Membrane proteins bind lipids selectively to modulate their structure and function

Arthur Laganowsky1*, Eamonn Reading1*, Timothy M. Allison1, Martin B. Ulmschneider2, Matteo T. Degiacomi3, Andrew J. Baldwin1 & Carol V. Robinson1

Previous studies have established that the folding, structure, and function of membrane proteins are influenced by their lipid environments8-10 and that lipids can bind to specific sites, for example, in potassium channels4. Fundamental questions remain however regarding the extent of membrane protein selectivity towards lipids. Here we report a mass spectrometry approach designed to determine the selectivity of lipid binding to membrane protein complexes. We investigate the mechanosensitive channel of large conductance (MscL) from Mycobacterium tuberculosis and aquaporin Z (AqpZ) and the ammonia channel (AmtB) from Escherichia coli, using ion mobility mass spectrometry (IM-MS), which reports gas-phase collision cross-sections. We demonstrate that folded conformations of membrane protein complexes can exist in the gas phase. By resolving lipid-bound states, we then rank bound lipids on the basis of their ability to resist gas phase unfolding and thereby stabilize membrane protein structure. Lipids bind non-selectively and with high avidity to MscL, all imparting comparable stability; however, the highest-ranking lipid is phosphatidylinositol phosphate, in line with its proposed functional role in mechanosensation8. AqpZ is also stabilized by many lipids, with cardioli pin imparting the most significant resistance to unfolding. Subsequently, through functional assays we show that cardiolipin modulates AqpZ function. Similar experiments identify AmtB as being highly selective for phosphatidylglycerol, prompting us to obtain an X-ray structure in this lipid membrane-like environment. The 2.3 Å resolution structure, when compared with others obtained without lipid bound, reveals distinct conformational changes that re-position AmtB residues to interact with the lipid bilayer. Our results demonstrate that resistance to unfolding correlates with specific lipid-binding events, enabling a distinction to be made between lipids that merely bind from those that modulate membrane protein structure and/or function. We anticipate that these findings will be important not only for defining the selectivity of membrane proteins towards lipids, but also for understanding the role of lipids in modulating protein function or drug binding.

Three membrane protein complexes were selected to give a range of topologies, oligomeric states and anticipated selectivity towards lipids: (1) the pentameric MscL from M. tuberculosis with two transmembrane helices per subunit and an intimate relationship with lipids10,11, (2) the tetrameric water efflux channel AqpZ from E. coli with six transmembrane helices for which associated lipid or detergent molecules have been revealed in crystal structures and in related homologues12,13, and (3) the trimeric AmtB from E. coli, with eleven transmembrane helices involved in the transport of ammonia, for which no lipid binding has been observed to date in crystal structures4.

Our first objective for studying membrane protein and lipid interactions by mass spectrometry was to maintain the folded state of these multimeric assemblies in the gas phase, devoid of detergent from the micelles required for their introduction14, such that stabilization by lipid binding could be deduced (Fig. 1a). This had not been possible previously as the conditions needed to disrupt the micelle perturbed the folded state of membrane protein complexes. We screened non-ionic detergents15 and employed ion mobility mass spectrometry (IM-MS) (Supplementary Video 1), a technique that measures the rotationally averaged collision cross-sections (CCS) of protein complexes in the gas-phase8. We found that in detergents commonly used for structural studies (n-dodecyl-β-D-maltoside (DDM), octyl glucose neopentyl glycol (OGNG) and n-nonyl-β-D-glucopyranoside (NG)), the CCS of all charge states were substantially greater when emerged from the micelle than those calculated for the crystal structure, indicative of gas-phase unfolding (Extended Data Fig. 1). Interestingly, for tetraethylene glycol monooyctyl ether (C8E4) lower average charge states were observed with CCS in agreement with those calculated for the crystal structures, thus providing optimal conditions for mass spectrometry of intact membrane protein complexes.

To assess individual lipid-binding events, we prepared a series of synthetic and natural lipid solutions and added these to protein complexes in C8E4 (Extended Data Fig. 2). MscL was tested first, revealing well-resolved mass spectral peaks corresponding to populations of MscL in complex with one to five lipids (Fig. 1b and Extended Data Fig. 3). Unlike other biological approaches, which report an average of all species in solution, whether bound or unbound, discrete lipid-bound states can be resolved and interrogated individually, allowing us to obtain CCS for individual lipid-binding events and to assess their effects on stability. Maintaining folded structures of these lipid-bound complexes enabled us to explore molecular dynamics simulations combined with filtering using CCS measurements16 to locate probable lipid-binding sites (Extended Data Figs 4 and 5 and Supplementary Discussion). We then applied collision-induced unfolding, analogous to that used previously for studying ligand binding to soluble complexes17. An increase in CCS, from the trend line calculated for the folded state, is observed with increasing collision voltage and is assigned to highly extended conformers (Fig. 1b and Supplementary Video 2). All species, apo and one to five phospholipid-bound forms, are measured within the same experiment enabling a direct comparison, and show less unfolding of native pentameric states as a function of lipid binding.

To extract quantitative values for the effects of lipid binding, we applied an equilibrium unfolding model, used extensively in solution studies involving chemical denaturation18,19. Plotting CCS against collision voltage reveals unfolding trajectories from the native state, through intermediates, to extended forms. Importantly, the oligomeric state is maintained throughout the unfolding experiment (Extended Data Fig. 6 and Fig. 1c). We identified four distinct intermediate states in the unfolding trajectories and by comparing the apo form with lipid-bound states we calculated the stabilization imparted at the transitions by each lipid (see Methods). The stabilization computed for bound lipids allows us to compare and rank seven different lipids, as well as synthetic phosphatidylcholine (PC) lipids with chain lengths varying from C14 to C24. Despite differences in these lipid structures, we found that all stabilized MscL to a similar degree, consistent with a previous fluorescence study with engineered Trp residues20 (Fig. 1d and Extended Data Fig. 3b,c). We also found that MscL avidly bound phosphatidylinositol phosphate

1Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 5QY, UK. 2Department of Materials Science and Engineering, Johns Hopkins University, Baltimore, Maryland 21218, USA. *These authors contributed equally to this work.

©2014 Macmillan Publishers Limited. All rights reserved
AqpZ is indiscriminately stabilized by lipids with the exception of cardiolipin (CDL), for which binding resulted in a substantial increase in protein stability compared to all other lipids. Intrigued by this observation, we performed functional assays of the channel reconstituted into liposomes of varying lipid composition. Water permeability assays of AqpZ reconstituted in a total polar lipid extract from a CDL-deficient E. coli strain (BKT22) showed a striking reduction in water transport compared to the channel reconstituted in wild-type E. coli polar lipid preparations (EPL) (Fig. 2c and Extended Data Fig. 8). Moreover, reconstituting AqpZ in a lipid preparation from strain BKT22 expressing cardiolipin synthase genes, restores CDL to near wild-type levels (mole percentage of 6.6 versus 9.1), and results in water transport

Figure 2 | AqpZ is indiscriminately stabilized by lipids with the exception of cardiolipin, a lipid that stabilizes the channel significantly and modulates its function. a, Mass spectra reveal AqpZ(POPC)_{10–5} resolved at 60 V with IM-MS arrival times in agreement with the CCS calculated from the crystal structure. b, Stabilization of AqpZ (+13) bound to lipids, calculated from unfolding parameters. Shown as described in Fig. 1d. One-way ANOVA (n = 3). **P < 0.01. c, Water permeability assay for AqpZ reconstituted in total polar lipid extracts from E. coli (EPL) (orange), or a cardiolipin-deficient strain (BKT22) (green), or BKT22 cells expressing ClsC and YmdB to restore cardiolipin (BKT22-YC) (cyan), compared with empty EPL liposomes (pink) (Extended Data Fig. 8). Rate constants of water transport (k_{wat}) and standard error of replicates (n = 5).
activity comparable to in EPL (Fig. 2c). These results clearly demonstrate CDL modulates the function of AqpZ.

