem strange as the $W_{14}A$ tetramer is less stable [32] and thus tends to favor lipid-protein interactions over protein-protein interactions. However, this is not the type of lipid interaction that we are observing. The experimental conditions mean that we are measuring the ability of lipids to compete with a large, about 1000 fold, excess of the C_8E_5 detergent. The differences between the mutants would appear to indicate that in each case, but particularly in the case of CL, the perturbation to the structure at the protein-protein interface caused by the $W_{14}A$ mutation disorganizes the lipid binding site and results in a preference for the short flexible chains of the C_8E_5 detergent.

Overall, these observations indicate the preference of AqpZ for anionic lipids, and in particular a strong binding of CL. They also show that there are few high affinity binding sites, one per monomer is adequate to describe the data. The sensitivity of the affinity to mutations in the protein-protein interface suggests that this is the location of specific lipid binding.

3.4. Molecular dynamics simulations

To gain molecular insight into the organization of lipids around AqpZ proteins, a series of molecular dynamics simulations were performed. A CG model of the *E. coli* internal membrane containing different AqpZ variants was constructed. Protein variants were placed, as described in the methods section in a simulated membrane environment containing 75.1% PE, 16.6% PG and 8.3% CL, homogeneously distributed between the two membrane leaflets. The molecular dynamics of these systems was then simulated for 17 μs, using the Martini force field [46]. Lipid binding sites were then studied by examining sites of excess lipid density and those preferred by specific headgroups, figure 5.

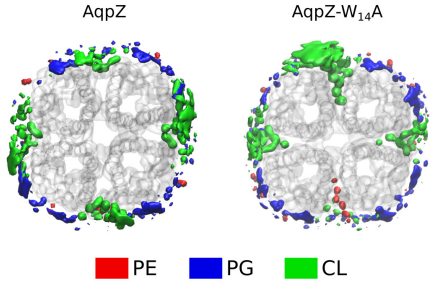

Figure 5: Molecular dynamics simulation of an *E. coli* coarse grain model membrane around AqpZ seen from the cytoplasmic side. Sites of phospholipid occupancy are depicted in colored densities according to the lipid type (PE in red, PG in blue and CL in green). Major binding sites can be seen particularly for PG and CL, at the monomer monomer interface.

The surface of the protein, particularly the cytoplasmic surface, where there are several positive charges, shows a strong affinity for PG and CL. This preferential binding reproduced above in our MS results, and those of Laganowsky *et al.* [33]. This is consistent with known affinity between membrane proteins and negatively charged lipids [13, 14], and the expectations from the positive inside rule.

However there are some differences worthy of note. In contrast to CL, PG localization sites seem to be more evenly distributed around AqpZ tetramer, without a clear specific location. In contrast to PG, CL is closely associated with the monomer-monomer interface. Indeed occasionally the CL molecules manage to penetrate into the interface. The major difference observed in the case of the $W_{14}A$ mutant is a deeper penetration of the CL molecules into this interface. Indeed, in the case of the $W_{14}A$ mutant, this deep penetration of lipids into the interface, resulted in a much reduced dissociation constant for bound phospholipids, which has as a consequence the observed assymmetry in figure 5, due to slower equilibration.

Thus the molecular dynamics simulations indicate that close to the surface of AqpZ there is an accumulation of negatively charged lipids, particularly on the cytoplasmic membrane leaflet. There appears to be relatively specific binding site for CL molecules at the interface between monomers, partially excluding other lipid species. This location for a specific binding site is entirely consistent with our MS observations presented above. The selection of CL appears to be stronger and more extensive in the $W_{14}A$ mutant.

Molecular dynamics simulations in the absence of CL show that the interfacial binding site does not bind either PG and PE particularly strongly in a membrane context. This appears to be at odds with our MS results, where we observe mutation sensitive binding sites for these lipids. We believe that this can be attributed to the differences in competition between lipids, where head-group chemistry is dominant, and lipid - detergent competition where hydrocarbon chain interactions play a larger role.

4. Discussion

Is it possible to integrate these results to gain a clearer picture of how AqpZ interacts with its lipid environment both in the native membrane and in lipid nanodiscs? To integrate our different results it is important to understand the differences between the systems studied. SMALP's are composed of an anionic polymer surrounding a mixture of lipid and protein, our analysis shows that these particles contain about 36% protein, 50% lipid and 14% polymer, by weight (supplementary table S1). The SMALP thus contains both protein bound lipids, and the lipid annulus, the latter possibly perturbed by the polymer. In contrast, in native MS we observe individual bound lipids that resist dissociation by competition with solvating detergent molecules, a situation different from the native one where the competition is with solvating lipid molecules in the annulus. Finally, our molecular dynamics simulations try to

mimic the native membrane, and thus include neither perturbations to the lipid annulus by polymer nor indications if bound lipids are detergent resistant.

