**The role of interfacial lipids in stabilising membrane protein oligomers**

**Kallol Gupta*[1], Joseph A.C. Donlan*[1], Jonathan T.S. Hopper*[1], Povilas Uzdavinys[2], Michael Landreh[1], Weston B. Struwe[1], David Drew[2], Andrew J. Baldwin[1], Phillip J. Stansfeld[3] and Carol V. Robinson[1]**

1. Department of Chemistry, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA.
2. Centre for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, Stockholm SE-106 91, Sweden
3. Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU
4. * These authors contributed equally to this manuscript

**Abstract**

**Oligomerisation of membrane proteins in response to lipid binding plays a critical role in many cell-signaling pathways [1] but is often difficult to define [2] or predict [3]. Here we develop a mass spectrometry platform to determine simultaneously the presence of interfacial lipids and oligomeric stability and discover how lipids act as key regulators of membrane protein association. We reasoned that stability would relate to free energy of oligomeric dissociation ($\Delta G°_{dissoc(m)}$), buried surface area and number of salt bridge interactions. Calculating these values for a dataset of 125 α-helical oligomeric membrane proteins revealed an absence of interfacial lipids in the mass spectra of 13 membrane proteins with high oligomeric stability. For the protein with the lowest oligomeric stability, the bacterial homologue of the eukaryotic bioamine transporters (LeuT) [4, 5], we found a precise cohort of lipids within the dimer interface. Delipidation, mutation of lipid binding sites or expression in cardiolipin (CDL) deficient *Escherichia coli*, abrogated dimer formation. Molecular dynamics simulation revealed that CDL acts as a bidentate ligand bridging across subunits. Subsequently, we show that for the sugar transporter SemiSWEET from *Vibrio splendidus* [6], another protein with low oligomeric stability, cardiolipin shifts the equilibrium from monomer to functional dimer. We hypothesised that lipids would be essential for dimerisation of the Na+/H+ antiporter NhaA from *E. coli,* but not for homologous NapA from *T. thermophilus,* which has a $\Delta G°_{dissoc(m)}$ considerably higher than that of NhaA. We found that lipid binding is obligatory for dimerisation of NhaA, whereas NapA has adapted to form an interface that is stable without lipids. Overall, by correlating free energy of dissociation with presence of interfacial lipids we provide a rationale for understanding the role of lipids in both transient and stable interactions within a range of α-helical membrane proteins, including GPCRs.**

The recent surge in structure determination of membrane proteins is providing details of protein-lipid binding [7] and yielding insight into the regulatory roles of lipids [8, 9]. The advent of mass spectrometry (MS) methods for characterising membrane proteins, individually [10], within interactomes [11], and in intact assemblies [12], is adding new information to potential roles of lipids inducing conformational changes [13], contributing to activity and modulating drug efflux, (reviewed in [14]). The role of lipids towards maintaining the oligomeric state of membrane proteins has however remained widely debated. To understand this phenomenon we performed a

bioinformatics analysis of all the α-helical oligomeric transmembrane proteins with known structures to rank them as per their oligomeric stability. We calculated the free energy of oligomeric dissociation ($\Delta G^\circ_{dissoc(m)}$) and separately validated the approach through calculation of the interface strength from buried surface area and number of salt bridge interactions across the interface (Fig. 1, Extended Data Fig. 1 and Supplementary Table S1). For 13 membrane proteins, with predicted high oligomeric stability, mass spectra revealed the measured masses are in agreement with that of their oligomeric masses, devoid of additional lipids (Fig. 1 and Extended Data Table 1). While membrane proteins with high $\Delta G^\circ_{dissoc(m)}$ values can still have substantial affinity towards lipids, lipid binding is not towards survival of the oligomeric entity [15].

By contrast, the weakest oligomeric interface was observed for the bacterial leucine transporter from *Aquifex aeolicus* (LeuT). LeuT is a sodium symporter that transports small aliphatic amino acids across the bacterial inner membrane and is the homolog of eukaryotic bioamine transporters [4]. A mass spectrum recorded following liberation of LeuT from octylglucoside (OG) micelles reveals its dimeric state (Fig. 2). The mass of the dimer was consistently greater than that of twice the monomer (126.0 kDa =2×59.3 + 7.4 kDa) over numerous preparations, with different constructs and collision energies, indicative of noncovalent binding of small molecules with high affinity and precise stoichiometry (Extended Data Fig. 2). Incubating LeuT in neopentyl glycol (NG), a detergent with high delipidation properties [16], and reconstituting into OG followed by MS, revealed exclusively monomeric LeuT (Extended Data Fig. 3a-c). Addition of *E. coli* polar lipids to delipidated monomeric LeuT, in OG, recovered a significant population of the dimer (Extended Data Fig. 3d). These observations imply that the additional mass associated with the dimer is comprised of lipid.

Conventional lipid identification experiments require extracting lipids from the proteo-micelle solution or cellular environment, followed by a chromatographic or MS step [17, 18]. These approaches report on the entire set of lipids present but fail to distinguish endogenous lipids binding to the membrane protein of interest from those in bulk solution. Identification of bound lipids simultaneously with oligomeric state requires a tandem MS (MS/MS) platform, akin to protocols developed to sequence peptides for top-down proteomics [19]. MS/MS in its current form cannot be applied to membrane proteins directly since the activation energy available in the collision cell is used to liberate membrane proteins from detergent micelles [20]. To overcome this problem, we developed an instrument platform (Fig. 2a) where high energy applied in the source region removes the detergent micelle, prior to entry into the collision cell, enabling isolation of discrete lipid-bound complexes in the quadrupole for subsequent MS/MS and lipid identification (Fig. 2a). Using this platform, we isolated the 23+ charge state of the dimer that incorporates the 7.4 kDa additional mass (Fig 2b). Activation of this species in the collision cell yields monomeric LeuT in apo and lipid-bound states, with one CDL and up to three phospholipids (PL). Increasing collision energy results in monomer retaining one CDL (Extended Data Fig. 4), enabling us to distinguish binding of one CDL from the possibility of two PLs. It is important to emphasize that all ions present in the MS/MS spectrum originate from dimeric LeuT parent ions with the additional 7.4 kDa, which can now be assigned to three PLs and one CDL per subunit (3.7 kDa or 7.4 kDa per dimer).

Expressing the wild type LeuT in a cardiolipin deficient *E. coli* strain [21] yielded exclusively delipidated monomeric LeuT (Extended Data Fig. 5), confirming that CDL is essential for dimerisation. Performing coarse-grained molecular dynamics (MD) simulation of the LeuT dimer in a bilayer revealed binding of CDLs and PLs at the dimer interface. Each CDL interacts with both

monomeric units, the bi-phosphate head group binding to basic residues on either side of the dimer interface (Fig 2c, Extended Fig 5-6, K376, H377 and R506). These results highlight the residues that form critical contacts with lipid head-groups to confer the specificity of these interactions. Mutating either the dimer interface or these basic residues that bind to CDL abolished dimer formation (Extended Data Fig. 5). We conclude that LeuT has a weak dimer interface, which is stabilised by CDL bridging the dimer interface and augmented by six phospholipids.

Bacterial sugar transporter SemiSWEET from *Vibrio splendidus*, a functional dimer [6], is another protein that has an exceptionally low oligomeric stability, comparable to that of LeuT. The mass spectrum of SemiSWEET shows the presence of both monomer and dimer (Extended Data Fig. 7a). To investigate if the presence of monomers is a consequence of the removal of lipids during protein purification, and whether the monomers are in equilibrium with the dimers, we prepared two mass distinct forms of SemiSWEET (+/- deca-His tag, Extended Fig 7). A time course MS experiment, subsequent to mixing of these two mass distinct forms in equal ratios, revealed rapid appearance of heterodimeric peaks, establishing a solution-phase monomer-dimer equilibrium (Fig. 3a). High energy MS/MS of SemiSWEET also identified endogenous bound lipids, a significant proportion of which is CDL (Extended Data Fig. 7c). Upon addition of CDL in increasing amounts to this equilibrium we observed preferential lipid binding to the dimer and a shift in the equilibrium towards the dimeric population (Fig 3b). In contrast addition of phosphatidylglycerol (PG) revealed no such preference towards any oligomeric forms (Extended Data Fig. 7d). We conclude that preferential lipid binding to the dimer drives the equilibrium towards the functionally active state of the protein.

Given our emerging hypothesis that lipids are critical for stabilising weak dimer interfaces, we searched for proteins with weak interfaces that might require lipids for dimerisation but have homologues with higher interface strengths. One such pair of proteins is the NhaA and NapA $Na^+/H^+$ antiporters from *E. coli* and *T. thermophilus* respectively (Fig. 1). While the oligomeric stability of dimeric NhaA is comparable to that of LeuT, dimeric NapA is substantially more stable. The mass spectrum of NhaA reveals an ensemble of lipid-bound dimeric species (Fig. 4a) and a complete absence of delipidated dimer. Performing MS/MS on the 15+ charge state of the lipid-bound NhaA dimer leads to stepwise losses of CDL and yields monomers with one CDL bound as well as an apo NhaA dimer that readily dissociates (Extended Data Fig. 8). The appearance of monomeric NhaA is coincident with the loss of the second CDL and consistent with CDL stabilizing the dimer structure. MD simulations support this observation by revealing that CDL can stabilize the dimeric state by binding at the interface (Extended Data Fig. 6). MS analysis of the homologous NapA reveals a striking contrast between lipid binding to NhaA and NapA (Fig. 4a). The NapA dimer is completely lipid-free, confirming its intrinsic stability in the absence of interfacial lipids. Proteins from thermophiles are known to be more stable than their non-thermophilic homologues. In NapA, an additional helix not present in NhaA strengthens the interface, essentially removing the requirement for lipids in dimer formation. These two proteins, with the same fold, physiological role and purified from identical membranes, demonstrate that membrane proteins can either (i) acquire structural elements to ensure greater contacts between subunits or (ii) recruit lipids to preserve their oligomeric state.

We anticipate that the ability to form a stable interface or recruit lipids to preserve oligomeric state might be a general phenomenon existing in other membrane protein systems, for example G-protein coupled receptors (GPCRs) [22]. Calculating the $\Delta G°_{dissoc(m)}$ of the two possible interfaces of μ-opioid

receptor we observed that TM5/TM6 dimer forms the stronger interface (16.5 kcal $mol^{-1}$) and TM1-TM2/H8 is considerably weaker (2.95 kcal $mol^{-1}$) (Fig. 4b) [23]. However, the tighter interface restricts the conformational flexibility required to attain the agonist bound state [24]. The weaker interface, , contains a cavity, akin to that found in NhaA, wherein a side chain of a fatty acid has been modeled in the crystal structure [24]. The weak dimeric interfaces involving TM1-TM2/H8 can also be constructed for many other GPCRs including the β1 adrenergic receptor and κ-opioid receptors (0.7 kcal $mol^{-1}$ and 6.1 kcal $mol^{-1}$ respectively) (Fig 4b). These very low oligomeric strengths are consistent with observations of transient oligomeric states [25] leading us to speculate that much of the controversy surrounding the oligomeric state of GPCRs stems from this ability to exist in multiple forms, with different interfaces modulated by interfacial lipids, analogous to the monomer - dimer equilibrium shown here for SemiSWEET.

While the intrinsic stability of the oligomers correlates with lipid binding to stabilise interfaces, a key question arises with respect to function. For both SemiSWEET and NhaA, existence of a stable dimeric state is critical for their mechanistic pathways [26]. Also for NhaA, it has been shown that under extreme stress conditions the dimer, observed predominantly with bound lipids, is functionally superior to the monomer [27]. By analogy with LeuT, dimerisation of the homologous eukaryotic dopamine (DAT) and serotonin transporters (SERT) might also be anticipated *in vivo* [28, 29]. Sequence alignment and superposition of the structures of LeuT and SERT reveals that the CDL binding residues identified here are conserved in all bioamine transporters (Extended Data Fig. 9). A key deviation between the X-ray crystallographic structures of LeuT and SERT arises in the C-terminal helix of SERT, which orients away from the subunit interface, preventing dimerization in the crystal form [30]. Nevertheless, a functional significance of dimers, or possibly higher oligomers, of SERT is well documented [31].

Overall our data shows how lipid binding at the interface stabilizes weak oligomers and provides direct and compelling evidence that altering the lipid composition in solution can propagate changes in oligomeric state. In the cellular environment such mechanisms are likely employed to regulate the abundance of functional forms of membrane proteins. As new structures of membrane proteins continue to emerge, the approach described here will help to resolve conflicts in oligomeric state and contribute to our understanding of their functional relevance, important considerations for the design of bio therapeutics and for drug targeting.

## Methods

Methods and associated references are available in the ONLINE METHODS SECTION.