Analogous IM-MS experiments carried out on AmtB show that resistance to gas-phase unfolding following binding of anionic lipids, phosphatidic acid (PA) and phosphatidylserine (PS), is not correlated with significant stabilization of the channel (Fig. 3). Zwitterionic lipids, PC and phosphatidylethanolamine (PE), confer only moderate stabilization. By contrast, addition of CDL or phosphatidylglycerol (PG) results in striking increases in stabilization with cumulative binding of PG increasing protein stability linearly (Fig. 3b, Extended Data Fig. 7 and Supplementary Video 4). These results therefore reveal a defined selectivity for PG-like headgroups.

To understand the origin of this observed selectivity, we crystallized AmtB in the presence of PG. A screen produced X-ray-grade crystals for PG-like headgroups. AmtB in the presence of PG showed a non-native, trimeric structure. These results therefore reveal a selectivity for PG-like headgroups.

To understand the origin of this observed selectivity, we crystallized AmtB in the presence of PG. A screen produced X-ray-grade crystals for PG-like headgroups. AmtB in the presence of PG showed a non-native, trimeric structure. These results therefore reveal a selectivity for PG-like headgroups.

To understand the origin of this observed selectivity, we crystallized AmtB in the presence of PG. A screen produced X-ray-grade crystals for PG-like headgroups. AmtB in the presence of PG showed a non-native, trimeric structure. These results therefore reveal a selectivity for PG-like headgroups.

By resolving individual lipid-bound states of membrane protein complexes within the mass spectrometer, and interrogating their stability through gas-phase unfolding, we obtained quantitative values to rank lipid-binding interactions. MscL binds lipids non-selectively, without regard to particular headgroup or chain length, and binding of any lipid imparts comparable stability, in accord with the ability of the channel to bind and respond promiscuously to lipid composition. The most significant cumulative effects on stability were observed upon binding of PI, proposed as the functional lipid in the Mycobacterium membrane involved in MscL mechanosensitivity. Similarly, AqpZ was found to be non-selective for a number of lipids suggesting comparable binding modes, in line with the two-dimensional crystals of Aqp0. The only exception to this non-selective lipid binding is CDL, which significantly stabilizes AqpZ and directly modulates function, as suggested in a previous in vivo study. Interestingly, CDL levels in E. coli depend on growth conditions, a higher abundance of CDL being present in the stationary phase, implying AqpZ function is fine-tuned through modulation of CDL binding.
CDL. In contrast to the other two systems, our finding that AmtB responds only weakly to the majority of lipids is consistent with its extensive transmembrane structure and inherent stability devoid of the lipid bilayer. However, its unexpectedly high selectivity for PG led to the first crystal structure of AmtB in complex with lipid, allowing effects on stability to be rationalized and further validated by mutation of residues forming the specific lipid-binding site.

By investigating a diverse set of membrane proteins and their lipid-binding properties we have extended many seminal studies1–4 by demonstrating different degrees of selectivity in lipid binding with AmtB > AqpZ > MscL. Despite these differences in selectivity, in all cases the most stabilizing lipids have a direct influence on the structure and function of the membrane proteins studied, as proposed for PI binding to MscL5 and demonstrated for CDL and PG binding to AqpZ and AmtB, respectively. As membrane proteins are intimately embedded in the bilayer, the relative lipid abundances, as well as differences in selectivity, can enforce the recruitment of a local lipid environment6, thus providing an elegant means for fine-tuning membrane protein structure and function.

METHODS SUMMARY

Detailed methods are described in Methods. In brief, membrane proteins were expressed as TEV protease cleavable fusion proteins containing a terminal His-tag in Escherichia coli. Membrane proteins were extracted from purified membranes, followed by affinity chromatography, TEV protease cleavage and gel filtration chromatography. Purified membrane complexes and lipids were prepared for IM-MS8 and demonstrated for CDL and PG binding to AqpZ and AmtB, respectively. As membrane proteins are intimately embedded in the bilayer, the relative lipid abundances, as well as differences in selectivity, can enforce the recruitment of a local lipid environment6, thus providing an elegant means for fine-tuning membrane protein structure and function.

Received 22 November 2013; accepted 28 April 2014.

1. Singer, S. J. & Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. Science 175, 720–731 (1972).
2. Cantor, R. S. The influence of membrane lateral pressures on simple geometric models of protein conformational equilibrium. Chem. Phys. Lipids 101, 45–56 (1999).
3. Hunte, C. & Richers, S. Lipids and membrane protein structures. Curr. Opin. Struct. Biol. 18, 406–411 (2008).
4. Lee, A. G. Biological membranes: the importance of molecular detail. Trends Biochem. Sci. 36, 493–500 (2011).
5. Sanders, C. R. & Mittendorf, K. F. Tolerance to changes in membrane lipid composition as a selected trait of membrane proteins. Biochemistry 50, 7858–7867 (2011).
6. Contreras, F. X., Ernst, A. M., Wieland, F. & Brugger, B. Specificity of intramembrane protein–lipid interactions. Cold Spring Harb. Perspect. Biol. 3, http://dx.doi.org/10.1101/cshperspect.a004705 (2011).
7. Whittle, J. P. Integral membrane proteins and bilayer proteomics. Anal. Chem. 85, 2658–2659 (2013).
8. Long, S. B., Tao, X., Campbell, E. B. & MacKinnon, R. Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. Nature 450, 376–382 (2007).
9. Zhong, D. & Blount, P. Phosphatidylglycerol is crucial for the mechanosensitivity of Mycobacterium tuberculosis Mscl. Biochemistry 52, 5415–5420 (2013).
10. Chang, G., Spencer, R. H., Lee, A. T., Barclay, M. T. & Rees, D. C. Structure of the Mscl homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel. Science 282, 2220–2226 (1998).
11. Haswell, E. S., Phillips, R. & Rees, D. C. Mechanosensitive channels: what can they do and how do they do it? Structure 19, 1356–1369 (2011).
12. Savage, D. F., Egea, P. F., Robles-Colmenares, Y., O’Connell, J. D. III & Stroud, R. M. Architecture and selectivity in aquaporins: 2.5A˚ X-ray structure of aquaporin Z. PNAS Biol. 1, 872 (2003).
13. Gonen, T. et al. Lipid–protein interactions in double-layered two-dimensional AQP0 crystals. Nature 438, 633–638 (2005).
14. Khademi, S. et al. Mechanism of ammonium transport by Amt/MEP/Rh: structure of AmtB at 3.5A˚. Science 305, 1587–1592 (2004).
15. Ferrera, N. P., Di Bartolo, N., Booth, P. J. & Robinson, C. V. Micelles protect membrane complexes from solution to vacuum. Science 321, 243–246 (2008).
16. Laganowsky, A., Reading, E., Hopper, J. T. & Robinson, C. V. Mass spectrometry of membrane protein complexes. Nature Protocols 8, 639–651 (2013).
17. Ruotolo, B. T. et al. Evidence for macromolecular protein rings in the absence of bulk water. Science 310, 1658–1661 (2005).
18. Baldwin, A. J. et al. The polydispersity of nα-crystallin is rationalized by an interconverting polyhedral architecture. Structure 19, 1855–1863 (2011).
19. Hyung, S. J., Robinson, C. V. & Ruotolo, B. T. Gas-phase unfolding and disassembly reveals stability differences in lipid-bound multiprotein complexes. Chem. Biol. 16, 382–390 (2009).
20. Pace, C. N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. Methods Enzymol. 131, 266–280 (1986).
21. Hong, H., Joh, N. H., Bowie, J. U. & Tamm, L. K. In Methods Enzymol Vol. 455 (eds Michael J. Johnson, Jo M. Holt, & Gary K. Ackers) 213–236 (Academic Press, 2009).
22. Powl, A. M., East, J. M. & Lee, A. G. Lipid-protein interactions studied by introduction of a tryptophan residue: the mechanosensitive channel Mscl. Biochemistry 42, 14306–14317 (2003).
23. Borgnia, M. J., Kozono, D., Caiarnita, G., Maloney, P. C. & Agre, P. Functional reconstitution and characterization of AqpZ, the e. coli water channel protein. J. Mol. Biol. 291, 1169–1179 (1999).
24. Tan, B. K. et al. Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. Proc. Natl Acad. Sci. USA 109, 16562–16567 (2012).
25. Yau, W. M., Wimley, W. C., Gavorch, K. & White, S. H. The preference of trytophan for membrane interfaces. Biochemistry 37, 14713–14718 (1998).
26. Isla, I. & Blount, P. Sensing and responding to membrane tension: the bacterial MscL channel as a model system. Biophys. J. 103, 169–174 (2012).
27. Hite, R. K., L.I. & Walz, T. Principles of membrane protein interactions with annular lipids deduced from aquaporin-0 2D crystals. EMBO J. 29, 1652–1658 (2010).
28. Romansov, T., Battle, A. R., Hendel, J. L., Martinac, B. & Wood, J. M. Protein localization in Escherichia coli cells: comparison of the cytoplasmic membrane proteins ProP, LacY, ProW, AqpZ, MscS, and MscL. J. Bacteriol. 192, 912–924 (2010).
29. Blaakey, D. et al. Purification of the Escherichia coli ammonium transporter AmtB reveals a trimeric stoichiometry. Biochem. J. 364, 527–535 (2002).
30. Engelmann, D. M. Membranes are more mosaic than fluid. Nature 438, 578–580 (2005).