Intact MS, as mentioned above, measures the preferences for lipids over the detergent that they displace. In the experiments we have reported here we show, in agreement with previous work [33], a high-affinity site for CL on AqpZ. The reduction of affinity in the $W_{14}A$ mutant and the molecular dynamics simulations both suggest that this binding site is at the monomer-monomer interface of the AqpZ tetramer. Functional measurements [33] suggest that CL binding to this site is important for the permeability of the AqpZ tetramer, and reinforce the idea that the tetrameric structure is important for activity perhaps due to dynamic considerations [32].

Our analysis of AqpZ containing SMALP's gives a lipid-:protein (wt/wt) ratio of about 1.5, however, in the bacterial cell cytoplasmic membrane the lipid:protein (wt/wt) ratio is typically considerably lower, 0.33 [63]. Thus, these nanodiscs are significantly lipid enriched. This lipid enrichment associated with SMA solubilization is entirely consistent with the recent observations of improved solubilization with lipid supplementation [64]. Surprisingly we found no evidence for lipid sorting based on chain lengths in the lipidomics analysis, despite the expected hydrophobic mismatch between AqpZ and the phospholipid membrane [25]. We believe that this is because the entropic cost of such sorting is too high in the context of a small protein in a membrane with mostly fluid unsaturated chains but could also be the result of membrane structural perturbations by the polymer.

We calculate that a typical nanodisc containing AqpZ will contain about 140 lipid molecules, thus specifically bound lipids are likely to be a minor component of the SMALP. Our calculated composition suggests that there are on average 4-5 CL molecules par nanodisc, this is sufficient to satisfy the specific binding of 1 CL per AqpZ monomer. However this leaves the lipid annulus in AqpZ containing SMALP's strongly depleted in CL at the expense of PE. This is in contrast to the previously observed enrichment of CL in SMALP's prepared with KcsA [13] or the SecYEG translocon [14], while no modification in lipid composition was observed for SecDE or YidC [14]-containing SMALP's. We note that these experiments were performed using a slightly less hydrophobic polymer, with a 2:1 rather than 3:1 Styrene:Maleic acid ratio. These observations can be explained in the context of dynamic exchange between SMALP's. Indeed, since their first use for solubilizing membrane proteins it has been argued that SMA provide a snapshot of the biological membrane around a protein. However, recent studies have modified this vision by showing the existence of rapid lipid exchange between SMALP's [65], and SMA exchange between SMALP's [58]. The presence of this dynamic exchange would suggest that the composition of SMALP's reflects not the immediate membrane surrounding proteins *in vivo*, but rather a re-equilibrated system after exchange of lipids between nanodiscs.

After membrane solubilization, before purification of AqpZ containing SMALP's, the composition of the lipids solubilized reflects that of the membrane as we show in figure 2. Furthermore, there is ample evidence that SMA needs lipids to solubilize membrane proteins [64]. Our vision of a SMALP is thus a central protein, with some bound lipid, a lipid barrier between the protein and the polymer, and then an external polymer belt in contact with some lipids. The overall composition of the SMALP will therefor depend on the composition and quantity of these different types of lipid. The molecular dynamics simulations suggest that anionic lipids are favored close to the protein, in particular with bound CL, supplementary table S3. It seems likely that cationic or zwitterionic lipids are favored in contact with the polymer. Therefor the belt of lipids, which are free to exchange between nanodiscs containing different proteins, is depleted of anionic lipids, and in particular CL. This depletion could be driven, at least in part, by electrostatic effects accompanying mixing of the polymer with the annular lipids. This mixing effect is expected to be slightly larger with the 3:1 copolymer that we have used than the 2:1 copolymer used previously, since the 3:1 polymer mixes better with lipids than the 2:1 polymer [66]. Thus KcsA and the SecYEG translocon would be expected to bind much more CL than AqpZ, while SecDE and YidC would be expected to bind slightly more CL than AqpZ. In this respect, it is interesting to note that the cytoplasmic loops of AqpZ contain 8 positive residues for 6 transmembrane helices, a ratio of 1.3. The corresponding ratios for KcsA, SecYEG, SecDF and YidC are much higher, 7.5, 2.7, 2.4 and 3.2 respectively, suggesting CL enrichment in SMALP's is related to positive charge density at the membrane interface. Overall, our different observations, and those of others [13, 14], can be reconciled in the context of a re-equilibrated system [65, 58] by suggesting that most of the CL in the SMALP's is bound to proteins. Possibly driven to associate with the most basic proteins due to mixing of the annular lipids and the anionic SMA.

