Membrane Protein Structure, Function and Dynamics: A Perspective from Experiments and Theory

Zoe Cournia¹, Toby W. Allen², Ioan Andricioaei³, Bruno Antonny⁴, Daniel Baum⁵, Grace Brannigan⁶, Nicolae-Viorel Buchete⁷, Jason T. Deckman⁸, Lucie Delemotte⁹, Coral del Val¹⁰, Ran Friedman¹¹, Paraskevi Gkeka¹, Hans-Christian Hege⁵, Jérôme Hénin¹², Marina A. Kasimova¹³,¹⁴, Antonios Kolocouris¹⁵, Michael L. Klein⁹, Syma Khalid¹⁶, M. Joanne Lemieux¹⁷, Norbert Lindow⁵, Mahua Roy⁸, Jana Selent¹⁸, Mounir Tarek¹³,¹⁹, Florentina Tofoleanu⁷, Stefano Vanni⁴, Sinisa Urban²⁰, David J. Wales²¹, Jeremy C. Smith²²,¹ and Ana-Nicoleta Bondar²³,¹

¹Biomedical Research Foundation, Academy of Athens, 4 Soranou Ephessiou, 11527, Athens, Greece ²School of Applied Sciences & Health Innovations Research Institute, RMIT University, GPO Box 2476, Melbourne, Vic, 3001, Australia; and Department of Chemistry, University of California, Davis, Davis, CA 95616, USA ³Department of Chemistry, University of California, Irvine, CA 92697 ⁴Institut de Pharmacologie Moléculaire et Cellulaire, Université de Nice Sophia-Antipolis and Centre National de la Recherche Scientifique, UMR 7275, 06560 Valbonne, France ⁵Department of Visualization and Data Analysis, Zuse Institute Berlin, Takustrasse 7, D-14195 Berlin, Germany ⁶Center for Computational and Integrative Biology and Department of Physics, Rutgers University-Camden, Camden, NJ, USA ⁷School of Physics and Complex and Adaptive Systems Laboratory, University College Dublin, Belfield, Dublin 4, Ireland ⁸Department of Chemistry, University of California, Irvine ⁹Institute of Computational and Molecular Science, Temple University, Philadelphia, Pennsylvania 19122, United States ¹⁰Department of Artificial Intelligence, University of Granada, E-18071 Granada, Spain ¹¹Linnaeus University, Department of Chemistry and Biomedical Sciences & Centre for Biomaterials Chemistry, 391 82 Kalmar, Sweden ¹²Laboratoire de Biochimie Théorique, IBPC and CNRS, Paris, France ¹³Université de Lorraine, SRSMC, UMR 7565, Vandoeuvre-lès-Nancy, F-54500, France ¹⁴Lomonosov Moscow State University, Moscow, 119991, Russian Federation ¹⁵Faculty of Pharmacy, Department of Pharmaceutical Chemistry, University of Athens, Panepistimioupolis-Zografou, 15771 Athens, Greece ¹⁶Department of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK ¹⁷Department of Biochemistry, Faculty of Medicine & Dentistry, Membrane Protein Disease Research Group, and Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada, T6G 2H7 ¹⁸Research Programme on Biomedical Informatics (GRIB), Department of Experimental and Health Sciences, Universitat Pompeu Fabra, IMIM (Hospital del Mar Medical Research Institute), Dr. Aiguader 88, E-08003 Barcelona, Spain ¹⁹CNRS, SRSMC, UMR 7565, Vandoeuvre-lès-Nancy, F-54500, France ²⁰Johns Hopkins University School of Medicine, Howard Hughes Medical Institute, Department of Molecular Biology & Genetics, 725 N.

¹Corresponding authors: Ana-Nicoleta Bondar, nbondar@zedat.fu-berlin.de, Tel: +49 30 838 53583, Fax: +49 30 838 56510; Zoe Cournia, zcournia@bioacademy.gr, Tel: +30 210 6597195, Fax: +30 210 6597545; Jeremy C. Smith, smithjc@ornl.gov, Tel: +1 865 574 9635, 1 865 274 1466, Fax: +1 865 576 7651.
Abstract

Membrane proteins mediate processes that are fundamental for the flourishing of biological cells. Membrane-embedded transporters move ions and larger solutes across membranes, receptors mediate communication between the cell and its environment and membrane-embedded enzymes catalyze chemical reactions. Understanding these mechanisms of action requires knowledge of how the proteins couple to their fluid, hydrated lipid membrane environment. We present here current studies in computational and experimental membrane protein biophysics, and show how they address outstanding challenges in understanding the complex environmental effects on the structure, function and dynamics of membrane proteins.