Supplementary Information

is available in the version of the paper.

Acknowledgements We thank J. Hobman and D. Lee for providing gene-encoding plasmids, Z. Guan and C. Li for providing E. coli strains and plasmids containing cardiolipin genes. We also thank D. Stanton and N. Housden for training on the stopped-flow apparatus; and T. Mize, J. Benesch, M. McDonough, T. Walton and D. Rees for discussions. We also gratefully acknowledge E. Lowe and S. Lea for organizing synchrotron proposals and Diamond Light Source beam line I04 and staff, the Medical Research Council (MRC), BBSRC and ERC advanced grant (IMPRESS) for funding. A.L. is a Nicholas Kurti Junior Research Fellow of Brasenose College, A.J.B. is a BBSRC David Phillips’ Fellow and C.V.R. is a Royal Society Professor.

Author Contributions A.L., E.R., and C.V.R. designed the research. A.L. and E.R. performed the experiments. T.M.A. assisted A.L. and E.R. in protein expression and purification. M.B.U. carried out molecular dynamics. M.T.D., A.J.B. and A.L. designed and performed post-molecular dynamics analyses. A.L., A.J.B. and A.L. developed IM-MS analysis software. A.L. and E.R. analysed the data. A.L., E.R. and C.V.R. wrote the paper with input from all the authors.

Author Information Atomic coordinates and structure factors for the crystal structure have been deposited with the Protein Data Bank (PDB) under accession code 4NH2. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.L. (art.laganowsky@chem.ox.ac.uk) or C.V.R. (carol.robinson@chem.ox.ac.uk).
METHODS

Plasmid construction. Two expression plasmids were constructed by subcloning the multiple cloning site region from XbaI (New England Biolabs) and BlpI (New England Biolabs) of pET23b (Novagen) into the backbone pET15b (Novagen). The resulting engineered vector was linearized by Ndel (New England Biolabs) and XhoI (New England Biolabs), gel purified (QIAquick Gel Extraction Kit, Qiagen), and used in subsequent Infusion cloning reactions (Clonetics) to generate a TEV protease cleavable C-terminal fusion to superfolder GFP (subcloned from Gandhi et al.18) followed by a 6× His-tag or a TEV protease cleavable N-terminal fusion to maltose binding protein (MBP) preceded by a secretion signal peptide (pelB) and 10× His-tag subcloned from Hili et al.20. The resulting N- and C-terminal fusion vectors were linearized with Nhel and XhoI and Ndel and Nhel digested, purified, and used in subsequent Infusion cloning reactions. AmtB (residues 26–428), AgpZ and Mscl genes were amplified by polymerase chain reaction (PCR) with Phusion high-fidelity DNA polymerase (New England Biolabs) from prepared E. coli BL21 (DE3) genomic DNA (Qiagen) or template plasmid DNA with primers designed for an Infusion cloning reaction using the manufacturer’s online tool. The PCR products were purified by agarose gel electrophoresis and extracted using the QIAquick gel extraction kit (Qiagen). The PCR products and linearized vectors (described above) were used in Infusion cloning reactions (Clonetics) to generate MBP–AmtB, AgpZ–GFP and Mscl–GFP. MBP–AmpB27,28,29N91A was achieved by two--rounds of site-directed mutagenesis using a QuickChange Lightning Multi Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer’s protocol. AgpZ harboring the TEV protease cleavable N-terminal 6× His-tag (NHIS–AgpZ) was constructed through an infusion cloning reaction of AgpZ23 and a custom pRSFDeutel-1 vector (Novagen). The N-terminal coding sequence between NcoI and BamHI was replaced with the following nucleotide sequence: CCATGAGACGAGCAGCTACATCATGACCAAGAAGCTATTGCTACCGGTTGATCC, linearized with BamHI and XhoI. All constructs were verified by DNA sequencing.

Generation of E. coli BL21(DE3) Amscl:kanR strain. Homologous recombination mediated chromosomal disruption of mscl was performed using the Gene Doctoring method30. Briefly, 900 base pairs flanking the 5′ and 3′ ends of the mscl gene were subcloned into the pDOCK-K plasmid resulting in these regions flanking a kanamycin resistance (Kan′) cassette. This plasmid and the recombineering plasmid, pACBSCB, were transformed into E. coli BL21 (DE3) Gold (Agilent). Homologous recombination was carried out by induction of λ-Red recombinease system and SceI endonuclease with arabinose. Recombinants harbouring gene disruption with Kan′ were selected on kanamycin and sucrose, and verified by colony PCR. Calcium chloride competent cells were made as described by Drew et al.18.