In conclusion, our analysis of lipids binding to AqpZ tetramers show a specific binding site at the interface between tetramers on the cytoplasmic surface of the membrane that has a preference for CL molecules. We suggest that binding of CL to this site is important for activity. The differences that we observe between CL accumulation around the protein in molecular dynamics simulations, and the depletion we observe in lipidomic analysis of SMALP's we attribute to low levels of free CL in the membrane with a significant proportion of the CL molecules being bound to specific high-affinity sites on membrane proteins. Surprisingly we fail to observe accumulation of short chain fatty acids close to the protein despite the predicted hydrophobic mismatch, suggesting that limiting the configurational space of the fluid lipids is less energetic than sorting the lipids in the membrane.

Acknowledgements

The authors thank V. Prima, S. Robert, H. Gaussier and the Proteomics plateform of the IMM for their aid and assistance. This work was financially supported by the CNRS, CINES and Espoir contre la Mucoviscidose.

References

[1] J. P. Overington, B. Al-Lazikani, A. L. Hopkins, How many drug targets are there?, Nature reviews. Drug discovery 5 (2006) 993–6.

[2] W. Van Klompenburg, I. Nilsson, G. von Heijne, B. De Kruijff, Anionic phospholipids are determinants of membrane protein topology, EMBO Journal 16 (14) (1997) 4261–4266. doi:10.1093/emboj/16.14.4261.

[3] G. von Heijne, Control of topology and mode of assembly of a polytopic MP by positively charged residues, Nature 342 (1989) 189–92. doi:10.1038/340301a0.

[4] S. Seppälä, J. S. Slusky, P. Lloris-Garcerá, M. Rapp, G. von Heijne, Control of membrane protein topology by a single c-terminal residue, Science 328 (5986) (2010) 1698–1700. doi:10.1126/science.1188950.

[5] C. Lange, J. H. Nett, B. L. Trumpower, C. Hunte, Specific roles of protein-phospholipid interactions in the yeast cytochrome bc 1 complex structure, EMBO Journal 20 (23) (2001) 6591–6600.

[6] E. Breukink, N. Nouwen, A. Van Raalte, S. Mizushima, J. Tommassen, B. De Kruijff, The c terminus of seca is involved in both lipid binding and secb binding, Journal of Biological Chemistry 270 (14) (1995) 7902–7907. doi:10.1074/jbc.270.14.7902.

[7] M. Opekarová, W. Tanner, Specific lipid requirements of membrane proteins - a putative bottleneck in heterologous expression, BBA-Biomembranes 1610 (1) (2003) 11–22. doi:10.1016/S0005-2736(02)00708-3.

[8] A. G. Lee, Lipid-protein interactions., Biochemical Society Transaction 39 (3) (2011) 761–6. doi:10.1042/BST0390761.

[9] M. Casiraghi, M. Damian, E. Lescop, E. Point, K. Moncoq, N. Morellet, D. Levy, J. Marie, E. Guittet, J.-L. Banères, L. J. Catoire, Functional modulation of a g protein-coupled receptor conformational landscape in a lipid bilayer, Journal of the American Chemical Society 138 (35) (2016) 11170–11175.

[10] M. Jamshad, V. Grimard, I. Idini, T. J. Knowles, M. R. Dowle, N. Schofield, S. Pooja, Y. Lin, R. Finka, M. Wheatley, O. R. T. Thomas, R. E. Palmer, M. Overduin, C. Govaerts, J.-M. Ruysschaert, K. J. Edler, T. R. Dafforn, Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins, Nano Research 8 (1995) 774–789.

[11] S. Paulin, M. Jamshad, T. R. Dafforn, O. Garcia-Lara, S. J. Foster, N. F. Galley, D. I. Roper, H. Rosado, P. W. Taylor, Surfactant-free purification of membrane protein complexes from bacteria: application to the staphylococcal penicillin-binding protein complex pbp2/pbp2a, Nanotechnology 25.

[12] M. Wheatley, J. Charlton, M. Jamshad, S. J. Routledge, S. Bailey, P. J. La-Borde, M. T. Azam, R. T. Logan, R. M. Bill, T. R. Dafforn, D. R. Poyner, GPCR-styrene maleic acid lipid particles (GPCR-SMALPs): their nature and potential., Biochemical Society transactions 44 (2) (2016) 619–623. doi:10.1042/BST20150284.

[13] J. M. Dorr, M. C. Koorengevel, M. Schäfer, A. V. Prokofyev, S. Scheidelaar, E. A. W. van der Cruijsen, T. R. Dafforn, M. Baldus, J. A. Killian, Detergent-free isolation, characterization, and functional reconstitution of a tetrameric k+ channel: The power of native nanodiscs, PNAS 111 (2014) 18607–18612.