## Acknowledgments

We thank Kevin Giles (Waters Corporation) and Justin Benesch towards the development of the high-energy source. Research in the Robinson group is funded by a Wellcome Trust Investigator Award (104633/Z/14/Z), an ERC Advanced Grant ENABLE (641317) and an MRC programme grant (MR/N020413/1). K.G. is a research fellow of the Royal Commission for the Exhibition of 1851 and Junior Research Fellow at St Catherine's College, Oxford. J.A.C.D. is supported by an EPSRC studentship, held at the Life Sciences Interface Doctoral Training Centre. M.L. is supported by an ERC Marie Curie Career Development Fellowship and is a Junior Research Fellow at St Cross College, Oxford. D.D. acknowledges support from EMBO through the EMBO Young Investigator Program and

the Knut and Alice Wallenberg foundation. The authors thank Eric Gouaux for the LeuT plasmid and Wolf B. Frommer and Liang Feng for the SemiSWEET plasmid. The authors thank Timothy Allison and Joseph Gault for helpful discussions and suggestions. A.J.B. acknowledges a BBSRC David Phillip's Fellowship, BB/J014346/1.

**Author contributions:** K.G., J.A.C.D. and C.V.R. designed the experiments. K.G. and J.A.C.D. performed the experiments. J.T.S.H. performed all the high energy experiments with K.G. K.G. and M.L. performed experiments on NapA and NhaA. P.U. expressed and purified NhaA and NapA under the guidance of D.D. J.A.C.D. purified SemiSWEET with the help of W.B.S. P.J.S. carried out molecular dynamics simulations. A.J.B. and K.G. did the theoretical calculation of determination of oligomeric strength. K. G. and C.V.R wrote the manuscript with contributions from all authors.

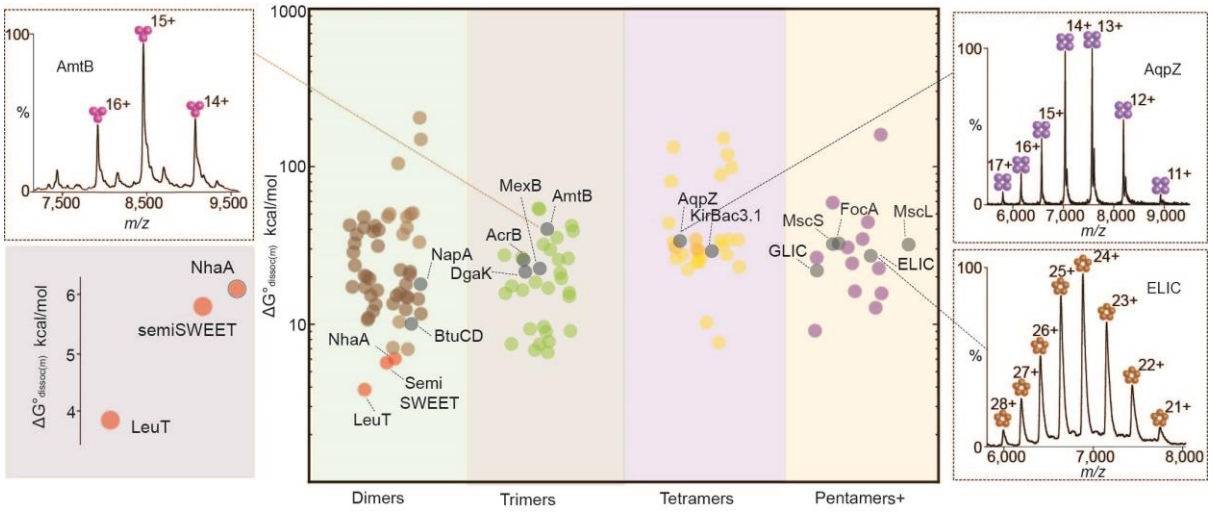



**Figure 1 | Plot of $\Delta G^{\circ}_{dissoc(m)}$ for oligomeric α-helical membrane proteins and representative mass spectra.** Plot of $\Delta G^{\circ}_{dissoc}$ with proteins, grouped by oligomeric state (pentamers+ oligomeric state ≥ 5). Membrane proteins for which native mass spectra have been acquired are labelled. LeuT, SemiSWEET and NhaA (bottom left expansion) have the lowest $\Delta G^{\circ}_{dissoc(m)}$. Two homologous proteins with different $\Delta G^{\circ}_{dissoc(m)}$, NapA and NhaA, are also labelled. Mass spectra of trimeric AmtB, tetrameric AqpZ and pentameric ELIC (insets clockwise from top left) are shown. A random horizontal jitter has been applied to all points to aid visualisation.

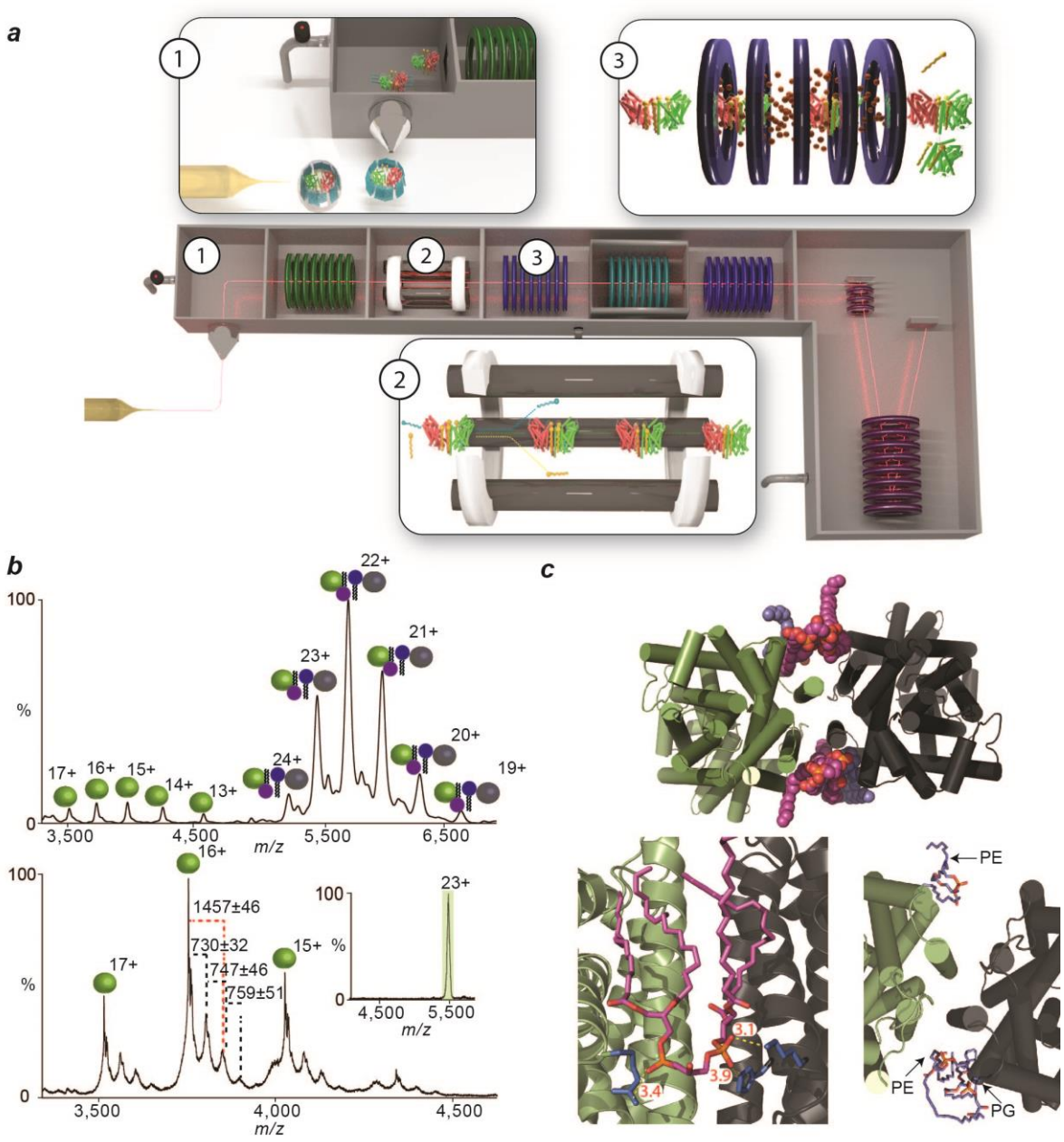



**Figure 2 | Schematic of the high-energy tandem mass spectrometry (HE-MS) platform, mass spectra and molecular dynamics (MD) simulations of the lipid-bound LeuT dimer.** ***a,*** Protein-lipid complexes (red/green rods, yellow) are desolvated and liberated from detergent micelles (blue) by a high potential difference applied to the cone (box 1). Quadrupole isolation of the protein-lipid complex is used to filter out detergent molecules and bulk lipids (box 2). Dissociation of the protein-lipid complex is then achieved via collision with neutral gas in the collision cell (box 3). Masses of the constituent protein and lipids are then measured in the Time-of-Flight mass analyser. ***b,*** Mass spectrum of LeuT liberated from OG micelles (upper panel) shows monomers (green) and dimers (green/grey) with a mass 7.4 kDa greater than the sequence mass. MS/MS of 23+ charge state of LeuT (bottom panel) reveals monomers with cardiolipin (CDL, purple head-group) and phospholipid (PL, blue head-group) retained. Masses of the lipids bound are marked in black (for PL) and red (for CDL) ***c,*** MD simulations of LeuT in an *E. coli* lipid bilayer reveal possible binding sites of interfacial

phospholipids and CDL (upper panel, viewed from cytoplasmic side of membrane). The CDL phosphate groups (orange) interact closely with positively charged residues (K376, H377, R506, blue) at the dimer interface. Interactions are shown (yellow dotted lines) with distances measured in Å (red). Phosphoethanolamine (PE) and phosphatidylglycerol (PG) also bind at the dimer interface (bottom right panel).

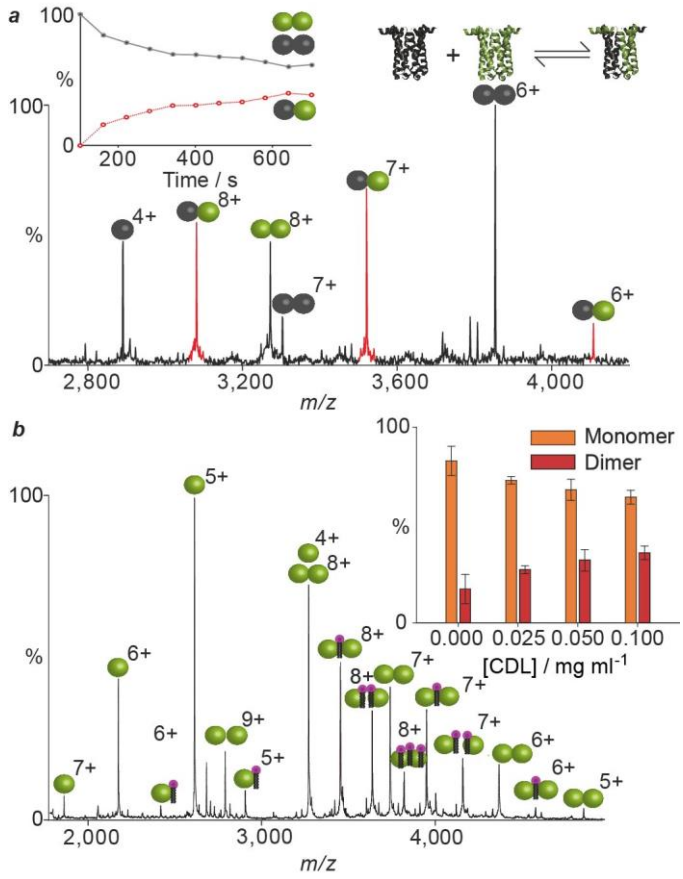



**Figure 3 | Mass spectra recorded for two different forms SemiSWEET that undergo subunit exchange and the effect of cardiolipin on the monomer dimer equilibrium.** ***a,*** Representative mass spectrum recorded after incubation in solution of an equimolar ratio of deca-His tagged and untagged SemiSWEET (green and black spheres, respectively), liberated from tetraethyleneglycolmonooctyl ether (C8E4) micelles. Plot of the percentage abundance of hetero- and homo- dimers over time (inset) SemiSWEET heterodimers (red trace, peaks highlighted red in mass spectrum) and homodimers (black trace), revealing the solution phase monomer-dimer equilibrium (right inset, PDB ID 4QND). ***b,*** Mass spectrum of SemiSWEET following incubation with cardiolipin (CDL, purple head-groups). Plot of CDL concentration (mg $ml^{-1}$) versus the percentage of monomer or dimer observed in mass spectra at various CDL concentrations is shown (bar chart).