Membrane protein expression. MBP–AmtB and AgpZ–GFP plasmids were transformed into E. coli BL21 (DE3) Gold (Agilent). MBP–AmpB27,28,29N91A was purified by agarose gel electrophoresis and extracted using the QIAquick gel extraction kit (Qiagen). The PCR products and linearized vectors (described above) were used in Infusion cloning reactions (Clonetics) to generate MBP–AmtB, MBP–AmpB27,28,29N91A and AgpZ–GFP were extracted from purified membranes in buffer B supplemented with 200 mM OG and incubated with gentle agitation overnight at 4°C. Mscl was extracted from purified membranes with glycerol-free buffer B supplemented with 1% ONG. Extracted membrane proteins were clarified by centrifugation at 20,000g for 20 min at 4°C. Supernatant was filtered before loading onto a 5 ml HiTrap-HP column (GE Healthcare, Piscataway, NJ) equilibrated in buffer E (200 mM sodium chloride, 10% glycerol, 20 mM imidazole, 0.025% DDM and 50 mM Tris, pH 7.4 at room temperature). After the clarified supernatant was loaded, the column was initially washed with 40–50 ml of DDM-free buffer E supplemented with 1% OG for MBP–AmtB and AgpZ–GFP and 0.5% ONG for Mscl–GFP. Membrane proteins were then exchanged into several column volumes of buffer E until a steady baseline was reached. Membrane protein fractions were eluted with a linear gradient to 100% in two column volumes of buffer F (100 mM sodium chloride, 10% glycerol, 500 mM imidazole, 0.025% DDM and 50 mM Tris, pH 7.4 at room temperature). Peak fractions were pooled, supplemented with 5 mM BME and His-tagged TEV protease32,33, and dialysed against buffer G (150 mM sodium chloride, 10% glycerol, 20 mM imidazole, 0.025% DDM and 50 mM Tris, pH 7.4 at room temperature) overnight at 4°C. After overnight incubation, samples were filtered and passed back over a 5 ml HiTrap-HP column equilibrated in buffer F. Flow-through containing the untagged membrane protein was collected and concentrated using a 100kDa MWCO concentration. Notably, tag-removed membrane proteins with fusions to the C terminus contained the additional protein sequence ASGENLYFQ, or GAS for N-terminally tagged proteins, resulting from the TEV protease recognition sequence and cloning restriction site. Concentrated protein was either used immediately or flash-frozen in liquid nitrogen and stored at −80°C. Protein concentration was measured using Biorad UV detector with the following calculated extinction coefficients of 1.37, 1.484, and 0.265 Mcm−1 for AmtB, AgpZ and Mscl, respectively. NHIS–AgpZ was purified as previously described18, Briefly, extracted and clari-fied membrane proteins were loaded onto a 5 ml HiTrap-HP column (GE Healthcare, Piscataway, NJ) equilibrated in buffer F (200 mM sodium chloride, 10% glycerol, 20 mM imidazole, 1% OG and 50 mM Tris, pH 7.4 at room temperature) over two column volumes. Peak fractions containing NHIS–AgpZ were concentrated before imidazole removal using two 5 ml HiTrap desalting columns (GE Healthcare, Piscataway, NJ) in tandem. Peak fractions were pooled and NHIS–AgpZ was concentrated using a 50 kDa MWCO concentrator.

Preparation of membrane proteins for native mass spectrometry. Either flash-frozen samples thawed on ice or fresh samples were detergent exchanged by gel filtration chromatography. Membrane protein samples were suspended in 200 GL 10/300 (GE Healthcare) column equilibrated in buffer H (130 mM sodium chloride, 10% glycerol and 50 mM Tris, pH 7.4 at room temperature) supplemented with either 0.5% of C6E6, 0.4% of NG, 0.116% ONG or 0.025% of DDM. Peak fractions containing the detergent-exchanged membrane protein complex were concentrated. A 50 kDa MWCO concentrator was used to concentrate samples in the detergent C6E6 and 100 kDa MWCO concentrator used on membrane protein complexes in all other detergents. Concentrated proteins were either used directly or flash-frozen in liquid nitrogen and stored at −80°C. Notably, we found no observable difference in mass spectra quality after a single freeze-thaw of membrane protein samples. Purified membrane proteins were initially washed in MS Buffer (two times the CMC of detergent of interest and 200 mM ammonium acetate, pH 7.2–8.0 with ammonium hydroxide) using a centrifugal buffer exchange device (Micro Bio-Spin 6, Bio-Rad) as previously described44. Membrane protein detergent exchanged into MS Buffer supplemented with C6E6 or ONG could be flash frozen without compromising mass spectra quality unlike MS Buffer supplemented with NG or DDM.

Preparation and titration of phospholipids. Phospholipids were purchased from Avanti Polar Lipids Inc., Alabana, USA) and prepared at stock concentrations of 10 mg ml−1 in 200 mM ammonium acetate pH 8.0 as previously described44. The titration of each individual membrane protein to phospholipid was optimized for each membrane protein complex to achieve consistency of nanoelec-trospray and resolved mass spectral peaks of bound phospholipid throughout the gas-phase unfolding series (Extended Data Fig 2b). Phospholipid prepara-tions were added to buffer exchanged membrane protein complexes followed by
equilibration at room temperature for 10 to 30 min before mass spectrometry analysis. Notably, by mass spectrometry we did not observe a significant increase in bound phospholipid with longer incubation times.

**Mass spectrometry.** Mass spectrometer settings were initially set to values for membrane proteins as previously described for a modified Q-TOF 2 mass spectrometer (Micromass, Manchester, UK) with a Z-spray source. Typical instrument values were 5–7 nC for source pressure, 1.5–1.8 kV for capillary voltage, 150–190 V for cone voltage, 1–10 V for extraction voltage, 180–200 V for collision voltage, argon for collisional gas and 0.2–0.3 MPa argon gas pressure.

**Ion mobility mass spectrometry.** Ion mobility measurements were performed on a Synapt G2 (where noted) or on a modified Synapt G1 HDMS instrument with the travelling-wave ion mobility cell replaced by an 18-cm drift cell with radial RF confinement and a linear voltage gradient to direct ions along the axis of transmission to the time-of-flight (TOF) mass analyser. The ion mobility mass spectrometer was typically set to a source pressure of 5–7 mBar, capillary voltage of 1.4–1.7 kV, capillary nanoflow of 0.03–0.2 mBar and argon as collision gas with flow rate set to 5–8 ml min$^{-1}$ (~6.6–6.7 e$^{-1}$ mBar). Collision voltage ranged from 50 to 240 V with measurements taken at 5 V steps for monitoring gas-phase unfolding of membrane protein complexes. The sample and extraction cone, and trap bias voltages, quadrupole profile, and collision (trap) gas pressure were optimized for maximal ion intensity of the target membrane protein complex. Helium was the drift cell gas and was set to a pressure of ~1.6–1.8 Torr (40 ml min$^{-1}$ flow rate). Pusher time, pulse width and TDC inhibit were set to 180, 70 and 7 μs, respectively. The ion guide/source (300 μs$^{-1}$, 10 V), trap (300 μs$^{-1}$, 0.2 V) and transfer (100 μs$^{-1}$, 10 V) travelling wave velocities and wave heights were kept constant. The mobility release time was set at 200 μs with a trap voltage of 30 V and extract voltage of 10 V. Pressure and temperature of the ion mobility cell was measured directly.