[14] I. Prabudiansyah, I. Kusters, A. Caforio, A. J. Driessen, Characterization of the annular lipid shell of the sec translocon, BBA-Biomembranes 1848 (2015) 2050–2056.

[15] J. Frauenfeld, J. Gumbart, E. O. V. D. Sluis, S. Funes, B. Beatrix, T. Mielke, O. Berninghausen, K. Schulten, R. Beckmann, Cryo–EM structure of the ribosome SecYE complex in the membrane environment, Nature Structural and Molecular Biology 18 (5) (2011) 614–621. doi:10.1038/nsmb.2026.Cryo.

[16] Y. Gao, E. Cao, D. Julius, Y. Cheng, S. Francisco, S. Francisco, S. Francisco, Trpv1 structures in nanodiscs reveal mechanisms of ligand and lipid action, Nature 534 (5) (2016) 347–351. doi:10.1038/nature17964.TRPV1.

[17] T. Gonen, Y. Cheng, P. Sliz, T. Hiroaki, Y. Fujiyoshi, S. C. Harrison, T. Walz, Lipid-protein interactions in double-layered two-dimensional aqp0 crystals, Nature 438 (2005) 633–638. doi:10.1038/nm1636.Mutational.

[18] H. Wang, J. Elferich, E. Gouaux, Structures of leut in bicelles define conformation and substrate binding in a membrane-like context, Nature Structural and Molecular Biology 19 (2) (2012) 212–219. doi:10.1038/nsmb.2215.

[19] C. Bechara, C. V. Robinson, Different modes of lipid binding to membrane proteins probed by mass spectrometry, Journal of the American Chemical Society 137 (16) (2015) 5240–5247.

[20] M. Zhou, N. Morgner, N. P. Barrera, A. Politis, S. C. Isaacson, D. Matak-Vinkovic, T. Murata, R. A. Bernal, D. Stock, C. V. Robinson, Mass Spectrometry of Intact V-Type ATPases Reveals Bound Lipids and the Effects of Nucleotide Binding, Science 334 (6054) (2011) 380–385. doi:10.1126/science.1210148.

[21] C. Bechara, A. Nöll, N. Morgner, M. T. Degiacomi, R. Tampe, C. V. Robinson, A subset of annular lipids is linked to the flippase activity of an abc transporter, Nature Chemistry 7 (3) (2015) 255–262. doi:10.1038/nchem.2172.

[22] J. Gault, J. A. C. Donlan, I. Liko, J. T. S. Hopper, K. Gupta, N. G. Housden, W. B. Struwe, M. T. Marty, T. Mize, C. Bechara, Y. Zhu, B. Wu, C. Kleanthous, M. Belov, E. Damoc, A. Makarov, C. V. Robinson, High-resolution mass spectrometry of small molecules bound to membrane proteins, Nature Methods 13 (4) (2016) 333–336. doi:10.1038/nmeth.3771.

[23] X. Cong, Y. Liu, W. Liu, X. Liang, D. H. Russell, A. Laganowsky, Determining Membrane Protein-Lipid Binding Thermodynamics Using Native Mass Spectrometry, JACS 138 (13) (2016) 4346–4349. doi:10.1021/jacs.6b01771.

[24] J. Marcoux, S. C. Wang, A. Politis, E. Reading, J. Ma, P. C. Biggin, M. Zhou, H. Tao, Q. Zhang, G. Chang, N. Morgner, C. V. Robinson, Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump, PNAS 110 (24) (2013) 9704–9709. doi:10.1073/pnas.1303888110.

[25] J. P. Duneau, J. Khao, J. N. Sturgis, Lipid perturbation by membrane proteins and the lipophobic effect, BBA-Biomembranes 1859 (1) (2017) 1–5. doi:10.1016/j.bbamem.2016.10.014.

[26] C. Arnarez, J.-P. Mazat, J. Elezgaray, S.-J. Marrink, X. Periole, Evidence for cardiolipin binding sites on the membrane-exposed surface of the cytochrome bc1, JACS 135 (8) (2013) 3112–3120. doi:10.1021/ja310577u.

[27] F. J.-M. de Meyer, M. Venturoli, B. Smit, Molecular simulations of lipid-mediated protein-protein interactions, Biophysical Journal 95 (2008) 1851–1865. doi:10.1529/biophysj.107.124164.

[28] G. van den Bogaart, K. Meyenberg, H. J. Risselada, H. Amin, K. I. Willig, B. E. Hubrich, M. Dier, S. W. Hell, H. Grubmüller, U. Diederichsen, R. Jahn, Membrane protein sequestering by ionic protein lipid interactions, Nature 479 (7374) (2011) 552–555. doi:10.1038/nature10545.