Keywords
membrane proteins; lipids; protein structure; protein function; protein dynamics; membrane-mediated interactions

Introduction

The lipid membranes surrounding cells and their various compartments host proteins that perform functions essential for both cell physiology and disease progression. Recent evidence indicates that lipid membranes can largely influence the local structure, dynamics, and even the activity of a membrane protein (Bondar et al., 2010; Bondar, del Val & White, 2009; Jardon-Valadez, Bondar & Tobias, 2010; Kalvodova et al., 2005; Urban & Wolfe, 2005). This discovery highlights the need to understand how membrane proteins operate in their physiological lipid environment, which can be very complex itself. For example, the outer membranes of Gram-negative bacteria are asymmetric in nature; the outer leaflet is composed of lipopolysaccharide, while the inner leaflet incorporates a mixture of approximately 25 phospholipid types.

The present perspective article arose from the workshop “Coupling between protein, water, and lipid dynamics in complex biological systems: Theory and Experiments”, organized within the framework of the Centre Européen de Calcul Atomique et Molecular (CECAM) in September 2013. The workshop aimed to place the most recent developments of computer simulation techniques into a biological context to identify future challenges for membrane protein computations, and to find unifying concepts in protein/lipid/water coupling. We review here selected aspects of membrane protein biophysics from both computational and experimental viewpoints. In what follows, we first introduce briefly the biomolecules that are specifically discussed in this perspective article, then provide details on the methodologies used to study them and finally discuss results.
Membrane proteins are recognized and inserted into the lipid bilayer by exquisite cellular machineries, such as the Sec protein translocase. A central component of the Sec protein translocase is the SecY/Sec61 protein translocon, which is thought to open a lateral gate and release transmembrane protein segments (du Plessis et al., 2009; Plath et al., 2004; Van den Berg et al., 2004). Opening of a lateral helical gate toward the lipid membrane is also observed in a different membrane protein, the GlpG rhomboid protease; in the case of GlpG, opening of a lateral gate allows docking of a transmembrane substrate (Baker et al., 2007).

Ion channels are fundamental molecular components of signaling in our nervous system, and have been the subject of major advances in structural determination over the last decade (see, for example (Brohawn, del Marmol & MacKinnon, 2012; Chen, Durr & Gouaux, 2014; Cuello et al., 2010; Gonzales, Kawate & Gouaux, 2009; Karakas & Furukawa, 2014; Kato et al., 2012; Lenaæus et al., 2014; Payandeh et al., 2012). The perspective of the community on the role played by lipids in channel modulation has recently started to shift: whereas earlier work only considered the membrane as an adaptable matrix for protein functioning, recent data suggest that lipid molecules play fundamental structural and functional roles in ion transport. For example, direct interactions of ligand-gated ion channels with cholesterol, which play a functional role, were observed (Hénin et al., 2014). Another challenging area in studying ion channel is how highly charged helical segments can transverse the cell membrane. For example, the S4 transmembrane helix of the voltage sensor domain (VSD) or channelrhodopsins contain highly charged helical segments, which, however, can be incorporated into the lipid membrane (Hessa, White & von Heijne, 2005) (Del Val et al., 2014), highlighting the need to understand how membrane protein segments partition into the lipid membrane.

Membrane proteins account for about two thirds of known druggable targets in the cell and about 50% of all known small molecule drugs bind to membrane proteins (Lappano & Maggiolini, 2011; Tautermann, 2014). In this regard, no discussion of biological membranes would be worth its salt without mentioning G protein-coupled receptors (GPCRs) and proteins related to amyloid diseases. For both GPCRs and amyloidogenic peptides, interactions with lipids are essential. GPCRs are sensitive to the hosting lipid environment (Brown, 1994; Goddard et al., 2013; Hille et al., 2014; Oates et al., 2012; Oates & Watts, 2011), and interactions between oligomers of amyloidogenic peptides and membranes or lipids appear central to the cellular toxicity of amyloid proteins (Tofoleanu & Buchete, 2012a; Walsh et al., 2002; Walsh & Selkoe, 2007).

Given the advances in computational methodologies and computer power, theoretical approaches are likely to become increasingly important in the study of membrane proteins and their reactions. Studying the potential energy landscape provides both conceptual and computational tools for understanding a wide range of observable properties in membrane protein science. In particular, we can exploit stationary points (minima and transition states) for structure prediction and analysis of global thermodynamic and kinetic properties. Upon passage through membrane pores, peptides undergo conformational transitions and sample intermediates that block the transmembrane current that would otherwise flow in an open pore under a potential drop. We consider here how these intermediate states can be considered “jammed” states, similar to the dynamical arrest of macroscopic granular matter.
and in macroscopic glasses. Finally, scaling up to mesoscale systems via coarse graining, coupled with three-dimensional membrane-coupled systems-level modeling, bioinformatics, and appropriate visualization techniques up to the cellular level will be required. New computational technologies will form the basis of our future understanding of integrated membrane structure and function.