Ion drift time (t$_d$) was determined as previously described. Briefly, t$_d$ was determined by subtracting the non-drift time component (t$_n$) from the ion arrival time (t$_a$)

\[ t_d = t_a - t_n \]

where z is the charge on the ion, N is the reduced mass of the ion-neutral in the drift chamber, N is the drift gas number density, k$_B$ is the Boltzmann constant, ε is the charge on the electron, V is the potential voltage across the drift cell, p is the pressure of the drift cell, T is the thermodinamic temperature of the drift cell, L is the length of the drift tube, T$_R$ is standard temperature and p$_R$ is standard pressure.

Travelling wave ion mobility measurements (TW-IMS) were performed on a commercially available Synapt G2 (Waters). Source pressure set to 5–7 mBar, capillary voltage of 1.4–1.7 kV, capillary nanoflow of 0.03–0.2 mBar and argon as collision (trap) gas at flow rate of 8.0 ml min$^{-1}$. Collision voltage ranged from 50 to 200 V with measurements taken at 5 V steps intervals for monitoring gas-phase unfolding of membrane complexes. The sample and extraction cone and trap bias voltages, quadrupole profile, and collision gas pressure were optimized for maximal ion intensity of the target membrane protein complex. Helium was the IMS entrance cell gas set to a flow rate of 180 ml min$^{-1}$ and the drift cell gas was nitrogen set to a flow rate of 90 ml min$^{-1}$. The EDC-delay coefficient was 1.57 ms. The trap wave velocity and wave height were 300 ms$^{-1}$ and 8.0 V, respectively. The IMS wave velocity and wave height were 300 ms$^{-1}$ and 30.0 V, respectively. The transfer wave velocity and wave height were 100 ms$^{-1}$ and 2.0 V, respectively. The mobility release time was set at 200 μs with a trap voltage of 30 V and extract voltage of 10 V. Gas-phase unfolding profiles are largely similar to those collected on a drift cell instrument (Extended Data Fig. 6d).

**Collision cross-section calculations.** The projection approximation (PA) calculation within a modified version of MOBCAL$^{14,44,45}$ used to calculate the ΩPA from respective crystal structure coordinates. CCS values were corrected Ωcorr, by the correction factor of 1.14 and missing residues between the protein construct (M̃_{eXp}) and crystal structure (M̃_{crystal})

\[ Ω_{corr} = 1.14Ω_{PA}(\frac{M_{exp}}{M_{crystal}})^{\frac{1}{2}} \]

This correction has been shown to give values that correlate with known values to < 3%$^{41,47}$.

**Gas-phase unfolding data analysis and modelling.** Mass spectra and ion mobility data were analysed using in-house software written in the Python programming language with a graphical user interface constructed using wxPython (manuscript in preparation). Briefly, MS-only data collected on a Q-TOF 2 instrument was smoothed with MassLynx software before being imported and linearized at a step size of one in m/Δz. Theoretical mass spectra were predicted following equations described in Stengel et al.$^{48}$, Mass, average charge state, width of charge state distribution, resolution, and concentration were fitted to experimental data with least squares regression of the pseudo-$x^2$ function.$^{46}$

Ion mobility data were imported using a custom script provided by Waters at an m/Δz resolution of four in drift data and one in mass spectra for fitting. Mass spectra were smoothed and theoretical mass spectra were fitted as described above. A standardized integration window was employed for extracting arrival time distributions (ATD) using a resolution (R = m/Δm + Δm/z) of 1,000 and a height cut-off of 25% of the modelled Gaussian to the respective m/Δz peak. To generate gas-phase unfolding plots, ATDs were extracted for the given m/Δz peak using the standardized integration window followed by normalization. To convert to a CCS axis, arrival time was directly converted to CCS using equations above. This procedure generates gas-phase unfolding data and are plotted using Matlablib$^{49}$. All ion mobility mass spectra are shown in linear scale.

To a first order approximation, gas-phase unfolding of protein complexes, like unfolding in solution, will be linearly dependent (m) on collision voltage (CV)

\[ ΔG = ΔG^0 - m(CV) \]

where AG is free energy of unfolding and ΔG$^{0}$ is the free energy of the system at zero collision voltage in the gas-phase. Using this assumption, two, three, four and five state equilibrium unfolding models were used to model the abundance of native and non-native species. Notably, two new free parameters are introduced for each new species. Two, three, four and five state unfolding models were fitted to the data, and an F-test was used to decide if the fitting quality is sufficiently improved as to justify the inclusion of additional fitting parameters. The most abundant charge state of the three membrane protein complexes agreed with a four state unfolding model. For example, in the case of three unfolding species we used the following thermodynamic denaturing unfolding model$^{20}$

\[ N = K_1 + K_2 + K_3 \]

Expansion of AG in terms gives the following,

\[ ΔG_{N≡1} = ΔG_{N≡2} - m_{N≡1}(CV) \]

where F is a molar fraction. The sum of mole fractions of native (F$_N$) and non-native species (F$_1$ and F$_2$) is equal to 1

\[ F_N + F_1 + F_2 = 1 \]

We can generalize this taking into account the entire relationships one can obtain the fraction of native and non-native given K$_1$ and K$_2$

\[ F_N = \frac{1}{1 + K_1 + (K_1K_2)} \]

\[ F_1 = \frac{K_1}{1 + K_1 + (K_1K_2)} \]

\[ F_2 = \frac{(K_1K_2)}{1 + K_1 + (K_1K_2)} \]

The respective ΔG$^{0}$ and m parameters were solved numerically. In the two species case one can obtain a closed form solution, however, it is both efficient and general to solve them numerically.

An irreversible kinetic unfolding model as an alternative for modelling gas-phase unfolding data was also explored. The irreversible unfolding model defines the degree of unfolding as being related to time. By changing the potential across the drift cell, the time ions spend in the ion mobility drift cell post collision-induced unfolding can be altered over several orders of magnitude. The relative abundances of unfolded intermediates species as a function of drift time was monitored for a membrane protein complex. There was no considerable change in relative abundances over a fourfold increase in time (Extended Data Fig. 6e, f) demonstrating that the unfolding mechanism does not agree with an irreversible kinetic model for the unfolding model and no further unfolding occurs post collision cell. Data are in excellent agreement however with a reversible model as described above.
Ion mobility intensities were extracted for native and unfolded CCS species at various collision voltages. With the assumption that signal intensity is approximately proportional to species concentration, the data can be converted to a molar fraction to generate two dimensional (2D) data, which can be fitted to an equilibrium unfolding model. Minimized 2D parameters were used as a seed for three dimensional (3D) fitting of gas-phase unfolding data.

From the unfolding curve, and parameters $m$, $\Delta G^\circ$, CCS$_{m}$ (mean of CCS species), CCS$_{50}$ (standard deviation of CCS species), the 3D data can be fit. First, a CCS range was selected for each native and unfolded species in the unfolding plot. These CCS ranges provided the initial parameters used to model various CCS species identified by inspection. Next, to make the calculations more efficient, the gas-phase unfolding data were trimmed to a minimum and maximum value derived by either subtraction or addition of 10% times the native and last unfolding CCS species, respectively. Trimming resulted in gas-phase unfolding containing on average around 900 data points. The equilibrium unfolding model described above was seeded with the parameters from the minimized 2D data and the molar fraction of species was normalized to match the normalization of experimental data. The fitting surface was rugged and required a sophisticated minimization procedure to maximize the chance of finding global minimum. First the 3D model was minimized using quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) available in Scipy. This provided the boundaries for bio-inspired algorithms; all parameters were set to boundaries of ± 25% with the exception of the CCS species peak centres which were set to ± 5%. The 3D model was minimized using the modified differential algorithm (de_1220) available in PyGMO, using the following parameters: population of 20; 12 evolutions; 250 generations; and 8 islands. Other algorithms, such as particle swarm and bee colony were screened however the modified differential algorithm produced similar or higher $R^2$ values in considerably less computation time.