[29] G. M. Preston, P. Agre, Isolation of the cdna for erythrocyte integral membrane protein of 28 kilodaltons: member of an ancient channel family, PNAS 88 (1991) 11110–11114.

[30] J. S. Jung, G. M. Preston, B. L. Smith, W. B. Gugginoll, P. Agre, Molecular structure of the water channel through aquaporin chip, Journal of Biological Chemistry 269 (1994) 14648–14654.

[31] M. Borgnia, D. Kozono, G. Calamita, P. C. Maloney, P. Agre, Functional reconstitution and characterization of aqpz, the *E. coli* water channel protein, Journal of molecular biology 291 (52) (1999) 1169–1179.

[32] V. Schmidt, J. N. Sturgis, Making monomeric aquaporin z by disrupting the hydrophobic tetramer interface, ACS Omega 2 (6) (2017) 3017–3027.
[33] A. Laganowsky, E. Reading, T. M. Allison, M. B. Ulmschneider, M. T. Degiacomi, A. J. Baldwin, C. V. Robinson, Membrane proteins bind lipids selectively to modulate their structure and function, Nature 510 (2014) 172–175.
[34] M. T. Marty, K. K. Hoi, J. Gault, C. V. Robinson, Probing the lipid annular belt by gas-phase dissociation of membrane proteins in nanodiscs, Angewandte Chemie International Edition 55 (2) (2016) 550–554. doi:10.1002/anie.201508289.
[35] B. K. Tan, M. Bogdanov, J. Zhao, W. Dowhan, C. R. H. Raetz, Z. Guan, Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates, PNAS 109 (41) (2012) 16504–16509. doi:10.1073/pnas.1212797109.
[36] E. G. Bligh, W. J. Dyer, A rapid method of total lipid extraction and purification., Can J Biochem Physiol 37 (8) (1959) 911–7. doi:10.1139/o59-099.
[37] S. Czolkoss, C. Fritz, G. Hölzl, M. Aktas, Two Distinct Cardiolipin Synthases Operate in Agrobacterium tumefaciens., PLoS ONE 11 (7) (2016) e0160373. doi:10.1371/journal.pone.0160373.
[38] C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis., Nature methods 9 (7) (2012) 671–5.
[39] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, Nature Methods 9 (7) (2012) 676–682. doi:10.1038/nmeth.2019.
[40] G. Vial, M.-A. Chauvin, N. Bendridi, A. Durand, E. Meugnier, A.-M. Madec, N. Bernoud-Hubac, J.-P. Pais de Barros, É. Fontaine, C. Acquaviva, S. Hallakou-Bozec, S. Bolze, H. Vidal, J. Rieusset, Imeglimin normalizes glucose tolerance and insulin sensitivity and improves mitochondrial function in liver of a high-fat, high-sucrose diet mice model, Diabetes 64 (6) (2015) 2254–2264. doi:10.2337/db14-1220.
[41] M. Baarine, P. Andréoletti, A. Athias, T. Nury, A. Zarrouk, K. Ragot, A. Vejux, J.-M. Riedinger, Z. Kattan, G. Bessede, D. Trompier, S. Savary, M. Cherkaoui-Malki, G. Lizard, Evidence of oxidative stress in very long chain fatty acid – treated oligodendrocytes and potentialization of ros production using rna interference-directed knockdown of abcd1 and acox1 peroxisomal proteins, Neuroscience 213 (2012) 1–18.
[42] M. M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding., Anal. Biochem. 72 (1976) 248–54.
[43] P. S. Chen, T. Y. Toribara, H. Warner, Microdetermination of phosphorus., Analytical Chemistry 28 (9) (1956) 1756–1758.
[44] M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. A. Hochberg, L. P. Justin, C. V. Robinson, Bayesian deconvolution of mass and ion mobility spectra : From binary interactions to polydisperse ensembles, Analytical Chemistry 87 (8) (2015) 4370–4376. doi:10.1021/acs.analchem.5b00140.Bayesian.
[45] M. Kito, S. Aibara, M. Kato, T. Hata, Differences in fatty acid composition among phosphatidylethanolamine, phosphatidylglycerol and cardiolipin of *Escherichia coli*, BBA-Lipids and Lipid Metabolism 260 (3) (1972) 475–478. doi:10.1016/0005-2760(72)90062-8.
[46] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, The MARTINI force field: Coarse grained model for biomolecular simulations, Journal of Physical Chemistry B 111 (27) (2007) 7812–7824. doi:10.1021/jp071097f.
[47] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, S. J. Marrink, The MARTINI coarse-grained force field: Extension to proteins, Journal of Chemical Theory and Computation 4 (5) (2008) 819–834. doi:10.1021/ct700324x.