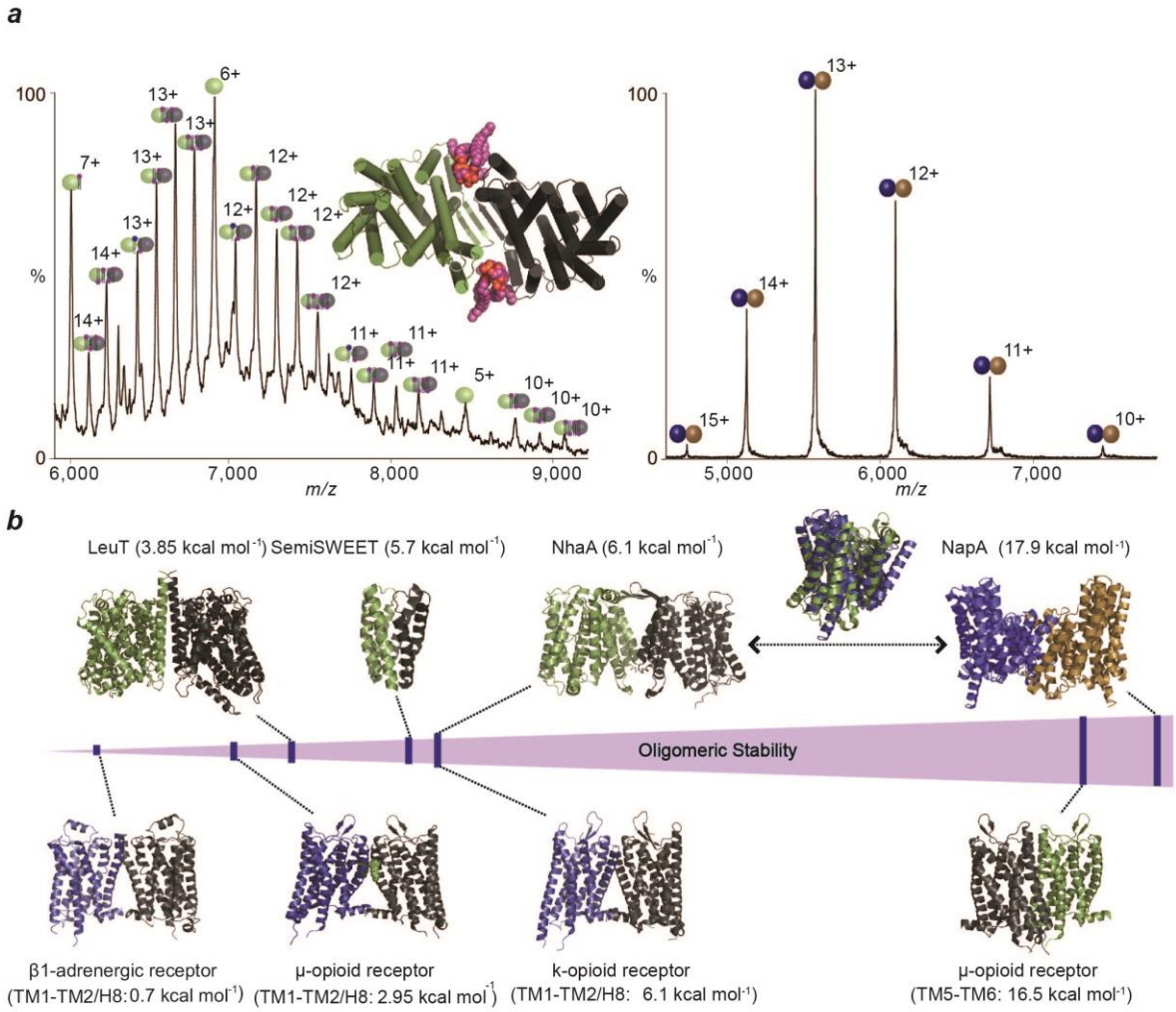



**Figure 4 | Comparison of mass spectra of NhaA with NapA and plot of $\Delta G_{dissoc}$ values of transporters studied here with G-protein coupled receptors.** ***a,*** Mass spectrum of NhaA (green/grey spheres), liberated from C8E4 micelles, reveals binding of cardiolipin (CDL, purple head-groups) and phospholipids (PL, blue head-groups) to the intact NhaA dimers. MD simulation of NhaA (green/black rods) in an *E. coli* lipid bilayer (inset) reveals interfacial CDL binding (orange/purple spheres). Mass spectrum of dimeric NapA (right), liberated from C8E4 micelles, shows NapA dimers (blue/brown spheres) with no lipid binding. ***b,*** Oligomeric stability scale (purple) annotated with crystal structures of proteins analysed in this study (above) and GPCRs (below) - $\Delta G°_{dissoc}$ values are given in brackets. Structures of the protomers of NhaA (green) and NapA (blue) are superimposed. This highlights the additional helix in NapA. PDB IDs: 2A65 (LeuT), 4QND (SemiSWEET), 4AU5 (NhaA), 4BWZ (NapA), 4GPO (β1 adrenergic receptor), 4DKL (μ-opioid receptor), 4DJH (κ-opioid receptor).

References


1. Arkhipov, A., et al., *Architecture and membrane interactions of the EGF receptor.* Cell, 2013. **152**(3): p. 557-69.
2. Wu, H., et al., *Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator.* Science, 2014. **344**(6179): p. 58-64.
3. Duarte, J.M., et al., *An analysis of oligomerization interfaces in transmembrane proteins.* BMC Struct Biol, 2013. **13**: p. 21.
4. Piscitelli, C.L., H. Krishnamurthy, and E. Gouaux, *Neurotransmitter/sodium symporter orthologue LeuT has a single high-affinity substrate site.* Nature, 2010. **468**(7327): p. 1129-32.
5. Penmatsa, A. and E. Gouaux, *How LeuT shapes our understanding of the mechanisms of sodium-coupled neurotransmitter transporters.* J Physiol, 2014. **592**(5): p. 863-9.
6. Xu, Y., et al., *Structures of bacterial homologues of SWEET transporters in two distinct conformations.* Nature, 2014. **515**(7527): p. 448-52.
7. Yeagle, P.L., *Non-covalent binding of membrane lipids to membrane proteins.* Biochim Biophys Acta, 2014. **1838**(6): p. 1548-59.
8. Hansen, S.B., X. Tao, and R. MacKinnon, *Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2.* Nature, 2011. **477**(7365): p. 495-8.
9. Gao, Y., et al., *TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action.* Nature, 2016.
10. Whitelegge, J.P., *Integral membrane proteins and bilayer proteomics.* Anal Chem, 2013. **85**(5): p. 2558-68.
11. Savas, J.N., et al., *Mass spectrometry accelerates membrane protein analysis.* Trends Biochem Sci, 2011. **36**(7): p. 388-96.
12. Konijnenberg, A., et al., *Extending native mass spectrometry approaches to integral membrane proteins.* Biol Chem, 2015. **396**(9-10): p. 991-1002.
13. Laganowsky, A., et al., *Membrane proteins bind lipids selectively to modulate their structure and function.* Nature, 2014. **510**(7503): p. 172-5.
14. Landreh, M., et al., *A sliding selectivity scale for lipid binding to membrane proteins.* Curr Opin Struct Biol, 2016. **39**: p. 54-60.
15. Pliotas, C., et al., *The role of lipids in mechanosensation.* Nat Struct Mol Biol, 2015. **22**(12): p. 991-8.
16. Ilgu, H., et al., *Variation of the detergent-binding capacity and phospholipid content of membrane proteins when purified in different detergents.* Biophys J, 2014. **106**(8): p. 1660-70.
17. Brugger, B., *Lipidomics: analysis of the lipid composition of cells and subcellular organelles by electrospray ionization mass spectrometry.* Annu Rev Biochem, 2014. **83**: p. 79-98.
18. Taki, T., *TLC-Blot (Far-Eastern Blot) and Its Application to Functional Lipidomics.* Methods Mol Biol, 2015. **1314**: p. 219-41.
19. Cox, J. and M. Mann, *Quantitative, high-resolution proteomics for data-driven systems biology.* Annu Rev Biochem, 2011. **80**: p. 273-99.
20. Barrera, N.P., et al., *Micelles protect membrane complexes from solution to vacuum.* Science, 2008. **321**(5886): p. 243-6.
21. Tan, B.K., et al., *Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates.* Proc Natl Acad Sci U S A, 2012. **109**(41): p. 16504-9.
22. Audet, M. and M. Bouvier, *Restructuring G-protein- coupled receptor activation.* Cell, 2012. **151**(1): p. 14-23.
23. Huang, W., et al., *Structural insights into micro-opioid receptor activation.* Nature, 2015. **524**(7565): p. 315-21.

24. Sounier, R., et al., *Propagation of conformational changes during mu-opioid receptor activation.* Nature, 2015. **524**(7565): p. 375-8.
25. Dorsch, S., et al., *Analysis of receptor oligomerization by FRAP microscopy.* Nat Methods, 2009. **6**(3): p. 225-30.
26. Drew, D. and O. Boudker, *Shared Molecular Mechanisms of Membrane Transporters.* Annu Rev Biochem, 2016. **85**: p. 543-72.
27. Herz, K., et al., *Beta-sheet-dependent dimerization is essential for the stability of NhaA Na+/H+ antiporter.* J Biol Chem, 2009. **284**(10): p. 6337-47.
28. Anderluh, A., et al., *Single molecule analysis reveals coexistence of stable serotonin transporter monomers and oligomers in the live cell plasma membrane.* J Biol Chem, 2014. **289**(7): p. 4387-94.
29. Zhen, J., et al., *Dopamine transporter oligomerization: impact of combining protomers with differential cocaine analog binding affinities.* J Neurochem, 2015. **133**(2): p. 167-73.
30. Coleman, J.A., E.M. Green, and E. Gouaux, *X-ray structures and mechanism of the human serotonin transporter.* Nature, 2016. **532**(7599): p. 334-9.
31. Kilic, F. and G. Rudnick, *Oligomerization of serotonin transporter and its functional consequences.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3106-11.

## Online methods

### Molecular cloning and plasmid construction

LeuT was expressed from a pET-16b vector, containing a thrombin cleavable C-terminal 8 x His tag. SemiSWEET was expressed from a pJexpress411 vector, containing a HRV-3C protease cleavable 10 x His tag. All point mutations were generated using a Quikchange Lightning Site-Directed mutagenesis kit, according to the manufacturer's protocol. For expression of LeuT in cardiolipin-deficient BKT22 *E. coli* strain, the LeuT gene was amplified by polymerase chain reaction using a Phusion Flex Hot Start Polymerase (New England Biolabs) with primers designed for Infusion cloning using the manufacturer's online tool. The PCR product was purified using gel agarose electrophoresis, then used in an Infusion cloning reaction (Clontech) with a linearised pBAD vector containing a 10 x His tag. The LeuT-eYFP fusion protein was expressed from a pET-15b/pET-23b hybrid vector, containing the LeuT gene followed by a TEV protease cleavage site, the eYFP gene (with mutation A206K to abolish eYFP dimerisation) and a 6 x His tag. To construct this plasmid, the LeuT gene was amplified by PCR as above, and used in an Infusion reaction with a pET-15b/pET-23b hybrid vector (Laganowsky, Nature 2014), cut with appropriate restriction enzymes, and a synthetic gene block (IDT) containing the TEV site and eYFP gene. All plasmid constructs were confirmed by DNA sequencing.

### Membrane protein expression and purification

The LeuT plasmid was transformed into C43 *E. coli* (Lucigen), and expressed and purified as reported previously [32]. Briefly, multiple colonies were used to inoculate 100 ml of Terrific Broth (TB) and grown overnight at 37 °C. 10 ml of overnight culture was used to inoculate each of 6 litres of TB, which were allowed to grow at 37 °C until the culture reached an $OD_{600}$ nm of 0.6. Isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM, and the cultures grown for 16 hours at 20 °C. Cells were harvested by centrifugation at 5,000 g, for 10 min at 4 °C, resuspended in lysis buffer solution (300 mM sodium chloride, 20 mM Tris, pH 7.4) supplemented with protease inhibitor cocktail (Roche), lysed using a M-110 PS microfluidizer (Microfluidics), and the cell debris pelleted by centrifugation at 20,000 g for 25 min at 4 °C. Membranes were pelleted by centrifugation of the supernatant at 100,000 g for 2h at 4 °C and subsequently resuspended in ice-cold Buffer soluition (100 mM sodium chloride, 20 mM Tris, 20 % glycerol, pH 8.0) and homogenised using a Potter-Elvehjem Teflon pestle and glass tube. DDM was added to resuspended membranes to a final concentration of 2% w/v and the suspension incubated with gentle agitation for 15 hours at 4 °C. Insoluble material was pelleted by centrifugation at 20,000g for 30 min and the supernatant filtered through 0.22 micron filters. LeuT was purified by immobilised metal ion-affinity chromatography using a HisTrap HP 5 ml column (GE healthcare) equilibrated with Buffer A (190 mM sodium chloride, 10 mM potassium chloride, 20 mM Tris, 20 mM imidazole, 10 % glycerol, 0.02 % DDM, pH 8.0) and eluted with Buffer B (190 mM sodium chloride, 10 mM potassium chloride, 20 mM Tris, 500 mM imidazole, 10 % glycerol, 0.02 % DDM, pH 8.0). The eluted protein was transferred to a dialysis cassette (100 kDa molecular weight cut-off) and dialysed against a dialysis buffer solution (190 mM sodium chloride, 10 mM potassium chloride, 20 mM Tris,10 % glycerol, pH 8.0) + 0.02% DDM overnight. A 100 kDa MWCO concentrator was used to concentrate the dialysed protein. LeuT was then injected onto a Superdex 200 Increase GL 10/300 column (GE Healthcare),

equilibrated in a buffer (190 mM sodium chloride, 10 mM potassium chloride, 20 mM Tris,10 % glycerol, pH 8.0) with 1 % OG. Peak fractions containing OG-solubilised LeuT were concentrated as above and used for further study. All protein concentration measurements were carried out using a UV/vis spectrophotometer (DS-11 +, DeNovix).