The resulting minimized 3D model enabled quantification of unfolding transitions and stabilization. An unfolding transition occurs when a species with a determined CCS transitions to another CCS species, for example see Extended Data Fig. 6c. The transition midpoint, CV$_{50}$ is determined by:

$$CV_{50} = \frac{\Delta G^\circ}{m}$$

where CV$_{50}$ is the value at which 50% of a specific transition state $i$ is depleted. This metric was found to be particularly useful for this application as it removes correlations between $\Delta G^\circ$ and $m$ inherent in the fitting procedure. The stabilization of a protein due to the presence of additional phospholipids was quantified according to:

$$\Delta CV_{50} = CV_{50}^{\text{PDB}} - CV_{50}^{\text{apo}}$$

This comparative approach takes into account to first order any variability in transition collision voltages between repeated measurements for a given protein.

To compare between various proteins and different unfolding states, as well as to average out any systematic and/or inherent variability in droplet formation, the average stabilization was calculated for the bound ligand by:

$$\text{Stabilization} = \frac{\sum_{i} \Delta CV_{50}}{\xi}$$

where $\xi$ is the number of transitions, $i$ is a specific transition, and $z$ is the charge state. Averages and standard error of the mean was calculated for three repeated measurements. Statistical significance between means was determined by a one-tailed (3D) fitting of gas-phase unfolding data.

**Identifying lipid binding sites from MD simulations and CCS measurements.** The object is to determine which MD frames are in best agreement with the experimental CCS data. For every simulated system, one frame was extracted every 0.2 ns generating 1,000 PDB files for a 500 ns simulation time. For each extracted frame, all lipids within 6 Å of the protein were identified. All possible combinations of these two lipids and protein were selected. For example for AqpZ and one lipid typically yielded 80 extracted combinations per frame (Extended Data Fig. 4c).

The CCS value of each combination per frame ($f$) was calculated using a modified version of MOBCAL. A ratio ($R_{calc}$) was determined using the calculated CCS values for each of the +1 lipid and the +2 lipid-bound states

$$R_{calc} = \frac{\text{CCS}_{\text{protein} + \text{l lipid}}}{{\text{CCS}_{\text{protein}}}}$$

where CCS$_{\text{protein} + \text{l lipid}}$ is CCS of protein plus number of lipid(s) (l) and CCS$_{\text{protein}}$ is CCS of protein. Using a ratio avoided any potential issues associated with taking the difference of uncorrected CCS values determined experimentally. These measures enable the calculation of three chi-squared ($\chi^2$) values for a given combination of two lipids and protein

$$\chi_i^2 = \frac{\left( R_{calc} - R_{exp} \right)^2}{\sigma_i^2}$$

where $R_{calc}$ and $\sigma$ are the average and standard deviation of the experimental CCS ratio for data with one or two lipid molecules bound from three repeated measurements, and $R_{exp}$ is the CCS ratio derived from theoretical calculations. The overall chi-squared values were calculated from:

$$\chi^2 = \chi_1^2 + \chi_2^2 + \chi_3^2$$

where $\chi_1^2$, $\chi_2^2$, and $\chi_3^2$ are the chi-squared values for each individual lipid molecule with the protein. In the case where only the +1 lipid CCS data are analysed, neglecting the +2 lipid data, only one $\chi^2$ value contributed towards the sum. The probability ($P$) that a given protein phospholipid pair with protein matches the data are given by the following Bayesian probability:

$$P = \frac{1}{\sqrt{2\pi \sigma^2}} \exp \left(-\frac{\left( \chi^2 - \chi^2_{\text{min}} \right)^2}{2\sigma^2} \right)$$

where $\chi^2_{\text{min}}$ is the minimum chi-squared value obtained across the entire simulation, whose corresponding $P$ is equal to 1, representing the most probable structure. It follows that all candidate lipid pair and protein combinations with a $P$ value greater than 0.785 will be one half standard deviation from the most probable structure. The ensemble of structures identified with $P > 0.785$ can be considered the most probable arrangement of lipids and protein found to be consistent with the experimental data. This ensemble of structures was projected onto the surface of the protein to identify the most probable locations of contact between the protein and the lipid.

**Total polar lipid extracts from cardiolipin-deficient E. coli strain, BKT22.** The cardiolipin-deficient E. coli strain (BKT22 (AclsA, AclsB, AclsC, AymdB-BanB)) and calcium chloride competent BKT22 cells were transformed with pBAD-YC (plasmid containing clsC and ymdB genes), referred herein as BKT22-YC, and grown overnight in LB media. Overnight cultures were diluted to an OD$_{600}$ of 0.03 in LB media with the exception that arabinose was added to a final concentration of 0.2% for BKT22-YC, and grown overnight in LB media. Overnight cultures were diluted to an OD$_{600}$ of 0.03 in LB media with the exception that arabinose was added to a final concentration of 0.2% for BKT22-YC to induce expression of cardiolipin genes, clsC and ymdB. The strains were grown for six hours to stationary phase (OD$_{600}$ ~2.0). Cells were collected by centrifugation at 3,500×g for 10 min. Cell pellets were resuspended in phosphate-buffered saline followed by centrifugation.

Lipids were extracted using the Folch method. Briefly, cells were resuspended in 20 ml of 2:1 (v/v) chloroform:methanol per gram of cells and agitated on a stir plate for 20 min. The mixture was then clarified using filter paper and washed with 0.2 volume of water. The extract was briefly vortexed before centrifugation at 800g.
to separate phases. The upper aqueous phase was removed and the interface was washed with 1:1 (v/v) methanol:water without disturbing the lower organic phase. The lower organic phase was recovered and evaporated under a stream of nitrogen. The dried total lipid extract was then resuspended in diethyl ether to extract polar lipids and evaporated under nitrogen then under vacuum overnight.

**Thin-layer chromatography.** HPTLC Silica gel 60 aluminium plates (Merck Millipore) were pre-developed with 1:1 (v/v) chloroform:methanol and left to air-dry overnight. The plates were then impregnated with an adsorbent modifier (1:1 (v/v) 1.2% boric acid in absolute ethanol-water) followed by air-drying at 100 °C for one hour. Samples were spotted and plates developed in 65:25:5 (v/v) chloroform:methanol:acetic acid. After air-drying for 30 min, plates were sprayed with 10% sulphuric acid in ethanol and charred at 200 °C (ref. 63). Molar per cent (mol %) of lipid species was determined from the intensity of lipid species using ImageJ 1.47v software.

**Proteoliposome/liposome preparation.** E. coli total polar lipid extract (aceton/ether preparation; Avanti Polar Lipids) was dissolved at 20 mg mL⁻¹ in 2:1 (v/v) chloroform:methanol to obtain a clear solution. Solvent was evaporated under a stream of nitrogen gas then under vacuum overnight. E. coli total polar lipid (EPL), BKT22 and BKT22-VC extracts were hydrated in 2 mM BME to a final concentration of 50 mg mL⁻¹ and left to incubate for 1 h at room temperature with constant agitation. These were then split into aliquots and frozen. Lipids were always handled under a nitrogen atmosphere.