MEMBRANE PROTEIN ASSEMBLY, INSERTION AND LIPID INTERACTIONS

Physical partitioning

Physical membrane protein partitioning properties directly determine membrane protein folding, stability, and function, and their understanding is vital for rational design of membrane-active peptides. However, whereas in the 1970s and 1980s the folding of proteins in membranes was considered in the context of the physical aqueous-membrane partitioning problem, it is now known that nascent transmembrane (TM) polypeptide segments are recognized and inserted into the lipid bilayer by cellular machineries such as the Sec protein translocase. The protein translocon (Sec61 in eukayotes, SecY in prokaryotes) is an essential component of the Sec protein secretion machinery in all organisms (for reviews see, e.g., Driessen & Nouwen, 2008; Rapoport, 2007; White & von Heijne, 2008). The translocon mediates the insertion of membrane proteins into the lipid membrane according to recognition rules that correlate strongly with physical hydrophobicity scales that describe the free energy of insertion of TM helices from water (Wimley & White, 1996). However, the exact relationship between the physical and biological scales is unknown, because solubility problems limit our ability to measure the direct partitioning of hydrophobic peptides across lipid membranes experimentally.

Experimentally, sufficiently hydrophobic peptides tend to aggregate, while atomistic computer simulations at physiological temperatures cannot yet reach the long time scales required to capture partitioning. Elevating temperatures to accelerate simulation of the dynamics has been avoided, as this was thought to lead to rapid denaturing. However, it has been shown that model TM peptides (WALP) are exceptionally thermostable (Ulmschneider et al., 2010). At these temperatures, sampling is ~50–500 times faster than at room temperatures (Ulmschneider et al., 2011), sufficient to directly simulate spontaneous partitioning at atomic resolution. Elevated temperature simulation ensembles further allow the direct calculation of the insertion kinetics. High-temperature simulations of experimentally validated thermostable systems suggest a new avenue for systematic exploration of peptide partitioning properties.

Using the above high-temperature methodology, in recent microsecond molecular dynamics (MD) simulations monomeric polyleucine segments of different lengths were allowed to partition spontaneously into and out of lipid bilayers (Ulmschneider et al., 2011). This approach directly revealed all states populated at equilibrium and allowed the free energy of insertion to be directly obtained from the relative occupancy of these states. The partitioning free energy of the polyleucine segments thus determined was found to correlate strongly with values from translocon experiments. Arguably, the most interesting result of this study was the close correlation between the spontaneous surface-to-bilayer partitioning equilibrium and the translocon-mediated insertion scale. However, the results show that ~2
fewer leucines than predicted by the translocon scale are required for the peptides to insert on their own. The probable explanation for this is that two different partitioning processes are involved. In contrast to partitioning between bilayer interface and hydrocarbon core, the translocon crystal structure suggests that hydrophobic peptides are released laterally from the protein-conducting channel into the membrane. This means that the experiments measure the partitioning of peptides between translocon channel and bilayer, rather than between water-soluble and TM configurations. This is consistent with a translocon-to-bilayer equilibrium being at the heart of the translocon-mediated insertion probability. The close correlation of the simulation results with translocon experiment suggests that the surface-bound helical state of the peptides might be located in a region of hydrophobicity similar to that of the internal translocon pore. Thus, the equilibria of spontaneous partitioning and translocon-mediated insertion are likely independent, with no thermodynamic cycle connecting the two insertion paths. Both differ highly from the much larger free energy changes involved in non-equilibrium water-to-bilayer partitioning.

**Dynamics of the translocon**

The functioning of the translocon is associated with significant rearrangements of the protein, and crystal structures provide valuable snapshots of the translocon in different conformational states (Egea & Stroud, 2010; Tsukazaki et al., 2008; Van den Berg et al., 2004; Zimmer, Nam & Rapoport, 2008). Particularly important is the geometry of helices 2 and 7 of the translocon, because these two helices are thought to function as gates that open towards the lipid membrane to release hydrophobic peptides (du Plessis et al., 2009; Plath et al., 2004; Van den Berg et al., 2004). In the closed state of the translocon, helices 2 and 7 hydrogen bond to each other and to helix 3, which hydrogen bonds to four other segments of the translocon (Bondar et al., 2010). The translocon has numerous hydrogen bonds that interconnect its various structural elements, and the gate helices 2 and 7 are part of a hydrogen-bonded network that stabilizes the closed state of the translocon (Bondar et al., 2010).