The *Vibrio sp.* semiSWEET plasmid was transformed into BL21 DE3 *E. coli* (Novagen), and expressed and purified as reported previously [6]. Briefly, multiple colonies were used to inoculate 100 ml of LB and grown overnight at 37 °C. 10 ml of overnight culture was used to inoculate each of 6 liters of LB, which were allowed to grow at 37 °C until the culture reached an OD600nm of 0.8. Isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM, and the culture grown for 15 hours at 22 °C. Cell harvesting, resuspension, lysis, membrane isolation and detergent extraction steps were identical to the LeuT purification. semiSWEET was first purified by IMAC using a HisTrap HP 5 ml column (GE healthcare) equilibrated with Buffer A (150 mM sodium chloride, 20 mM Tris, 20 mM imidazole, 10 % glycerol, 1 mM DTT, 0.02 % DDM, pH 8.0) and eluted with Buffer B (150 mM sodium chloride, 20 mM Tris, 500 mM imidazole, 10 % glycerol, 1 mM DTT, 0.02 % DDM, pH 8.0). The eluted protein and HRV 3C protease (Novagen) were transferred to a dialysis cassette (30 kDa molecular weight cut-off), and dialysed against buffer (150 mM sodium chloride, 20 mM Tris,10 % glycerol, 1 mM DTT, pH 8.0) with 0.02% DDM overnight. For preparation of 10 x His semiSWEET, no HRV 3C protease was added to the dialysis cassette. A 50 kDa MWCO concentrator was used to concentrate the dialysed protein. semiSWEET was then injected onto a Superdex 200 Increase GL 10/300 column (GE Healthcare), equilibrated in a buffer solution (150 mM sodium chloride, 20 mM Tris, 10 % glycerol, 1 mM DTT, 0.02 % DDM, pH 8.0) + 0.5 % C8E4. Peak fractions containing C8E4-solubilised semiSWEET were concentrated as above and used for further study.

NapA, NhaA, AqpZ, AmtB and ELIC were expressed as described before [13, 33-35].

**Non-denaturing mass spectrometry**

Samples were prepared for non-denaturing mass spectrometry by buffer-exchange into MS buffer (200 mM ammonium acetate, 2 x CMC of detergent of interest, pH 7.4) using a centrifugal buffer exchange device (Micro Bio-Spin, Biorad). For semiSWEET experiments, 1 mM DTT was also added to the MS buffer.

Mass spectrometry measurements were performed on a Synapt G1(Waters) with a Z-spray source, using nanoelectrospray capillaries prepared in-house [36]. The source pressure was set to 4-7 mBar, with a capillary voltage of 1.4-1.7 kV, capillary nanoflow of 0.05 - 0.2 mBar and argon as collision (trap) gas at flow rate of 1.5-8.0 ml $min^{-1}$. Other parameters, including the sample and extraction cone and trap bias voltages, collision voltages and quadrupole profile were optimised for maximal ion intensity and minimal dissociation of the target membrane protein complex. Data was processed using MassLynx software.

**High energy non-denaturing mass spectrometry**

To allow the use of higher voltages on the extraction cone in the source region of the Synapt G2, the configuration file for the extraction cone was modified to increase the maximum voltage setting from 10V to 200V. However, altering this setting alone would restrict the maximum sample cone

voltage that could be accessed due to the limits imposed by the power supplies. To overcome this limitation, the capability to drive the sample cone voltage from an external supply was implemented. This took the form of a patch cable introduced between the instrument lens control PCB and the source ion block. The ion block contains both heater elements and a thermocouple in addition to supporting the extraction cone and these were decoupled with the patch cable to prevent possible electrical breakdown with use of higher cone voltages. The extraction cone was patched directly through from the lens PCB whilst the sample cone voltage was decoupled and a new wire connection made to an external power supply.

**Delipidation**

Purified, OG-solubilised LeuT was incubated in 2% NG overnight at 4 °C. Subsequently, the sample passes through a Superdex 200 Increase GL 10/300 column (GE Healthcare) equilibrated in 200mM ammonium acetate with twice CMC amount of NG, to remove the excess OG and NG. This sample was subjected to MS analysis. The delipidated LeuT in NG was re-exchanged back in OG using the above protocol with, the 200mM ammonium acetate containing 1% OG. *E. coli* polar lipid stocks were made from powder (Avanti Polar Lipids Inc., Alabama USA) at a concentration of 10mg/ml using previously published methods [13] and subsequently diluted 50 times in 200mM ammonium acetate solution containing 1%OG. A dilyso-CDL stock of (10 mg/ml in 200 mM ammonium acetate) was prepared as described previously [35]; aliquots of this stock were diluted 50 times in 200mM ammonium acetate solution containing 1%OG for each MS experiment.

**Preparation and titration of phospholipids**

Purified semiSWEET-$His_{10}$ in MS buffer was diluted to an oligomer concentration of 20 µM. A cardiolipin stock (10 mg/ml in 200 mM ammonium acetate) was prepared as described previously [35]; aliquots of this stock were diluted in MS buffer to a CDL concentration 2x that required for each MS experiment. Diluted semiSWEET and cardiolipin solutions were then mixed 1:1 and incubated on ice for 5 mins (for the measurement without CDL, semiSWEET was mixed with MS buffer). 2 µL of this mixture was then used for each of three MS measurements at the 4 lipid concentrations.

Data was acquired for 100 scans, which were summed in MassLynx software and processed using UniDec deconvolution software [37]. Relative monomer and dimer abundances were calculated by taking the sum of the respective deconvoluted intensities in all lipidation (and non-lipidated) states, normalised to the total intensity of all semiSWEET species, and averaged over the 3 repeats (error bars are +/- 1 standard deviation). A plot of relative monomer and dimer abundances against cardiolipin concentration was generated using SigmaPlot.

MS measurements were performed on a Synapt G1 as described above. The source pressure was set to 4.2 mBar, capillary nanoflow 0.1 mBar, trap collision voltage 50 V, transfer collision voltage 10 V and collision gas flow rate 1.8 ml $min^{-1}$.

**Subunit exchange**

Purified, C8E4-solubilised semiSWEET and semiSWEET-His10 were separately buffer-exchanged into MS buffer as described above, then concentrated to 40 µM using a 30 kDa MWCO concentrator. 3 µL

of semiSWEET and semiSWEET-$His_{10}$ were then mixed briefly on ice. 3 µL of this equimolar mixture was immediately transferred to a nanoelectrospray capillary for MS data acquisition. MS data was acquired continuously for 10 minutes.

Data was processed using Xcalibur software (Thermo Scientific) as follows: spectra were extracted from summation of the chromatogram in 30 second scan windows centred on each minute e.g. 0.75 - 1.25 min for the 1 min time point. Relative abundances of the semiSWEET and semiSWEET-$His_{10}$ homodimers and the semiSWEET.semiSWEET-$His_{10}$ heterodimer were calculated for each time point using UniDec. A plot of relative homodimer abundance and heterodimer abundance against time was generated using SigmaPlot.

MS measurements were performed on a modified Q-Exactive orbitrap mass spectrometer (Thermo Fisher, Bremen Germany) [38] modified and optimised for non-denaturing MS of membrane protein complexes. Spectra were acquired in "Native Mode" with maximum RF applied to all ion optics, -3.2 kV to the central electrode of the Orbitrap and with ion trapping in the HCD cell. Ions were generated in positive ion mode from a static nanospray source using gold-coated capillaries prepared in-house. Transient times were 64 ms and AGC target was $1\times10^6$. Spectra were acquired with 1 microscan, a noise level parameter set to 3 and HCD cell voltage of 75 V; no in-source activation was applied. The collision gas was Argon and pressure in the HCD cell was maintained at approximately $1\times10^{-9}$ mbar.

**Molecular dynamics simulations**

All MD simulations were performed using GROMACS v5.1.2 [39]. The MemProtMD pipeline [40] was used with the Martini 2.2 force field [41] to run five repeats of a 1 µs Coarse Grained (CG) MD simulation of the dimeric protein complexes. Last 800ns of each of these simulation trajectories were considered for further analysis. The proteins were centered within the simulation system to permit the assembly and equilibration of 10 % cardiolipin with either a 20% 1-palmitoly, 2-oleoyl, phosphatidylglycerol (POPG): 70% 1-palmitoyl, 2-oleoyl, phosphatidylethanolamine (POPE) or 90 % 1-palmitoyl, 2-oleoyl, phosphatidylcholine (POPC) bilayers. Systems were neutralised with a 150 mM concentration of NaCl. All simulations were performed at 323K, with protein, lipids and solvent separately coupled to an external bath using the velocity-rescale thermostat [42]. Pressure was maintained at 1 bar with a semi-isotropic compressibility of $5 \times 10^{-6}$ using the Berendsen barostat [43]. All bonds were constrained with the P-LINCS algorithm [44]. Electrostatics was measured using the Reaction Field method [45], while a Verlet cut-off scheme to permit GPU calculation of non-bonded contacts was used for Lennard-Jones parameters [46]. Simulations were performed with an integration time-step of 20 fs. Atomistic snapshots at the end-point of the CGMD simulations were created by using CG2AT [47] in combination with Alchembed [48]. The lipid densities and contacts with the protein during the MD simulations were calculated using MDAnalysis [49], and locally written code. All images and animations were generated using Pymol [50].

**Estimation of Oligomeric stability and box plot analysis: Calculation of $\Delta G°_{dissoc(m)}$ , buried interface, salt bridge and $\Delta G°_{stab}$ calculation**

A database of integral oligomeric membrane proteins, for which crystal structures have been obtained, was created from the MPStruc database (*http://blanco.biomol.uci.edu/mpstruc/*).

Subsequently, $\Delta G^{\circ}_{dissoc}$ , buried surface area and the number of salt bridge of interactions were calculated using PDBePISA webserver [51].

Further, as we are effectively analysing the equilibrium of the form $n\,M_1 \rightleftharpoons M_n$ , the obtained $\Delta G^{\circ}_{dissoc}$ values cannot be compared across different oligomeric states.

To address this concern, we adjusted the equilibria to the form $M_1 \rightleftharpoons \frac{1}{n} M_n$, which can be physically interpreted as a free energy change per monomer on joining the complex. This adjustment approximately corrects, for example, for unbalanced translational free energy (PNAS, 2010, 107: 2007).

So the new $\Delta G^{\circ}_{dissoc(m)}$ is calculated as: $\Delta G^{\circ}_{dissoc(m)} = \frac{1}{n} \Delta G^{\circ}_{dissoc}$ Where n = Oligomeric state, as described in and $\Delta G^{\circ}_{dissoc}$ is the value obtained from PISA.

$\Delta G^{\circ}_{dissoc(m)}$ were visualised using a scatter generated using the Matplotlib library for the Python programming language. Superimposed on this is a circle for each protein, whose vertical position was determined by its $\Delta G^{\circ}_{dissoc(m)}$ value and horizontal position by its oligomeric state. To reduce overlap of points, an arbitrary horizontal jitter (sampled randomly from a uniform distribution) was applied to each point and the Y-axis was plotted in log scale. Same protocol was followed to plot total buried surface area and the number of salt bridges against the oligomeric state.

We further validated this approach by calculating interface strength through buried surface area and salt bridge interactions, both of which were proposed to be reasonable proxies of interface strength (*J. Mol. Biol*, 1976, 105, 1; Nature, 2008, 453: 1262; J R Soc. Interface, 2012, 10: 20120835; Protein Science, 2013, 22: 510).

In PISA, the effect of buried surface area and number of slat bridges are cumulated by multiplying the buried surface area by atomic solvation parameters (ASP) and number of salt bridge by $E_{sb}$

$$\Delta G^{\circ}_{stab} = \sum (\sigma \times \omega) + \sum (N_{sb} \times E_{sb}) \quad \text{(i)}$$

Where σ = buried surface area, ω= Atomic solvation parameter (ASP), $N_{sb}$ = total number of salt bridges, $E_{sb}$ = energy contribution of each of the salt bridges and $\Delta G^{\circ}stab$ = stabilization energy. The more positive the value of $\Delta G^{\circ}_{stab}$ the more stable the oligomeric species

In PISA, the $E_{sb}$ value is assigned as 0.15 kcal/mol based on a bulk water dielectric constant of 80. As suggested by the reviewer, for the current data set of integral membrane proteins, such interactions in the bilayer would scale up due to the very low dielectric of the medium. We have thus now modified this value by a factor of 40 (0.15kcal/mol × 40 = 6 kcal/mol) to account for the lower dielectric constant in the bilayer (ε = 2) (PNAS USA, 1984, 84: 5412). Likewise, the buried surface area contribution was converted to an energy contribution based on an ASP value of 7 cal/Å (PNAS USA, 2011, 31: E359).