[48] M. Dahlberg, Polymorphic phase behavior of cardiolipin derivatives studied by coarse-grained molecular dynamics, Journal of Physical Chemistry B 111 (25) (2007) 7194–7200. doi:10.1021/jp071954f.
[49] T. A. Wassenaar, H. I. Ingolfsson, R. A. Bockmann, D. P. Tieleman, S. J. Marrink, Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations, Journal of Chemical Theory and Computation 11 (5) (2015) 2144–2155. doi:10.1021/acs.jctc.5b00209.
[50] D. H. de Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. a. Wassenaar, L. V. Schäfer, X. Periole, D. P. Tieleman, S. J. Marrink, Improved Parameters for the Martini Coarse-Grained Protein Force Field, Journal of Chemical Theory and Computation 9 (1) (2013) 687–697. doi:10.1021/ct300646g.
[51] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation, Computer Physics Communications 91 (1-3) (1995) 43–56. doi:10.1016/0010-4655(95)00042-E.
[52] M. J. Abraham, T. Murtola, R. Schulz, S. Pall, J. C. Smith, B. Hess, E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19–25. doi:10.1016/j.softx.2015.06.001.
[53] D. H. De Jong, S. Baoukina, H. I. Ingólfsson, S. J. Marrink, Martini straight: Boosting performance using a shorter cutoff and GPUs, Computer Physics Communications 199 (2016) 1–7. doi:10.1016/j.cpc.2015.09.014.
[54] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity-rescaling, The Journal of chemical physics 126 (1) (2008) 14101–14107. doi:10.1063/1.2408420.
[55] M. Parrinello, a. Rahman, Polymorphic Transitions in Single Crystals: a New Molecular Dynamics Method., Journal of Applied Physics 52 (12) (1981) 7182–7190. doi:10.1063/1.328693.
[56] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, Journal of Molecular Graphics 14 (1) (1996) 33–38. doi:10.1016/0263-7855(96)00018-5.
[57] J. M. Dorr, S. Scheidelaar, M. C. Koorengevel, J. J. D. Pardo, M. Schäfer, C. A. van Walree, J. A. Killian, The styrene–maleic acid copolymer: A versatile tool in membrane research, European Biophysics Journal 45 (2016) 3–21.
[58] V. Schmidt, J. N. Sturgis, Modifying styrene-maleic acid copolymer for studying lipid nanodiscs., BBA-Biomembranes 1860 (3) (2018) 777–783.
[59] C. R. Raetz, Enzymology, genetics, and regulation of membrane phospholipid synthesis in Escherichia coli, Microbiological Reviews 42 (3) (1978) 614–659.
[60] Y.-Y. Chang, J. E. Cronan, Membrane cyclopropane fatty acid content is a major factor in acid resistance of *Escherichia coli*, Molecular Microbiology 33 (2) (1999) 249–259. doi:10.1046/j.1365-2958.1999.01456.x.
[61] D. W. Grogan, J. E. Cronan, Cyclopropane ring formation in membrane lipids of bacteria, Microbiology and Molecular Biology Reviews 61 (4) (1997) 429–441.
[62] M. Bogdanov, J. Xie, W. Dowhan, Lipid-protein interactions drive membrane protein topogenesis in accordance with the positive inside rule, Journal of Biological Chemistry 284 (15) (2009) 9637–9641. doi:10.1074/jbc.R800081200.
[63] T. de Vrije, R. L. de Swart, W. Dowhan, J. Tommassen, B. de Kruijff, Phosphatidylglycerol is involved in protein translocation across Escherichia coli inner membranes., Nature 334 (6178) (1988) 173–5. doi:10.1038/334173a0.
[64] D. J. Swainsbury, S. Scheidelaar, N. Foster, R. van Grondelle, J. A. Killian, M. R. Jones, The effectiveness of styrene—maleic acid (sma) copolymers for solubilisation of integral membrane proteins from sma accessible and sma resistant membranes, BBA-Biomembranes 1859 (10) (2017) 2133–2145. doi:10.1016/j.bbamem.2017.07.011.
[65] R. Cuevas Arenas, B. Danielczak, A. Martel, L. Porcar, C. Breyton, C. Ebel, S. Keller, Fast collisional lipid transfer among polymer-bounded nanodiscs, Scientific Reports 7 (2017) 45875. doi:10.1038/srep45875.
[66] A. Grethen, A. O. Oluwole, B. Danielczak, C. Vargas, S. Keller, Thermodynamics of nanodisc formation mediated by

styrene/maleic acid (2:1) copolymer., Scientific Reports 7 (2017) 115–117.