Liposomes and proteoliposomes were prepared following previously established protocols. Briefly, lipid stocks were diluted into borosilicate tubes to a final concentration of 45 mg mL⁻¹ in 100 mM MOPS–NaH₂PO₄ pH 7.5 and pulsed in a bath sonicator until a clear suspension was obtained. A reconstitution mixture (400 μL) was prepared in a borosilicate tube at room temperature by sequentially adding 100 mM MOPS–NaH₂PO₄ (pH 7.5), 1.25% (wt/vol) OG, purified Ni₃His-AggZ for proteoliposomes (final concentration 100 μg mL⁻¹), and 10 mg mL⁻¹ sonicated lipids. After incubation for 1 h at room temperature, the mixture was centrifuged 25-fold into 20 mM HEPES pH 7.5. The liposomes were collected by centrifugation 140,000 × g for one hour and resuspended to a final volume of 1.5 mL in 20 mM HEPES pH 7.5. Liposomes and proteoliposomes were stored on ice before measurement. EPL, BKT22 and BKT22-VC liposomes/proteoliposomes had an average diameter of 140, 149 and 160 nm determined from dynamic light scattering (ViscoTek 802).

**AggZ water permeability assay.** Liposome shrinkage was induced by rapidly mixing 1:1 (v/v) proteoliposomes/liposomes with osmolyte buffer (20 mM HEPES pH 7.5, 570 mM sucrose; a 285 milliosmolarity gradient) using a stopped-flow apparatus (Applied Photophysics SX20). Water transport/permeability was monitored by measuring the light scattering (λ = 600 nm) of the preparation upon mixing at 8 °C. An increase in signal reflects liposome shrinkage. Light scattering (L) data from repeated measurements (n = 5) were fit to a single exponential rise equation to obtain the rate of water transport (kₛₑ) using OriginPro 8.5.1.

$$\frac{L(t)}{L_0} = A e^{-k_s t} + L_n$$

where t is time in msec, A is a constant, kₛₑ is rate of water transport, and Lₙ is the maximum light scattering (typically one for normalized data).

AmitB crystallization and structure determination. AmitB was purified as described for native mass spectrometry experiments with one minor modification. AmitB was detergent exchanged into buffer H containing 0.5% C₈E₄ using a Superdex 200 10/300 gel filtration column (GE Healthcare). Peak fractions of AmitB from gel filtration were pooled and concentrated to ~15 mg mL⁻¹ using a 50 kDa MWCO concentrator (Millipore). A tenfold molar excess of phosphotidylglycerol was added to AmitB before crystallization. Crystals of AmitB were grown at 20 °C in hanging drop plates with crystallization solution 15% PEG 4000, 0.8 M potassium formate, and 0.1 M sodium acetate pH 4.6. Hexagonal plate crystals appeared after one month.

Single crystals were mounted with CrystalCap HT Cryoloops (Hampton Research, Aliso Viejo, CA) before flash frozen. Data were collected at Diamond Light Source and 0.1 M sodium acetate pH 4.6. Hexagonal plate crystals appeared after one month.

**Comparison of simple potential functions for simulating liquid water.** J. Chem. Theor. Comput. 5, 926–935 (1983). Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the **Comparative** potential functions for simulating liquid water. J. Chem. Theor. Comput. 7, 50. Jones, E. *et al.* Scipy: open source scientific tools for Python http://www.scipy.org/ (2001).

1. Izzo, D., Rucinski, M. & Biscani, F. The generalized island model. Studies in Computational Intelligence 151, 151–169 (2012).
2. Izzo, D. PyGMO and PyKEP: open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization). Proceed. Fifth International Conf. Astrodyn. Tools and Techniques, ICATT (2012).
3. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation using a mass spectrum deconvolution approach. Chem. Biol. 19, 599–607 (2012).
4. Hunter, J. D. Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95 (2007).
5. Jones, E. *et al.* Scipy: open source scientific tools for Python http://www.scipy.org/ (2001).

1. Izzo, D., Rucinski, M. & Biscani, F. The generalized island model. Studies in Computational Intelligence 151, 151–169 (2012).
2. Izzo, D. PyGMO and PyKEP: open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization). Proceed. Fifth International Conf. Astrodyn. Tools and Techniques, ICATT (2012).
3. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation using a mass spectrum deconvolution approach. Chem. Biol. 19, 599–607 (2012).
4. Hunter, J. D. Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95 (2007).
5. Jones, E. *et al.* Scipy: open source scientific tools for Python http://www.scipy.org/ (2001).

1. Izzo, D., Rucinski, M. & Biscani, F. The generalized island model. Studies in Computational Intelligence 151, 151–169 (2012).
2. Izzo, D. PyGMO and PyKEP: open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization). Proceed. Fifth International Conf. Astrodyn. Tools and Techniques, ICATT (2012).
3. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation using a mass spectrum deconvolution approach. Chem. Biol. 19, 599–607 (2012).
4. Hunter, J. D. Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95 (2007).
5. Jones, E. *et al.* Scipy: open source scientific tools for Python http://www.scipy.org/ (2001).

1. Izzo, D., Rucinski, M. & Biscani, F. The generalized island model. Studies in Computational Intelligence 151, 151–169 (2012).
2. Izzo, D. PyGMO and PyKEP: open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization). Proceed. Fifth International Conf. Astrodyn. Tools and Techniques, ICATT (2012).
3. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation using a mass spectrum deconvolution approach. Chem. Biol. 19, 599–607 (2012).
4. Hunter, J. D. Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95 (2007).
5. Jones, E. *et al.* Scipy: open source scientific tools for Python http://www.scipy.org/ (2001).

1. Izzo, D., Rucinski, M. & Biscani, F. The generalized island model. Studies in Computational Intelligence 151, 151–169 (2012).
2. Izzo, D. PyGMO and PyKEP: open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization). Proceed. Fifth International Conf. Astrodyn. Tools and Techniques, ICATT (2012).
3. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation using a mass spectrum deconvolution approach. Chem. Biol. 19, 599–607 (2012).
4. Hunter, J. D. Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95 (2007).
5. Jones, E. *et al.* Scipy: open source scientific tools for Python http://www.scipy.org/ (2001).