Hence the new $\Delta G^{\circ}_{stab}$ revised for the bilayer environment is calculated as

$$\Delta G^{\circ}_{stab} = \frac{\sigma \times 7}{1000} + N_{sb} \times 6 \quad \text{(ii)}$$

For reasons mentioned above the obtained was divided by the oligomeric state and plotted against the oligomeric state using the same protocol.

**LeuT/SERT Structure alignment**

The structure of LeuT (PDB ID 2A65) and SERT (5I6Z) were superimposed using PyMOL and the sequence alignment of LeuT with the BATs were done using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). The individual sequences were obtained from Uniprot.

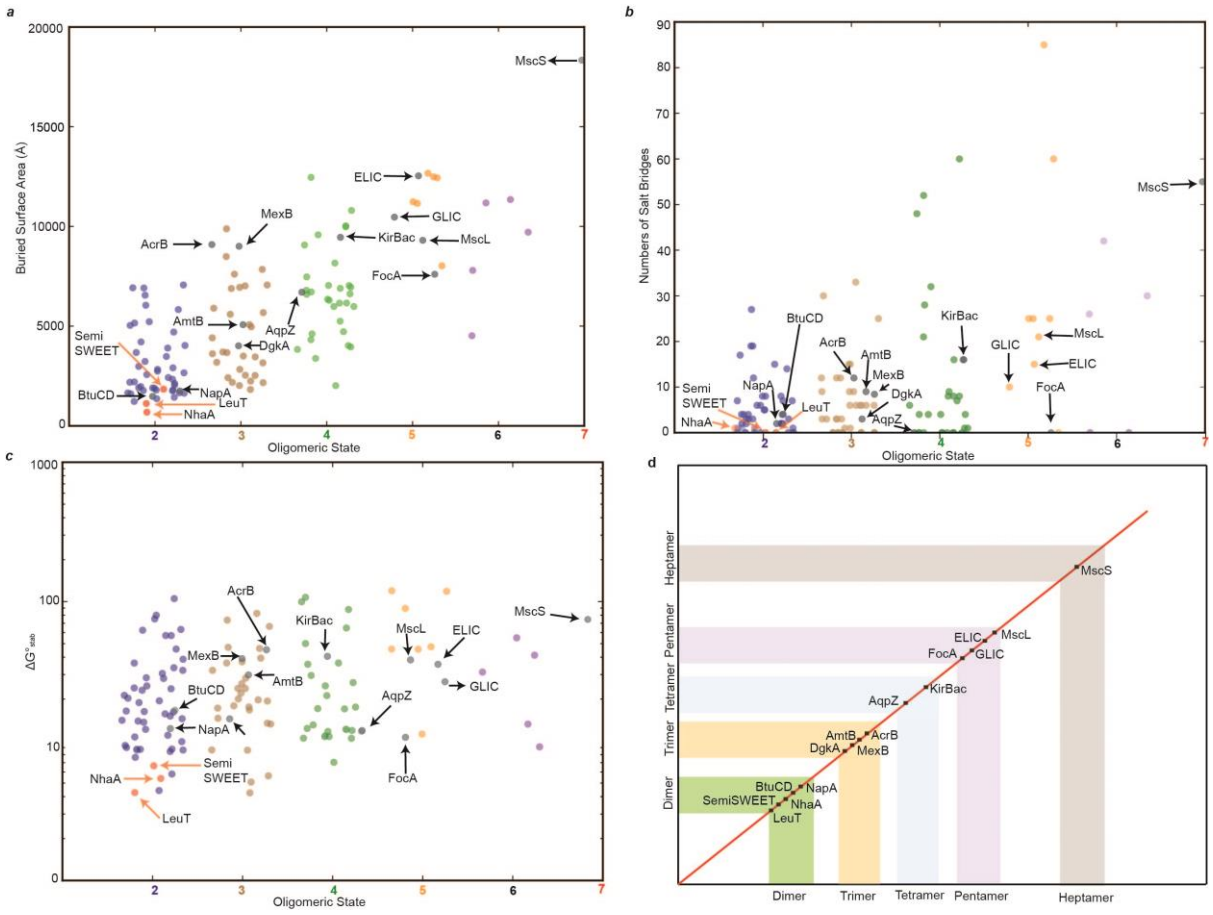

1. Arkhipov, A., et al., *Architecture and membrane interactions of the EGF receptor.* Cell, 2013. **152**(3): p. 557-69.
2. Wu, H., et al., *Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator.* Science, 2014. **344**(6179): p. 58-64.
3. Duarte, J.M., et al., *An analysis of oligomerization interfaces in transmembrane proteins.* BMC Struct Biol, 2013. **13**: p. 21.
4. Piscitelli, C.L., H. Krishnamurthy, and E. Gouaux, *Neurotransmitter/sodium symporter orthologue LeuT has a single high-affinity substrate site.* Nature, 2010. **468**(7327): p. 1129-32.
5. Penmatsa, A. and E. Gouaux, *How LeuT shapes our understanding of the mechanisms of sodium-coupled neurotransmitter transporters.* J Physiol, 2014. **592**(5): p. 863-9.
6. Xu, Y., et al., *Structures of bacterial homologues of SWEET transporters in two distinct conformations.* Nature, 2014. **515**(7527): p. 448-52.
7. Yeagle, P.L., *Non-covalent binding of membrane lipids to membrane proteins.* Biochim Biophys Acta, 2014. **1838**(6): p. 1548-59.
8. Hansen, S.B., X. Tao, and R. MacKinnon, *Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2.* Nature, 2011. **477**(7365): p. 495-8.
9. Gao, Y., et al., *TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action.* Nature, 2016.
10. Whitelegge, J.P., *Integral membrane proteins and bilayer proteomics.* Anal Chem, 2013. **85**(5): p. 2558-68.
11. Savas, J.N., et al., *Mass spectrometry accelerates membrane protein analysis.* Trends Biochem Sci, 2011. **36**(7): p. 388-96.
12. Konijnenberg, A., et al., *Extending native mass spectrometry approaches to integral membrane proteins.* Biol Chem, 2015. **396**(9-10): p. 991-1002.
13. Laganowsky, A., et al., *Membrane proteins bind lipids selectively to modulate their structure and function.* Nature, 2014. **510**(7503): p. 172-5.
14. Landreh, M., et al., *A sliding selectivity scale for lipid binding to membrane proteins.* Curr Opin Struct Biol, 2016. **39**: p. 54-60.
15. Pliotas, C., et al., *The role of lipids in mechanosensation.* Nat Struct Mol Biol, 2015. **22**(12): p. 991-8.
16. Ilgu, H., et al., *Variation of the detergent-binding capacity and phospholipid content of membrane proteins when purified in different detergents.* Biophys J, 2014. **106**(8): p. 1660-70.
17. Brugger, B., *Lipidomics: analysis of the lipid composition of cells and subcellular organelles by electrospray ionization mass spectrometry.* Annu Rev Biochem, 2014. **83**: p. 79-98.

18. Taki, T., *TLC-Blot (Far-Eastern Blot) and Its Application to Functional Lipidomics.* Methods Mol Biol, 2015. **1314**: p. 219-41.
19. Cox, J. and M. Mann, *Quantitative, high-resolution proteomics for data-driven systems biology.* Annu Rev Biochem, 2011. **80**: p. 273-99.
20. Barrera, N.P., et al., *Micelles protect membrane complexes from solution to vacuum.* Science, 2008. **321**(5886): p. 243-6.
21. Tan, B.K., et al., *Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates.* Proc Natl Acad Sci U S A, 2012. **109**(41): p. 16504-9.
22. Audet, M. and M. Bouvier, *Restructuring G-protein- coupled receptor activation.* Cell, 2012. **151**(1): p. 14-23.
23. Huang, W., et al., *Structural insights into micro-opioid receptor activation.* Nature, 2015. **524**(7565): p. 315-21.
24. Sounier, R., et al., *Propagation of conformational changes during mu-opioid receptor activation.* Nature, 2015. **524**(7565): p. 375-8.
25. Dorsch, S., et al., *Analysis of receptor oligomerization by FRAP microscopy.* Nat Methods, 2009. **6**(3): p. 225-30.
26. Drew, D. and O. Boudker, *Shared Molecular Mechanisms of Membrane Transporters.* Annu Rev Biochem, 2016. **85**: p. 543-72.
27. Herz, K., et al., *Beta-sheet-dependent dimerization is essential for the stability of NhaA Na+/H+ antiporter.* J Biol Chem, 2009. **284**(10): p. 6337-47.
28. Anderluh, A., et al., *Single molecule analysis reveals coexistence of stable serotonin transporter monomers and oligomers in the live cell plasma membrane.* J Biol Chem, 2014. **289**(7): p. 4387-94.
29. Zhen, J., et al., *Dopamine transporter oligomerization: impact of combining protomers with differential cocaine analog binding affinities.* J Neurochem, 2015. **133**(2): p. 167-73.
30. Coleman, J.A., E.M. Green, and E. Gouaux, *X-ray structures and mechanism of the human serotonin transporter.* Nature, 2016. **532**(7599): p. 334-9.
31. Kilic, F. and G. Rudnick, *Oligomerization of serotonin transporter and its functional consequences.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3106-11.
32. Yamashita, A., et al., *Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters.* Nature, 2005. **437**(7056): p. 215-23.
33. Coincon, M., et al., *Crystal structures reveal the molecular basis of ion translocation in sodium/proton antiporters.* Nat Struct Mol Biol, 2016. **23**(3): p. 248-55.
34. Lee, C., et al., *Crystal structure of the sodium-proton antiporter NhaA dimer and new mechanistic insights.* J Gen Physiol, 2014. **144**(6): p. 529-44.
35. Laganowsky, A., et al., *Mass spectrometry of intact membrane protein complexes.* Nat Protoc, 2013. **8**(4): p. 639-51.
36. Hernandez, H. and C.V. Robinson, *Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry.* Nat Protoc, 2007. **2**(3): p. 715-26.
37. Marty, M.T., et al., *Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles.* Anal Chem, 2015. **87**(8): p. 4370-6.
38. Gault, J., et al., *High-resolution mass spectrometry of small molecules bound to membrane proteins.* Nat Methods, 2016.
39. Abraham, M.J., et al., *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1–2**: p. 19-25.
40. Stansfeld, P.J., et al., *MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes.* Structure, 2015. **23**(7): p. 1350-61.
41. de Jong, D.H., et al., *Improved Parameters for the Martini Coarse-Grained Protein Force Field.* Journal of Chemical Theory and Computation, 2013. **9**(1): p. 687-697.
42. Bussi, G., D. Donadio, and M. Parrinello, *Canonical sampling through velocity rescaling.* J Chem Phys, 2007. **126**(1): p. 014101.