Supplementary Information

- Table S1: Analysis of the composition of AqpZ containing SMA nanodiscs.
- Table S2: Lipid composition of the coarse grained *E. coli* inner membrane model.
- Table S3: Compositional analysis around AqpZ in coarse grained molecular dynamics simulations.
- Figure S1: Intact mass spectra of AqpZ-C_{20}S and W_{14}A in DDM detergent.
- Figure S2: Plots of mole fractions and K_D fittings of lipid binding to both AqpZ mutants.

Component wt %	Sample 1	2	3	Average %
Protein	35	36	34	35.0±1.0
Lipid	48	50	47	48.3±1.5
SMA	16	14	19	16.3±2.5

Table S1: Compositional analysis of AqpZ containing SMA nanodiscs. Three different preprations of nanodiscs were analyzed. Protein and SMA concentrations were calculated from UV absorption spectra. Lipid concentration was measured as phosphate using a molybdenum blue phosphate assay.

Lipid type	Notation	Coarse grained acyl beads	Number of double bonds	Number of acyl carbons	%
PE	PE8:0	8	0	32—36	6.2
PE	PE8:1	8	1	32—36	50.2
PE	PE8:2	8	2	32—36	12.5
PE	PE6:0	6	0	24—28	6.2
PG	PG8:0	8	0	32—36	1.3
PG	PG8:1	8	1	32—36	12
PG	PG8:2	8	2	32—36	3.3
CL	CL20:4	20	4	64—72	3.7
CL	CL20:2	20	2	64—72	4.6

Table S2: Lipid composition of the coarse grained *E. coli* inner membrane model. Acyl beads take into account the overall number of acyl beads of the phospholipid. Unsaturations represents the total number of unsaturations of the phospholipid including all acyl chains.

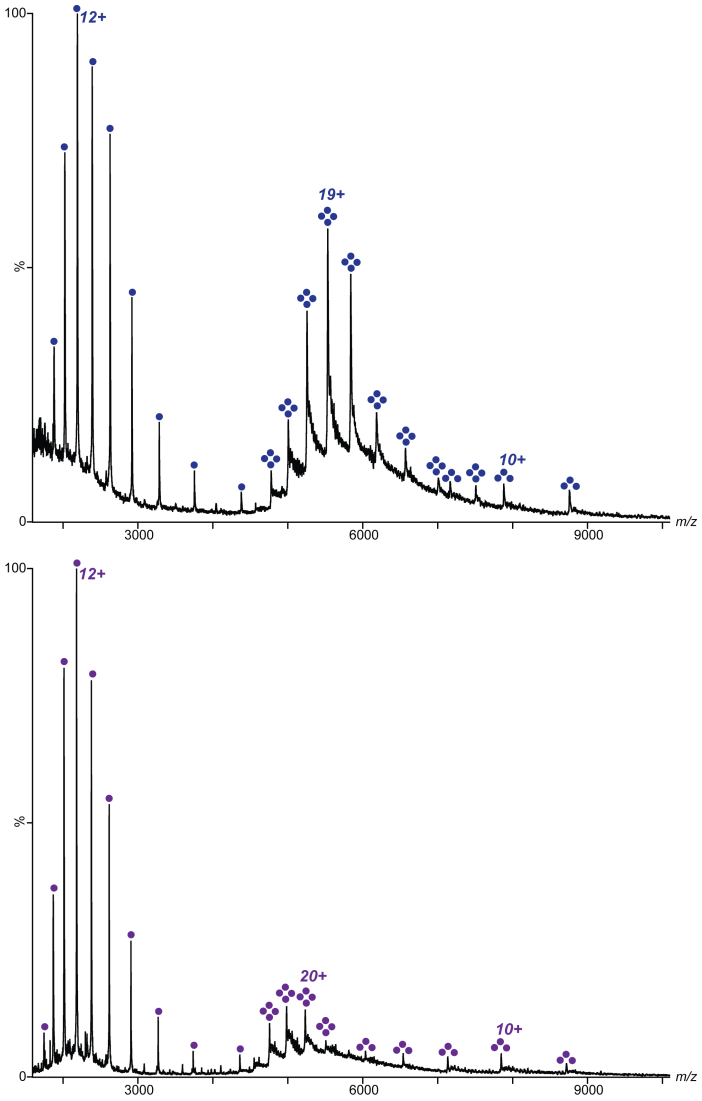

Figure S1: Intact mass spectra of AqpZ-C_{20}S (top) and W_{14}A (bottom) in DDM detergent, revealing tetramers, stripped trimers and monomers in both cases. Protein concentration was 7 μM, in 200 mM ammonium acetate buffer pH 8 supplemented with 0.02% DDM. The lowest collision voltage needed to detect the tetramers (80 V) was used.