The dynamics of the central hydrogen-bonded cluster of helices 2, 3 and 7 has some remarkable features that could not have been anticipated from the crystal structure: of the five hydrogen bonds of the cluster, three have a multi-state character, breaking and reforming on the picosecond-nanosecond timescale (Bondar & White, 2012). Importantly, in crystal structures that glimpse into open states of the translocon (Egea & Stroud, 2010; Tsukazaki et al., 2008; Zimmer et al., 2008) groups of the central hydrogen-bonded cluster have rearranged such that helix 7 appears to have broken away from hydrogen bonding with helices 2 and 3 (Bondar & White, 2012). The comparison of the hydrogen-bond analysis and of the crystal structure then suggests that hydrogen bonds with complex conformational modes are implicated in the control of the protein conformational state (Bondar & White, 2012).

**Targeting cellular organelles by amphipathic helices**

One crucial membrane property that has been suggested to favour translocon-assisted membrane protein folding is the extremely low cholesterol concentration of the endoplasmic reticulum (ER) of eukaryotic cells (Bretscher & Munro, 1993). The peculiar lipid
composition of the ER, abundant in conical shaped lipids, such as diacylglycerols or those containing monounsaturated fatty acids (Bigay & Antonny, 2012), appears thus well adapted to decrease the energetic barriers for protein insertion thanks to its intrinsic plasticity. In contrast, late secretory pathway membranes (plasma membrane, late endosomes), are much stiffer, because of their enrichment in cholesterol and sphingolipids.

Remarkably, not only proteins that partition to the interior of the bilayer, but also those that partition to its surface, are sensitive to the lipid composition of the various membrane organelles, and can take advantage of the remarkable plasticity of early secretory pathway (ER, Golgi apparatus) membranes (Bigay & Antonny, 2012). The most common protein motifs able to recognize these specific membrane properties are amphipathic helices, short protein motifs that are usually unfolded in solution and that fold into an alpha helix upon binding with the membrane surface.

To accurately characterize how amphipathic helices can sense membrane properties that depend on lipid composition, all-atom MD simulations, in combination with a topographical method able to scan the bilayer surface at the lipid-water interface, have been used to show that bilayers containing a high percentage of conical lipids are also enriched in defects in lipid packing (Vamparys et al., 2013) and that these defects resembles those generated by curvature.

These defects can be recognized by amphipathic helices, such as the membrane-curvature sensing Amhipathic Lipid Packing Sensor (ALPS) motif: all-atom MD simulations show that the ALPS motif's partition into the bilayer through repetitive insertions of hydrophobic residues into pre-existing lipid packing defects and the partitioning is dramatically slowed down in the presence of tightly packed lipid bilayers (Vanni et al., 2013). These studies suggest that computer simulations are a powerful methodology to investigate, with atomic-level resolution, how membrane properties change as a function of lipid composition and how these properties can be sensed by protein motifs (Vanni et al., 2014).

**PROTEOLYSIS IN THE LIPID MEMBRANE PLANE: RHOMBOID PROTEASES**

The protein translocon discussed above uses a lateral gate to release transmembrane helices into the lipid bilayer. A lateral gate to the membrane is also used by rhomboid proteases to dock their transmembrane substrates. The tight communication between the lipids and the protease is highlighted by the observation from experiments that the proteolytic activity of the rhomboid proteases is drastically affected by the composition of the lipid bilayer – for example, the *Escherichia coli* GlpG rhomboid protease is active in phosphatidylethanolamine (PE) lipids, but not in phosphatidylcholine (PC) (Urban & Wolfe, 2005). This marked dependency of the proteolytic activity on lipids may be explained in part by the different hydrogen-bonding properties of different lipid head groups: MD simulations of GlpG in POPE vs. POPC lipid bilayers indicated that the lipid membrane composition affects not only the local hydrogen bonding between lipid head groups and protein surface amino acid residues, but it also influences hydrogen bonding between specific protein groups, and the orientation of the protease relative to the lipid membrane (Bondar et al., 2009). A key open question is how the transmembrane substrate is
docked and cleaved by the protease. This question is addressed below from the viewpoint of X-ray crystallography, biochemistry and cell biology.