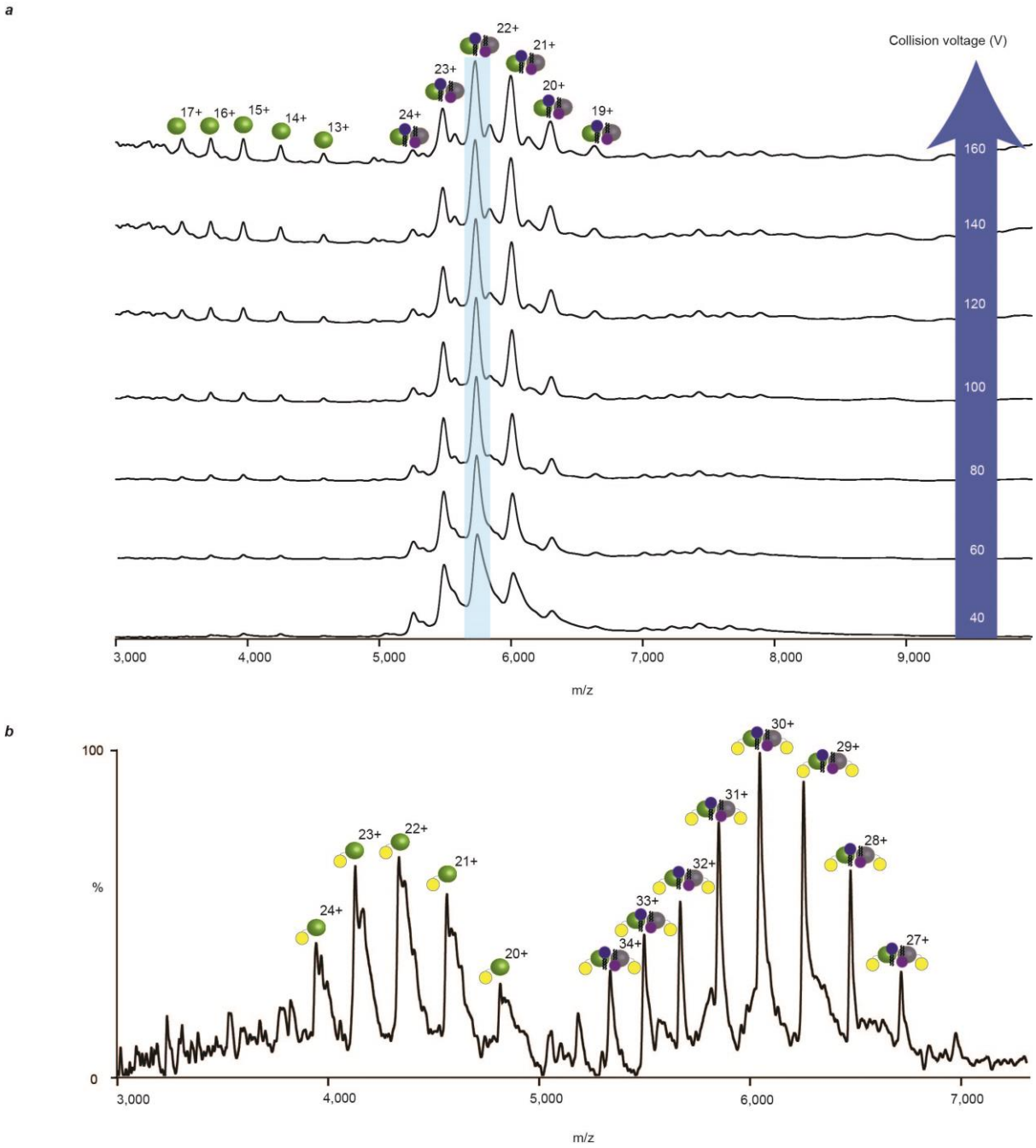
43. Berendsen, H.J.C., et al., *Molecular-Dynamics with Coupling to an External Bath.* Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.
44. Hess, B., *P-LINCS: A parallel linear constraint solver for molecular simulation.* Journal of Chemical Theory and Computation, 2008. **4**(1): p. 116-122.
45. Tironi, I.G., et al., *A Generalized Reaction Field Method for Molecular-Dynamics Simulations.* Journal of Chemical Physics, 1995. **102**(13): p. 5451-5459.
46. Pall, S. and B. Hess, *A flexible algorithm for calculating pair interactions on SIMD architectures.* Computer Physics Communications, 2013. **184**(12): p. 2641-2650.
47. Stansfeld, P.J. and M.S. Sansom, *Molecular simulation approaches to membrane proteins.* Structure, 2011. **19**(11): p. 1562-72.
48. Jefferys, E., et al., *Alchembed: A Computational Method for Incorporating Multiple Proteins into Complex Lipid Geometries.* J Chem Theory Comput, 2015. **11**(6): p. 2743-2754.
49. Michaud-Agrawal, N., et al., *Software News and Updates MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations.* Journal of Computational Chemistry, 2011. **32**(10): p. 2319-2327.
50. DeLano, W.L., *The PyMOL Molecular Graphics System.* 2002. **http://www.pymol.org**.
51. Krissinel, E. and K. Henrick, *Inference of macromolecular assemblies from crystalline state.* J Mol Biol, 2007. **372**(3): p. 774-97.
52. Sturt, H.F., et al., *Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry--new biomarkers for biogeochemistry and microbial ecology.* Rapid Commun Mass Spectrom, 2004. **18**(6): p. 617-28.

**Extended Data Table 1**: Summary of the mass spectral analysis and $\Delta G_{dissociation}$ membrane proteins forming strong oligomers

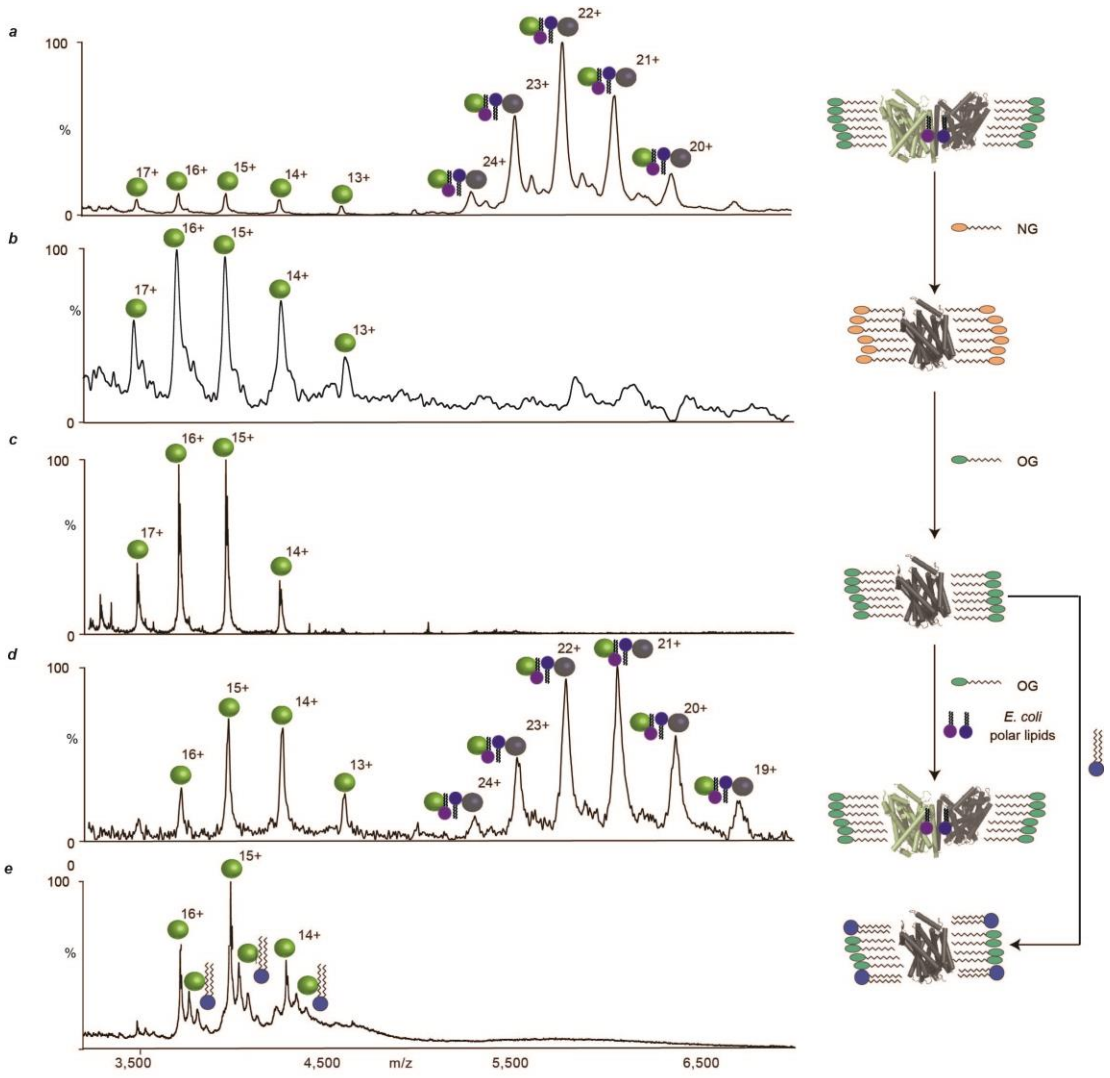
| Name | Expected Oligomeric State | Observed Oligomeric State | $\Delta G_{dissociation}$ Kcal/mol | Expected Mass (kDa) | Observed Mass k(Da) | Ref |
|---|---|---|---|---|---|---|
| AqpZ | 4 | 4 | 137.5 | $98.8x10^3$ | $98.9x10^3$ | This Work |
| MscS | 7 | 7 | 223.8 | $223.7x10^3$ | $224.3x10^3$ | [15] |
| MscL | 5 | 5 | 162.8 | $85.2 x10^3$ | $85.5 x10^3$ | [13, 53] |
| ELIC | 5 | 5 | 135.8 | $185.6 x10^3$ | $185.7 x10^3$ | This Work |
| AmtB | 3 | 3 | 120.5 | $126.8 x10^3$ | $126.7x10^3$ | This Work |
| NapA | 2 | 2 | 35.8 | $82.1x10^3$ | $82.1x10^3$ | This Work |
| DgKa | 3 | 3 | 64.5 | $42.7 x10^3$ | $42.7 x10^3$ | [54] |
| MexB | 3 | 3 | 67.1 | $342.4 x10^3$ | $344.2x10^3$ | [55] |
| KirBac 3.1 | 4 | 4 | 117 | $134.9 x10^3$ | $134.9 x10^3$ | [56] |
| FocA | 5 | 5 | 161.5 | $158.6 x10^3$ | $158.6 x10^3$ | [53] |
| AcrB | 3 | 3 | 78.4 | $342.9 x10^3$ | $342.6 x10^3$ | [53] |
| $BtuC_2D_2$ | 2 (membrane dimer) | 2 | 20.1 | $129.5 x10^3$ | $129.6 x10^3$ | [20] |
| EmrE | 2 | 2 | 20.9 | $24.7x10^3$ | $24.7x10^3$ | [55] |



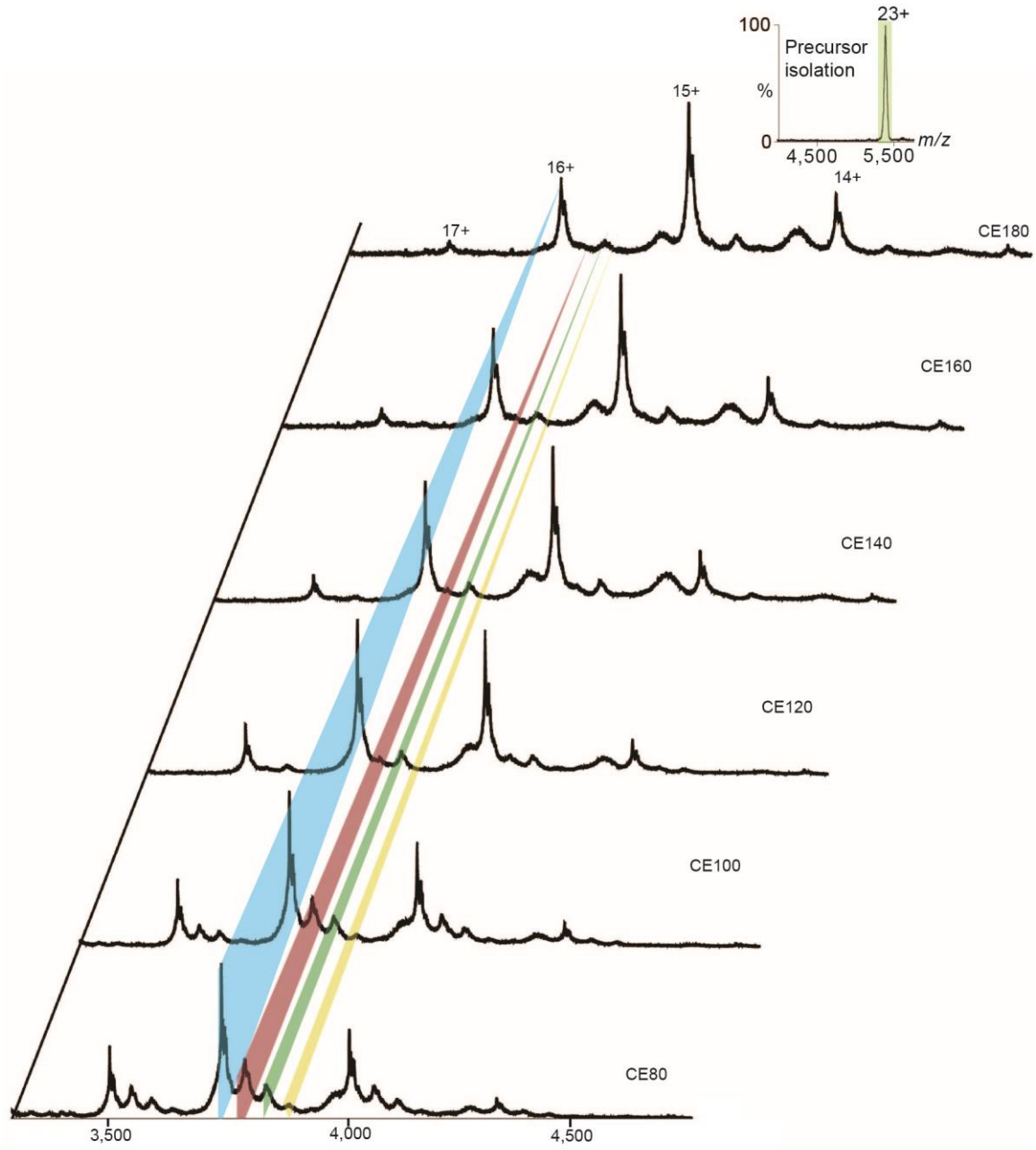

**Extended Data Figure 1 | Oligomeric interface stability of membrane proteins. a**, Buried surface area, **b**, the number of salt bridges and **c**, $\Delta G°_{stab}$ of oligomeric membrane proteins as a function of oligomeric state. Data points of different oligomeric states are independently coloured. The data points for LeuT, NhaA and SemiSWEET are coloured red. The analysis shows buried surface are, bumber of salt bridges or combine ($\Delta G°_{stab}$), irrespective of the matrix used that LeuT, NhaA and SemiSWEET rank amongst the weakest oligomeric membrane proteins. These three proteins contain no salt bridges. The 13 other proteins for which lipid-free assemblies have been observed by mass spectrometry are also labelled. **d**, shows the for all the 16 proteins used the PISA predicted oligomeric state matched with that of the experimentally determined oligomeric state.