Component	Distance Å			
wt %	6.0	12.0	20.0	∞
PE (WT)	61.6	66.0	67.2	69.3
PG (WT)	18.1	17.5	17.1	15.2
CL (WT)	20.3	16.5	15.7	15.5
Number (WT)	58.1	116.5	181.7	na
PE (W_{14}A)	61.1	64.3	67.3	69.3
PG (W_{14}A)	17.4	16.5	16.6	15.2
CL (W_{14}A)	21.4	19.2	16.1	15.5
Number (W_{14}A)	57.9	118.8	180.4	na

Table S3: Compositional analysis around AqpZ in coarse grained molecular dynamics simulations. Lipids falling at less than the specific distanced from the protein molecule were sorted by type and counted. Average values for the composition at different distances from the protein were determined. To allow direct comparison with the experimental data (supplementary table S1) the values have been converted to wt %, i.e. taking into account that CL is about twice as large as PE or PG. Also shown is the average number of lipid molecules within the shells.

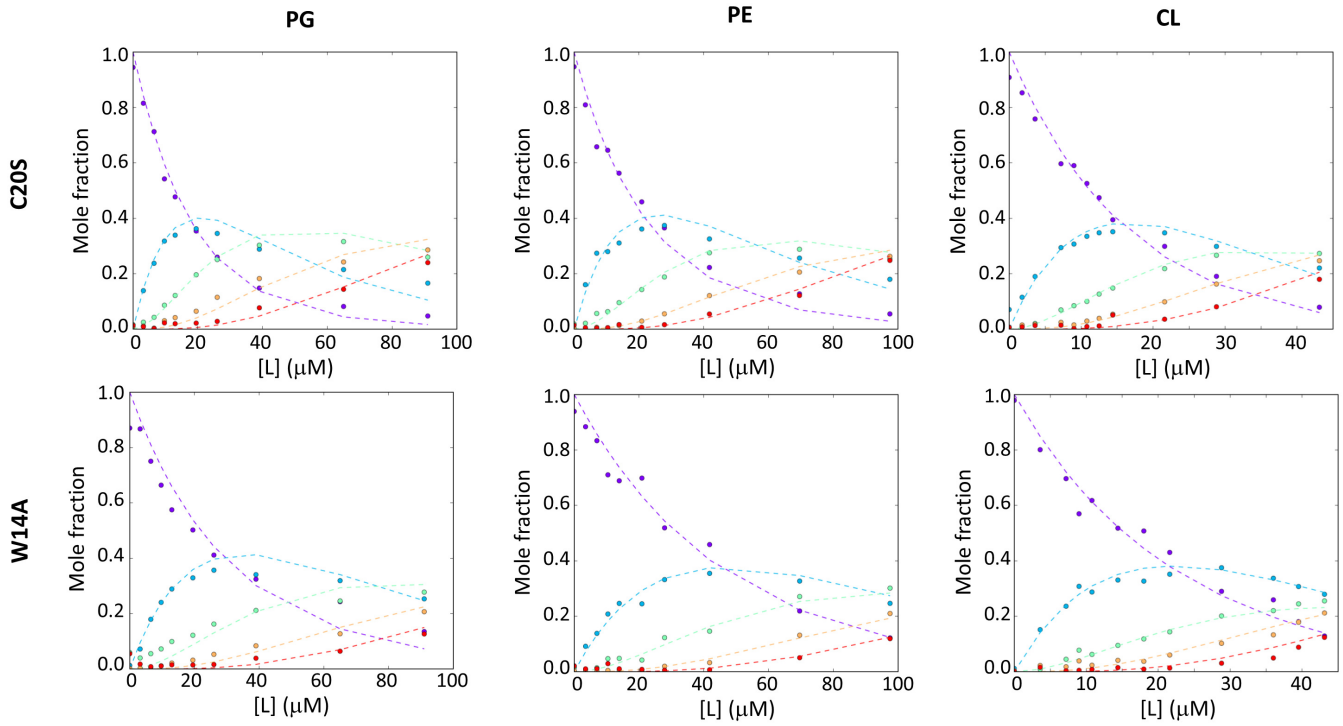

Figure S2: Plots of mole fractions and K_D fittings of lipid binding to both AqpZ mutants. In purple the *apo*-protein tetramer, blue, green, orange and red corresponds to the tetramer +1, +2, +3 and +4 lipids respectively. The best model in UniDec Data Collector module was used for all the fittings, corresponding to a sequential binding model with all free K_D values for each binding site.