**Rhombooids have low affinity for their transmembrane substrates**

While rhomboid enzymes are the best understood intramembrane proteases structurally (Baker & Urban, 2012), the mechanistic details of their catalytic activity inside the membrane are still unclear. A new reconstitution system has been developed to study the kinetics of rhomboid proteolysis inside the membrane (Dickey et al., 2013). In this system prior hurdles were systematically overcome by substrate turnover and allowing the reaction to be initiated synchronously. Fluorescence of the released product provided a sensitive and quantitative means to measure initial rates within the first 12 minutes in real time. This is a new achievement, because current intramembrane protease assays are performed in detergent and over the course of hours. The resulting steady-state kinetics revealed rhomboid proteases have very low affinity for their substrates in the membrane, which was confirmed by measuring the $K_d$ directly and validated by following cleavage in living cells. The reaction is also slow. These two features revealed a surprising similarity to DNA repair enzymes that monitor substrate dynamics. In a similar way, rhomboid proteases may target their substrates by identifying transmembrane segments that are able to unfold inside its active site. This property might also explain why no enzyme-substrate complex structure has been achieved despite nearly a decade of effort. Consequently, excellent opportunities exist for MD simulations to explore how substrates behave as they approach the protease, enter the active site, and are ultimately cleaved.

**Dimer model for the full-length E. coli rhomboid intramembrane protease**

To date the only crystal structures of rhomboid available are determined from *Haemophilus influenzae* (hiGlpG) (Lemieux et al., 2007), and the core *E. coli* GlpG membrane domain (Ben-Shem, Fass & Bibi, 2007; Brooks et al., 2011; Dickey et al., 2013; Lemieux et al., 2007; Vinothkumar, 2011; Vinothkumar et al., 2010; Wang & Ha, 2007; Wang et al., 2007; Wang, Zhang & Ha, 2006; Wu et al., 2006; Xue et al., 2012; Xue & Ha, 2012). Both are found in the same phylogenetic branch (Urban & Dickey, 2011), and not surprisingly, have similar three dimensional structures (Lemieux et al., 2007). Both consist of six TM domains that assemble to form an $\alpha$-helical bundle. A catalytic Ser-His dyad buried 10–12 Å deep within the bilayer is located within this bundle. Between helices 1 and 2 is a loop, containing some invariant residues, that partially extends into the lipid bilayer.

The E. coli rhomboid GlpG represents a classical prokaryotic rhomboid that has a soluble N-terminal cytoplasmic domain in addition to the six transmembrane segments. The physiological function for GlpG is unknown, and furthermore the role for this cytoplasmic domain in GlpG is unclear. No structural information exists for the full-length *E. coli* rhomboid. We have determined that the presence of the cytoplasmic domain does not affect the catalytic rate or specificity of GlpG (Lazareno-Saez et al., 2013). The crystal structure of the E. coli rhomboid cytoplasmic domain reveals a domain swapped dimer. This structure may be physiologically relevant as it is dimeric not only in solution but when in contact with the membrane domain (Arutyunova et al., 2014). The membrane domain of prokaryotic rhomboid is also dimeric independent of this cytoplasmic domain (Sampathkumar et al.,
We hypothesize that dimerization is a key step for regulation of intramembrane proteolysis.

Rhomboids are synthesized as active enzymes and are known to cleave substrates with minimal specificity. A mechanism must exist to achieve specificity for this enzyme. Movements in loop 5 and helix 5 around the buried active site are thought to provide access to the catalytic machinery (Baker et al., 2007; Brooks et al., 2011; Wang & Ha, 2007), and may provide a means for substrate specificity. Dimerization of the cytoplasmic domain would stabilize the dimer of the catalytic membrane domain, and may provide the structural changes necessary to promote conformational changes in the active site. Here we present a model where dimerization occurs at both the membrane domain and the cytoplasmic domain (Figure 1). The two domains are connected by a flexible linker. The dimerization interface of the membrane domain, here modeled to occur between the conserved loop 1, needs to be confirmed both experimentally and with molecular dynamics.

ION CHANNELS

Studying ion channel function using fully atomistic molecular dynamics simulations

Ion channels catalyze rapid and selective transport of salt ions, such as Na⁺, K⁺, Cl⁻ and Ca²⁺, across cell membranes to enable electrical and chemical activity in the body. They are associated with many neurological, muscular and cardiac disorders, which establishes them as important pharmacological targets (Clare, 2010).