1. Izzo, D., Rucinski, M. & Biscani, F. The generalized island model. Studies in Computational Intelligence 151, 151–169 (2012).
2. Izzo, D. PyGMO and PyKEP: open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization). Proceed. Fifth International Conf. Astrodyn. Tools and Techniques, ICATT (2012).
3. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation using a mass spectrum deconvolution approach. Chem. Biol. 19, 599–607 (2012).
4. Hunter, J. D. Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95 (2007).
5. Jones, E. *et al.* Scipy: open source scientific tools for Python http://www.scipy.org/ (2001).
62. Fine, J. B. & Sprecher, H. Unidimensional thin-layer chromatography of phospholipids on boric acid-impregnated plates. J. Lipid Res. 23, 660–663 (1982).
63. Tan, B. K. et al. Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. Proc. Natl Acad. Sci. USA 109, 16504–16509 (2012).
64. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9, 671–675 (2012).
65. Kabach, W. XDS. Acta Crystallogr. D 66, 125–132 (2010).
66. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
67. Padilla, J. E. & Yeates, T. O. A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. Acta Crystallogr. D 59, 1124–1130 (2003).
68. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D 58, 1948–1954 (2002).
69. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D 60, 2126–2132 (2004).
70. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D 53, 240–255 (1997).
71. Schrödinger, LLC The PyMOL Molecular Graphics System, Version 1.3r1 (2010).
Extended Data Figure 1 | Maintaining intact native membrane protein complexes in the mass spectrometer. a, b, Gas-phase unfolding plots (left, 5 V steps) and mass spectra (right) of detergent stripped AqpZ and AmtB ions (charge inset) from different non-ionic detergent solutions; C8E4 (pink), NG (green), OGNG (purple) and DDM (orange). Membrane protein complexes from NG, OGNG and DDM possessed CCS values substantially larger than those calculated from crystal structures. c, Complete removal of C8E4 at low-collision voltages reveals CCS values consistent with those calculated from their respective crystal structures. d, Reported are measured masses, standard deviations, and empirical t0 values used for direct CCS calculation of membrane proteins studied.
Extended Data Figure 2 | Phospholipid abbreviations and optimization of phospholipid binding experiments. **a**, Phospholipid abbreviations and headgroup structures. **b**, Protein to phospholipid to detergent ratios (P:L:D) for each membrane protein, number of resolvable phospholipids bound with their masses. Conditions were optimized empirically to maintain nanospray, sufficient mass spectral quality and phospholipid binding. Masses for one lipid bound to the protein complex were measured using MassLynx V4.1 software (Waters).
Extended Data Figure 3 | Phospholipid binding to MscL reveals an insignificant impact on protein gas-phase stabilization independent of lipid alkyl chain length. a, Representative mass and ion mobility spectra of MscL bound to various phospholipids species (inset top right). b, Representative mass and ion mobility spectra of MscL bound to phospholipids of varying alkyl chain length. c, Cumulative stabilization of MscL is seen for one (blue) to two (red) bound lipid molecules. Reported are the means ± s.e.m. from repeated measurements (n = 3).
Extended Data Figure 4 | Summary of statistics for molecular dynamics (MD) simulations followed by CCS filtering. a, Representative filtering procedure of AqpZ in a PC bilayer from the MD (top). For each frame, lipids are extracted within 6 Å of the protein (bottom left) then filtered by the experimental CCS of AqpZ bound to a single PC molecule (bottom right); see Methods. b, Summary of statistics. c, Ratios of phospholipid bound to apo CCS values for membrane protein complexes. Reported are the mean and standard deviation ($n = 3$). d, MD filtered by CCS resulted in similar patches of PC molecules across different time points within the simulation indicating the systems were equilibrated.
Extended Data Figure 5 | MD simulations filtered by CCS reveals probable one (1×) and two (2×) phospholipid binding sites. a, For each candidate structure, protein and 1× or 2× PC molecule(s), the ratio of their calculated CCS values was determined (CCS ratio). This procedure generated a large number of candidate structures (grey bars) that were filtered using our CCS measurements (cyan line). The structures in grey that intercept this curve are essentially the ones selected as our most probable ensemble. b, The intersection between the simulated lipid complexes and the experimental data are then projected onto the surface of the protein to identify the most probable binding sites. Probable lipid locations for MscL resembled an annular belt, with no specific patches of lipids probably stemming from the relatively cylindrical geometry of this complex. By contrast, for AqpZ and AmtB the most probable location of the lipid molecules were localized to the interfacial regions between protein subunits, as well as other probable locations on individual monomers. c, X-ray derived PG (blue spheres with white tails) located at the subunit interfaces agrees with the predicted PC binding site.
Extended Data Figure 6 | Modelling and quantification of gas-phase unfolding pathways. a, Representative ion mobility mass spectra are collected over a range of collision voltages. b, Ion arrival time is converted to CCS before generating unfolding plots. c, Model fitting process from 2D to 3D data (see Methods). d, Both a Synapt2 modified with a linear drift cell (DT-IMS) and the commercially available travelling wave SynaptG2 (TW-IMS) produce qualitatively similar unfolding data. e, A contour plot representing the variance of CCS of two gas-phase unfolding species as a function of ion mobility drift cell potential. f, Stacked plots of arrival time distributions for two gas-phase unfolding species as a function of drift cell potential. The lifetime of unfolding protein complexes in the drift tube ranges from 4 to 15 ms depending on the drift cell potential. No additional unfolding post activation occurs implying that the unfolding mechanism is not consistent with an irreversible unfolding model. Such a mechanism would predict time dependence on the population of unfolded species. By contrast the unfolding mechanism is well described by the reversible unfolding mechanism (see Methods).
Extended Data Figure 7 | AqpZ, AmtB and AmtB<sup>N72A/N79A</sup> bound to various phospholipid species. a, Representative mass and ion mobility spectra of AqpZ bound to phospholipids. b, Representative mass and ion mobility spectra of AmtB bound to phospholipids and AmtB<sup>N72A/N79A</sup> bound to PG.
Extended Data Figure 8 | Summary of water permeability assays and analysis of lipid extracts. 

**a**, HPTLC analysis of total polar lipid extract from wild-type *E. coli* (EPL), cardiolipin-deficient strain (BKT22), or BKT22 cells expressing ClsC and YmdB (BKT22-YC) to restore cardiolipin. Lipids were quantified by densiometry. 

**b**, Reported are the rate constants ($k_{\text{wat}}$) and standard error of replicates ($n = 5$) for empty liposomes (−) and AqpZ proteoliposomes (+) reconstituted in differing *E. coli* lipid compositions.
Extended Data Figure 9 | Structural analysis of AmtB bound to PG.

a, Crystal packing with six AmtB (multicoloured) and eight PG (orange) molecules located in the asymmetric unit cell and symmetry related molecules shown in grey and light orange, respectively. 
b, $F_o-F_c$ and $2F_o-F_c$ electron density maps after refinement without lipid and near-lipid water molecule (if present) contoured at 2.0 and 1.0 sigma, respectively. 
c, Comparison of AmtB bound to PG (green chain, this work) aligned with the AmtB structure (maroon chain, PDB: 1U7G). 
d, Structure overlay of AmtB–GlnK complex bound to octylglucoside (purple chain, PDB: 2NS1) aligned with AmtB bound to PG reveals a distinct conformational change. The lipid–water interface (grey plane) was determined from coordinates of phosphate atoms from bound PG.
## Extended Data Table 1 | Summary of X-ray data collection and refinement statistics

| AmiB bound to PG |
|------------------|
| **Data collection** |
| Space group | C222₁ |
| Cell dimensions |
| a, b, c (Å) | 116.2, 201.2, 232.5 |
| α, β, γ (°) | 90, 90, 90 |
| Resolution (Å) | 40-2.3 (2.4-2.3) |
| Rmerge (%) | 12.4 (120.4) |
| I/σ(I) | 13.7 (2.5) |
| Completeness (%) | 99.9 (99.5) |
| Redundancy | 11.0 (6.9) |
| **Refinement** |
| Resolution (Å) | 38.7-2.3 (2.36-2.3) |
| No. reflections | 119,766 |
| Rwork / Rfree | 20.2 / 23.4⁹ |
| No. atoms |
| Protein | 15,856 |
| Ligand/ion | 293 |
| Water | 111 |
| B-factors |
| Protein | 45.3 |
| Ligand/ion | 64.7 |
| Water | 29.5 |
| R.m.s. deviations |
| Bond lengths (Å) | 0.016 |
| Bond angles (°) | 1.71 |

---

⁹ Values in parentheses are for highest-resolution shell.

⁹ Rfree calculated using 5% of the data.