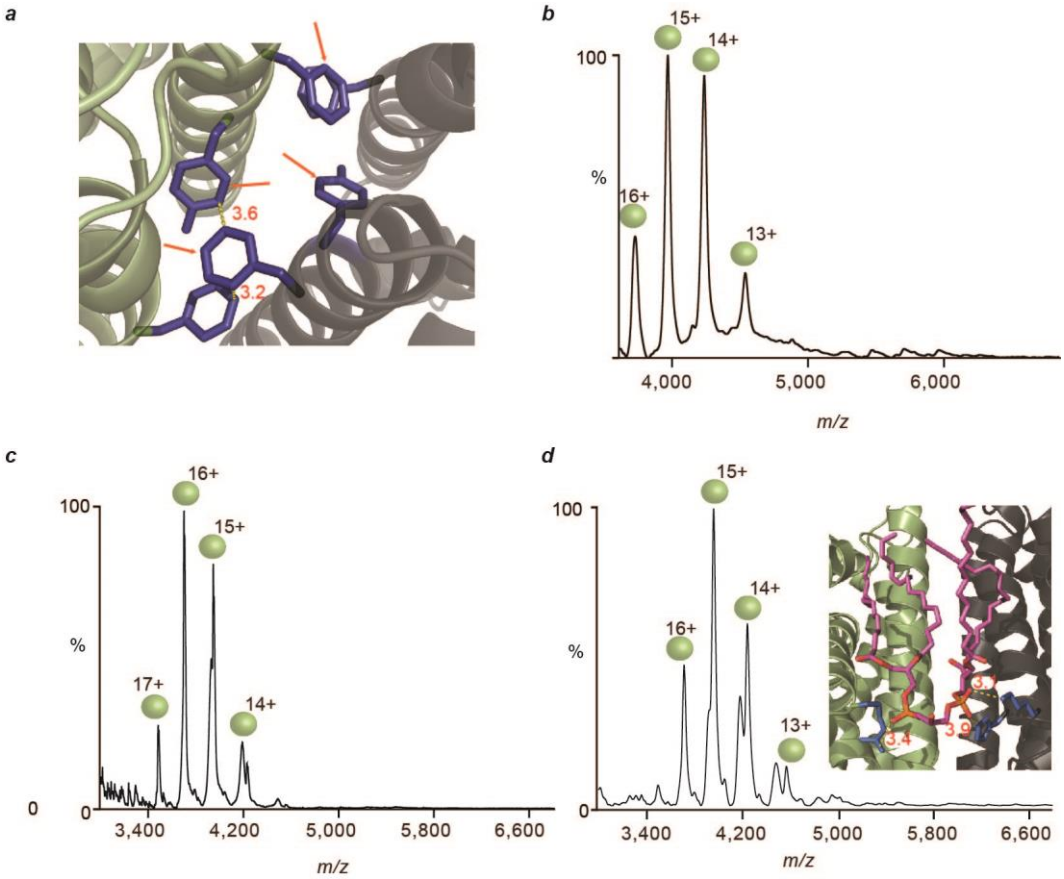

**Extended Data Figure 2 | Mass spectra of LeuT recorded with increasing collision voltages and of a LeuT fusion protein construct**. **a,** Mass spectra of LeuT, liberated from OG micelles, (green/grey spheres, most abundant charge state highlighted in pale blue), show that the 7.4 kDa lipid adduct (blue/purple head groups), is retained throughout the trap collision energy range (white, blue arrow) of the mass spectrometer. **b,** Mass spectra of LeuT expressed as a fusion protein with eYFP (LeuT-eYFP yellow circles), liberated from OG micelles, show that the dimer is similarly associated with a 7.4 kDa adduct.



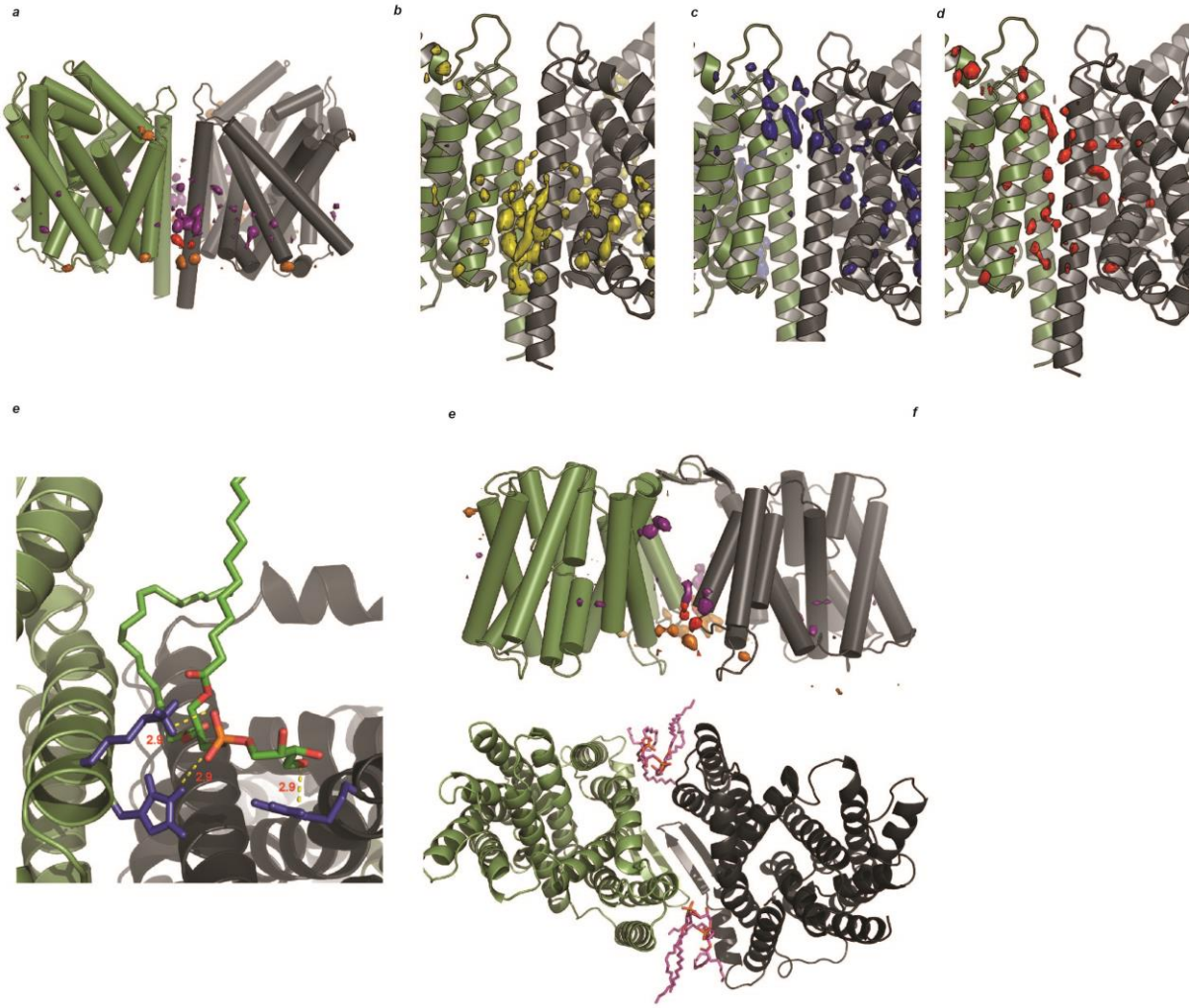

**Extended Figure 3 | Mass spectra of LeuT following incubation with delipidating detergents and *E. coli* polar lipids**. **a,** Mass spectrum of LeuT liberated from OG micelles (green head groups) shows low-abundance, delipidated monomers (green spheres, 59.3 kDa) and high-abundance, lipid-bound dimers (green/black spheres, 126.0 kDa). **b,** Mass spectrum of LeuT after incubation with neopentyl glycol (NG, orange head-groups) shows only delipidated monomers. **c,** Mass spectrum of LeuT after incubation with NG, then OG shows only delipidated monomers. **d,** Mass spectrum recorded after incubation of delipidated LeuT monomers with OG and *E. coli* polar lipids (blue/purple head-groups) shows delipidated monomers and lipid-bound dimers. **e**, Mass spectrum recorded after adding dilysocardiolipin (blue head-groups) to delipidated monomeric LeuT (**c**) shows no dimerisation in the presence of this lipid.



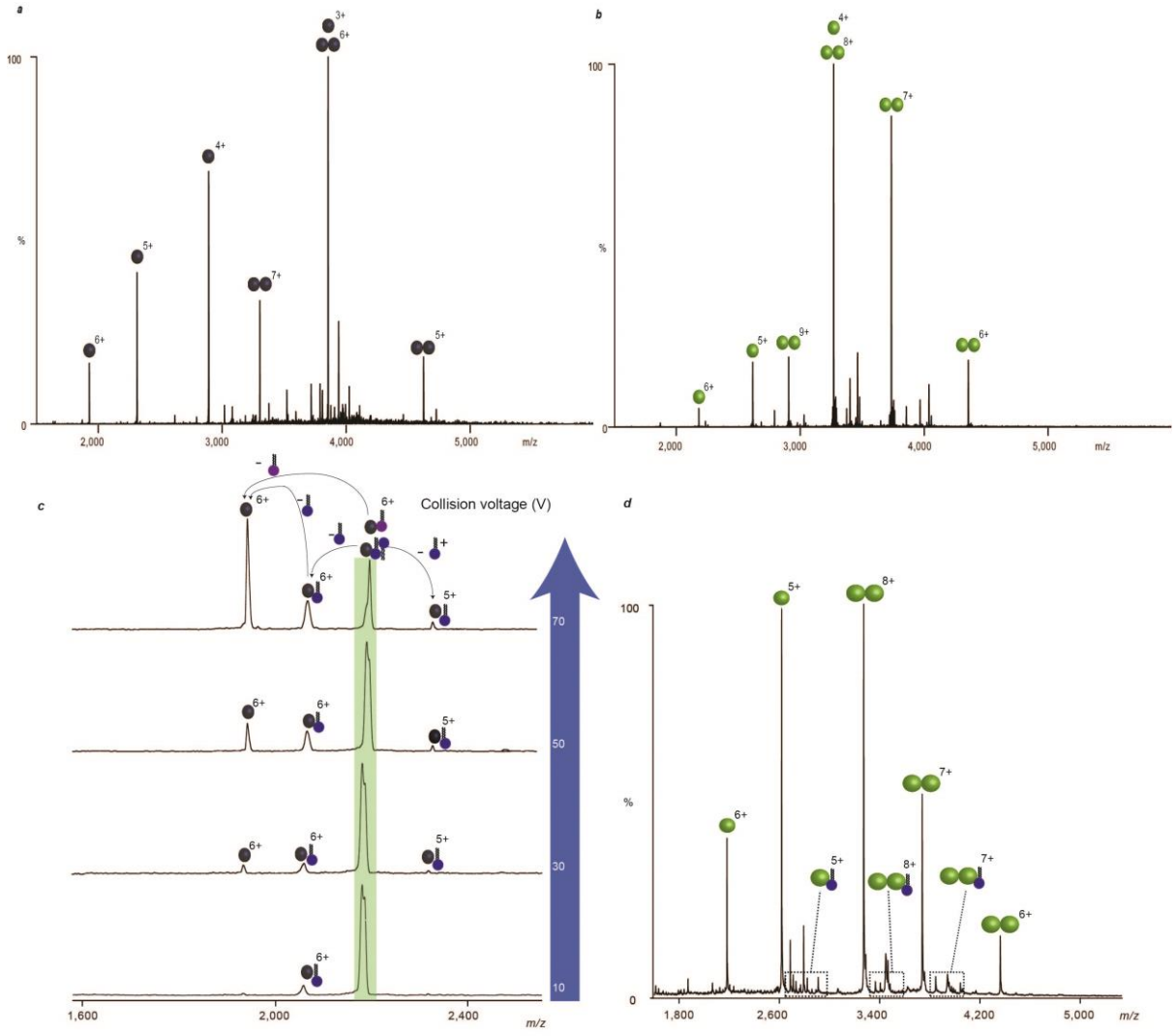

**Extended Figure 4**: **High energy MS/MS experiment of the 23+ charge state of dimeric LeuT, with the 7.4 kDa adduct, as a function of collision voltage.** Three satellite peaks represent the lipid bound states arising through the dissociation of the monomer. The naked monomer is highlighted in blue, while the three satellite peaks assigned to one phospholipid (PL), one CDL and two PL bound species (red, green and yellow respectively). Under higher energy, only the CDL bound species remains, discounting the mathematical possibility of two PL bound species. Inset shows the isolated 23+charge state of the lipid bound dimer. Presence of bound CDL at higher energy, over PL indicates a higher binding energy of CDL over the latter, plausibly owing to greater ionic and hydrophobic interactions.