The solution of the first high-resolution crystal structure in 1998 for the bacterial K⁺ channel, KcsA (Doyle et al., 1998) (see Figure 2a), opened the door to molecular-level insight into ion channel conduction. In the extracellular-facing pore segment, called the selectivity filter, the protein backbone forms a narrow, ion-sized pore, presenting a sequence of carbonyl cages to coordinate almost dehydrated K⁺ ions, and is responsible for ion discrimination. It has long been thought that this filter is inherently better at stabilizing K⁺ over Na⁺ ions (Neyton & Miller, 1988), suggesting selective conduction via a ‘thermodynamic’ mechanism of selective binding. Based on the KcsA structure, it was argued that this preference is due to a ‘snug fit’ of the K⁺ ion inside the selectivity filter, which is too big for a Na⁺ ion (Doyle et al., 1998). The high-resolution structure in itself does not, however, tell the full story. Proteins are often pictured as rigid entities, but they are in fact highly flexible, with even the most rigid protein backbones fluctuating (and able to deform) on the Å-scale, critical for ion channel conduction (Allen, Andersen & Roux, 2004; Noskov, Berneche & Roux, 2004). For the protein to prohibit ~0.4 Å adjustment of carbonyl groups to account for K⁺-Na⁺ size differences, despite the attractive interactions between ion and carbonyl groups, it would have to be as stiff as diamond (Allen et al., 2004). Clearly mechanisms of ion channel selectivity must emerge from models that incorporate the thermal fluctuations of the protein, from picosecond vibrations to structural isomerizations that span up towards the microsecond scale.

Much progress has been made with MD simulation, based on this structure, to understand the mechanisms by which K⁺ channels achieve selective conduction (Andersen, 2011). Although there has been considerable attention given to the preference of carbonyl cage
geometries for K$^+$ ions, focusing on the field strength created by the ligands, the strain energies associated with ligand-ligand repulsion, and the role of thermal fluctuations (e.g. (Noskov et al., 2004), it has been suggested that the selective binding picture is oversimplified. Recent calculations have shown that this preference is smaller than previously thought, simply because the cage of carbonyl groups is the K$^+$ binding site, but that the selectivity filter prefers to house the smaller Na$^+$ ions at adjacent plane of carbonyl sites (such that focusing only on cage sites inherently biases calculations to favor K$^+$).

Instead, on average, there is only a small overall preference for K$^+$ in the middle of the selectivity filter, and surprisingly, a preference for Na$^+$ at the entrances of that region (Kim & Allen, 2011). It has been shown that selective permeation can instead arise from differences in the barriers for the concerted motions of multiple ions through the channel (Figure 2b and c; (Nimigean & Allen, 2011; Thompson et al., 2009). This ‘kinetic view’ of selective exclusion (Bezanilla & Armstrong, 1972) implies that one must elucidate the full free energy landscape governing the movements of all of the ions in the pore.

With this in mind, and armed with a new crystal structure for a bacterial Na$^+$ channel (Payandeh et al., 2011), research has rapidly moved towards uncovering the contrasting mechanisms of selectivity in Na$^+$ channels. Simulations have revealed a surprising degree of protein flexibility for this channel, especially in the important selectivity filter, and have shown that is essential to invest in sufficient sampling times, if one wants to obtain quantitative measures of function (Boiteux, Vorobyov & Allen, 2014a; Chakrabarti et al., 2013). Use of purpose-built computational architectures to simulate large membrane protein systems, for multiple microseconds, up to milliseconds, e.g. (Jensen et al., 2012), have indeed revealed protein conformational changes relevant to activity, e.g. (Boiteux et al., 2014a; Boiteux et al., 2014b).

Despite the ability to carry out long enough simulations to begin to observe function, we still seek quantitative and reproducible measures to understand mechanisms and compare to experiments. An obvious route is to carry out equilibrium sampling to produce a free energy surface, via its connection to the probability distribution, or the potential of mean force (PMF) along the reaction coordinate. The calculation is typically done using Umbrella Sampling (Roux, 1995), where a sequence of independent simulations, distributed along the coordinate, biased to ensure local sampling within every region of that coordinate, can then be recombined after simulation in an optimal way, such as via the Weighted Histogram Analysis Method, WHAM (Kumar et al., 1992). Interestingly, however, with the use of new computer architectures that allow multi-microsecond unbiased ion channel simulations, complete sampling of the multipleion mechanisms of conduction have been obtained without Umbrella Sampling (Boiteux et al., 2014a). Those same long simulations have also revealed conformational changes involving trans-membrane helix bending and gate collapses that relate to slow inactivation in Na$_v$ channels, despite these processes typically occurring on much longer timescales. They have also uncovered the interplay between the protein and the lipid membrane via fenestrations in the pore domain. Importantly, they have revealed the distribution of lipids and lipophilic drugs around the channel pore domain and fenestrations, as well as key details of the interactions responsible for drug binding and pathways (Boiteux et al., 2014a; Boiteux et al., 2014b). The field of biomolecular simulation
is therefore beginning to uncover the quantitative mechanisms of ion channel function and modulation, enabling future developments in drug discovery.

**Voltage Sensor Domain (VSD) Activation**

The VSD is a transmembrane molecular device used by cells to transduce an electrical signal across the plasma membrane. Composed of a solvated four transmembrane helix bundle (Figure 3), this device can be coupled to an ion channel pore (voltage-gated cation channel – VGCC), to a phosphatase (voltage-sensing phosphatase-VSP) or can be a self-standing selective channel (proton voltage-gated channel – Hv). Amongst the four helices, S4 stands out with its large number (4 to 7) of positive charges (Arg and Lys) spread along its lumen facing side: as such, it constitutes the primary voltage-sensor (Palovcak et al., 2014). These charges are stabilized in a transmembrane position through their interaction, not only with the negative charges of S1–S3, but also with the negatively charged phosphate groups of the lipid membrane, creating a favorable solvated environment. When a hyperpolarizing electric field is applied, S4 is dragged downwards in a ratchet-like movement, as the salt bridge pairings between the S4 positive charges and their counterparts rearrange in a step-wise manner. Because of the hourglass volume accessible to water within the VSD, the electric field is focused over a narrow hydrophobic region (sometimes called the hydrophobic plug), made up by several hydrophobic residues, including a conserved PHE residue on S2 and ILE on S1. In VGCCs, an intracellular interfacial helix (the S4–S5 linker) couples the VSD to the pore helices (S5–S6) helping therefore to convert electrical energy of the membrane potential into the mechanical energy needed to open and close the S6 gate.

Kv1.2 is the first mammalian VGCC of which the open state crystal structure was revealed. Using an MD based protocol to trigger deactivation (closing) of the channel in a model POPC membrane the details of the S4 sensor downwards movement were revealed (Delemotte, Klein & Tarek, 2012; Delemotte et al., 2011; Tarek & Delemotte, 2013). Interestingly, several of the S4 positive charges are stabilized by their interaction with the lipid head groups, not only in the activated state, where the two top charges are bound to top leaflet phosphate groups, but also in the intermediate and resting states: in the middle intermediate state, the outermost charge is in contact with the top PO₄-group and the innermost charge in contact with the inner leaflet PO₄-groups and in the resting state, the innermost three charges interact with the inner leaflet phosphate group.

**VGCC modulation by lipids**

The perspective on how lipids modulate ion channels has recently started to change (Kasimova et al., 2014). Indeed, earlier work only considered the membrane as an adaptable matrix for protein functioning. Recent data, however, revealed that lipid molecules have fundamental structural and functional roles (Sands & Sansom, 2007; Schmidt, Jiang & MacKinnon, 2006; Treptow & Tarek, 2006). For example, the activation of the VSD is modified by the alteration of sphingomyelin head group of the outer membrane leaflet by sphingomyelase (Combs et al., 2013; Ramu, Xu & Lu, 2006; Xu, Ramu & Lu, 2008). Such a mechanism is taken advantage of by the spider *Loxoceles reclusa*, which carries this enzyme in its venom and uses it to paralyze its prey. Another well-studied example is the
specific role of PIP2, a signaling lipid of the membrane inner leaflet, which modulates several processes in the working cycle of VGCCs (Kruse, Hammond & Hille, 2012; Logothetis et al., 2010; Suh & Hille, 2008).

In principle, two hypotheses can be proposed to rationalize the macroscopic VGCC response at a molecular level: altering the lipid head groups may 1) modify the global properties of the membrane, and more specifically the electrostatic environment or 2) change the local interaction pattern between protein residues and their environment. Insights gained into this question are discussed using MD simulations.

1. Sphingomyelin alteration—Sphingomyelin is an abundant lipid (forms up to 15%) of the outer layer of mammalian cell membranes (Combs et al., 2013). Recently, Lu and co-workers have shown that enzymes present in the venom of Loxoceles reclusa target VGCCs, causing thereby paralysis (Combs et al., 2013; Ramu et al., 2006; Xu et al., 2008). The mode of action of these enzymes is very peculiar since they target the lipid matrix rather than the protein itself. Specifically, Smase D removes the sphingomyelin choline group leaving a negatively charged lipid (ceramide-1-phosphate) that promotes activation of the channel. Smase C, on the other hand, removes the entire sphingomyelin zwitterionic head group and leaves behind a polar molecule (ceramide). This favors the closed state of the channel.