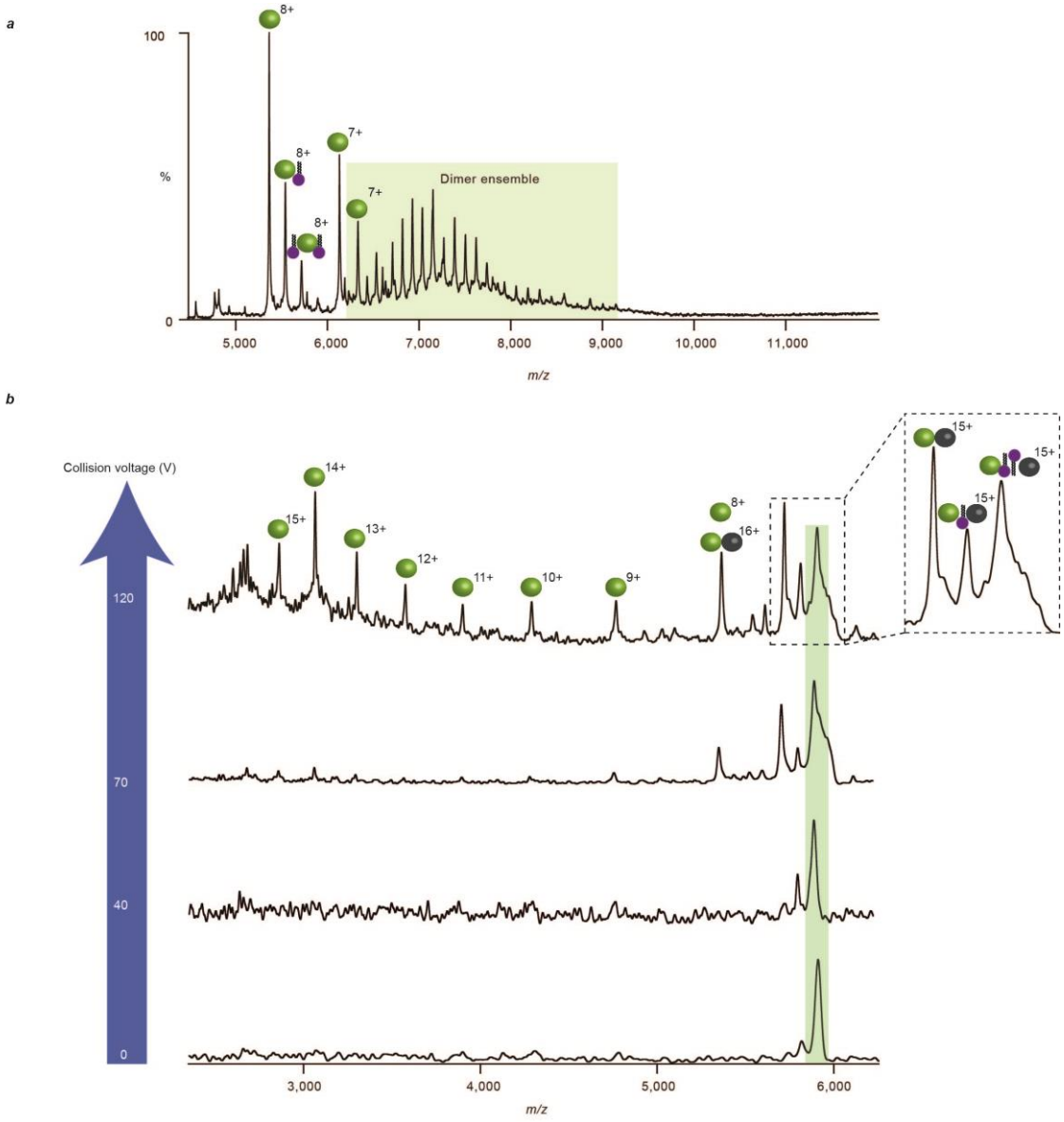

**Extended Data Figure 5 | Site directed mutagenesis of selected residues at the LeuT dimer interface, resulting mass spectra and MD simulations. a,** The LeuT dimer interface, showing π-stacking interactions (yellow dotted lines, distances labelled in red) between aromatic residues (purple). When residues F488 and Y489 (orange arrows) are mutated to alanine the π-stacking interactions are abolished and LeuT cannot dimerize. **b,** Mass spectrum of LeuT F488AY489A, liberated from OG micelles, reveals monomeric LeuT (green spheres). **c,** Mass spectrum of LeuT expressed in a CDL-deficient *E. coli* strain (BKT22), liberated from OG micelles, reveals monomeric LeuT, implying that CDL is required for LeuT dimerisation. **d,** Mass spectrum of LeuT K376AH377A, liberated from OG micelles, reveals monomeric LeuT. Insert: MD simulations of LeuT reveal CDL phosphate groups (orange sticks) interacting closely with positively charged residues (K376, H377, R506, blue sticks) at the dimer interface. Interactions are shown as dotted lines with distances measured in Å.



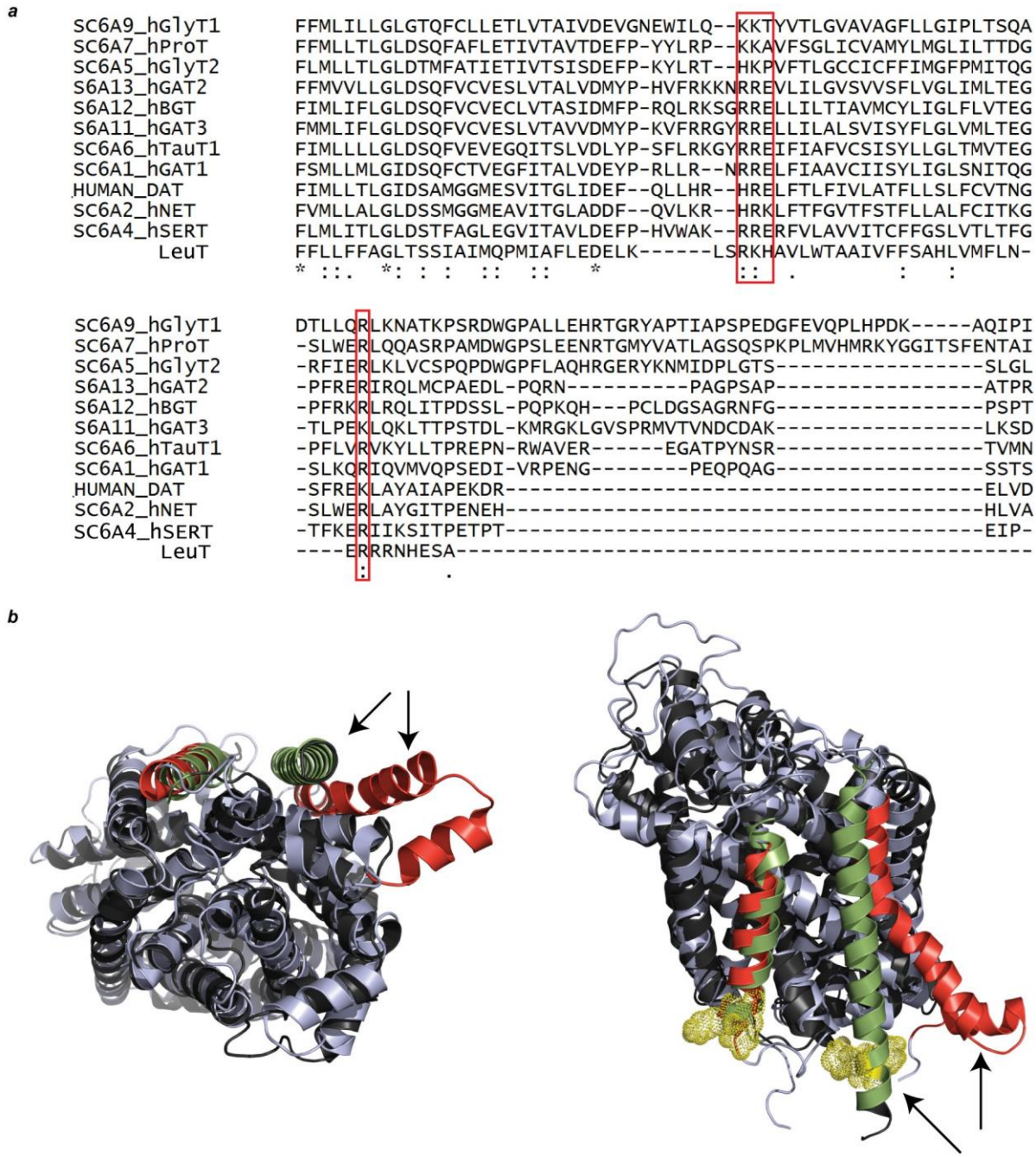

**Extended Data Figure 6 |** Particle densities from five repeats of 1 µs Coarse-grained MD simulations for CDL around (**a**) LeuT. The surface densities represent the most occupied positions from the simulations of the phosphate (orange), glycerol (red) and alkyl tails (purple) particles of CDL. As the figures show the proposed binding sites at the interface are the only places where CDL shows considerable population density. Comparative of particle densities of (**b**) CDL, (**c**) PG and (**d**)PE at the LeuT dimeric interface, summed over the simulations show no/minimal densities of PG and PE at the CDL binding site. (**a**)-(**d**) together shows that the proposed binding sites of CDL at the interface are sites of specific bindings. **e**, Dimeric structure of LeuT with modelled APT (aminopentanetetrol, aminophospholipids) classes of lipid present in the *A. Aeolicus* [52]. The lipid was drawn in ChemDraw and subsequently modelled by superimposing it on the CDL to the CDL bound dimeric structure. The favorable van der Waals' distances show that it is capable of bridging the dimeric entity, through the same sets of residues that were found to be critical towards CDL binding, in an endogenous environment lacking CDLs. **f**, Particle densities from five repeats of 1 µs Coarse-grained MD simulations for CDL around NhaA binding of two CDLs at the dimer interface. **g**,




**Extended Data Figure 7 | Mass spectra of His-tagged and unmodified SemiSWEET and identification of endogenous and exogenous lipid binding a,** Mass spectrum of unmodified SemiSWEET, liberated from tetraethyleneglycolmonooctyl ether (C8E4) micelles, reveals SemiSWEET monomers and dimers (black spheres). **b,** Mass spectrum of deca-His tagged SemiSWEET, liberated from C8E4 micelles, reveals SemiSWEET monomers and dimers (green spheres). **c,** High energy MS/MS of unmodified SemiSWEET, liberated from dodecylmaltoside (DDM) micelles, allows isolation of the 6+ charge state (black spheres) of the SemiSWEET monomer (black spheres) bound to endogenous lipids. Fragmentation of the lipid-bound species leads to loss of either cardiolipin (1470 ± 26 Da, purple head-groups), 1 or 2 neutral phospholipids (each 756 ± 22 Da, blue head-groups), or a positively charged phospholipid. Trap collision voltages shown in white inside blue arrow. **d,** Mass spectrum of deca-His SemiSWEET, liberated from C8E4 micelles and incubated with phosphatidylglycerol (PG, blue head-groups). PG binds to both monomers and dimers (dotted boxes highlight lipid-bound peaks) without substantial preference.




**Extended Data Figure 8 | Mass spectrum and high-energy MS/MS of NhaA at a range of collision voltages a,** Mass spectrum of NhaA, liberated from C8E4 micelles, reveals NhaA monomers (green spheres) bound to CDL (purple head-groups) and an ensemble of NhaA dimer species in different lipidation states (highlighted in green). **b,** MS/MS of the 15+ charge state (green) of the NhaA dimer (green/black spheres) bound to 2 CDL liberated from C8E4 micelles. Increasing collision voltage applied to the 2 x CDL-bound species leads to: loss of 1 CDL to form NhaA dimers bound to 1 CDL (40 V); loss of 2 CDL to form delipidated NhaA dimers, with concomitant generation of NhaA monomers (70 V) and further dissociation of NhaA dimers into monomers (120 V). Trap collision voltages (white inside blue arrow).




**Extended Data Figure 9 | Sequence and structure alignment of LeuT with other eukaryotic bioamine transporters (BATs)**. **a,** The basic residues of LeuT that are involved in lipid binding (red box) are conserved across the BATs. **b**, Two views of the superimposed structures of LeuT (PDB ID 2A65, black) and SERT (PDB ID 5I6Z, light blue) show the differences in the dimer interface. Dimer interface helices are highlighted with arrows and coloured (LeuT green, SERT red); basic residues responsible for lipid binding in LeuT (yellow mesh). One of the interface helices in SERT swings away from the interface, negating the possibility of lipid-induced oligomerisation analogous to that proposed for LeuT.