The electrostatics of asymmetric bilayers in which the top leaflet is made of 100% of one of the three sphingomyelin derivatives (sphingomyelin itself, ceramide-1-phosphate and ceramide) and the bottom one of 100% phosphatidylcholine were studied. Interestingly, the contributions to the electrostatic potential profile along the normal to the bilayer from the individual components of the top layer (lipids, water and counterions) are system-dependent, i.e. the most polarizable elements counteract the rather static dipole potential from the lipid head groups. However, surprisingly, summing these different contributions leads to the same overall transmembrane potential of ~200 ± 50 mV. The Kv1.2 VSD (Delemotte et al., 2011; Long, Campbell & Mackinnon, 2005) in three different states (activated, intermediate and resting) in these three asymmetrical bilayers were then inserted. The gating charge, i.e. the quantity of charge transported by the VSD during activation was subsequently calculated (Delemotte et al., 2012; Treptow, Tarek & Klein, 2009). Through its energetic formulation, the gating charge is an indirect measurement of the rearrangement of the electric field acting on the protein upon external voltage application (Roux, 1997). Therefore, it can be computed for each transition from one state to another. In this particular case, the analyses have shown that the gating charges do not vary much with the environment, i.e. with the lipid composition. Taken together, these results indicate that the reshaping of the electric field is independent of the lipid composition and rule out any significant global effect of the electrostatic environment on the VSD functioning subsequent to modification of the lipid.

On the other hand, close inspection of the systems reveals lipid-dependent interactions with the S4 outermost positive residues, providing support to the local effect hypothesis. The presence of a phosphate in the sphingomyelin head group provides a stable binding site for the top two S4 arginines when the VSD is activated. In ceramide-1-phosphate, this interaction is expected to be more stabilizing due to the absence of a countercharge in the lipid head group. In ceramide, since there is no phosphate group, it was hypothesized that
the interaction between the VSD and the lipids head groups is much weaker. Relaxation of the three systems using unbiased MD simulations exceeding the 100 ns time scale under large hyperpolarizing potentials were not enough to reveal any significant destabilization of the interactions between the VSD and the lipid head groups. Free energy calculations are now under way to fully characterize the interactions involved, and link the changes in activation kinetics to experimental observables.

2. PIP2 modification—Phosphatidylinositol-(4,5)-bisphosphate (PIP2), a high negatively-charged phospholipid found in the inner membrane leaflet, has been shown to modulate the function of voltage-gated potassium channels of the Shaker-family (Abderemane-Ali et al., 2012; Rodriguez-Menchaca et al., 2012). The application of PIP2 to these channels leads both to a gain-of-function, namely an increase of the ionic current, and to a loss-of-function manifested by the right shift of G/V and I/V curves, describing respectively the gating and ionic current voltage dependencies.

Because PIP2 is a rare component of the membrane (1% of the inner leaflet of most cells) any global effect was ruled out. Rather, the focus was placed on the specific interactions between the channel and the lipids in a system in which the ring of phosphatidylcholine molecules of the inner leaflet embedding the channel was replaced by PIP2 lipids. From 100 ns MD simulations of the Kv1.2 channel embedded in such a model lipid bilayer, it was shown that PIP2 molecules interact electrostatically with positively charged residues of the channel in a state-dependent manner. In particular, PIP2 lipids were found to interact with two different sites of the channel that bear an excess amount of positive charges. The first site (I) is composed by the S4–S5 linker outermost residues (LYS312, ARG326). The other site (II) is located in the middle of the S4–S5 linker (LYS322) and is occupied by residues in a state-dependent manner. In the closed state, aside from LYS322, site II encloses the S4 lower basic residues ARG303, LYS306 and ARG309, consistent with PIP2 stabilizing the OFF position of the voltage sensor. In the open state, ARG419 of the S6 C-terminal are in close contact with the middle of the S4–S5 linker (witnessed in three subunits out of four) occupying therefore site II. This is consistent with PIP2 stabilizing the open gate conformation through a mechanism similar to the one described for Kir channels (Hansen, Tao & MacKinnon, 2011; Whorton & MacKinnon, 2011).

In the up state, all S4 basic residues are engaged in salt bridges with other residues of the VSD, or with the upper leaflet lipid head groups. Within the first ON to OFF transition, as S4 translate down, the lowest residue ARG309 comes in close contact with the lipids of the lower leaflet (Delemotte et al., 2011). Consistent with this, one expects that the presence of PIP2 in this lower leaflet would modify the kinetics associated with this early transition, i.e. leads to accelerated deactivation. Accordingly, the models seem to agree well with the electrophysiology data (Abderemane-Ali et al., 2012; Rodriguez-Menchaca et al., 2012) and provide an insight into the molecular mechanisms involved in the seemingly contradictory gain- and loss-of-function effects of intracellular PIP2.

This work is being currently extended to other channels of the VGKC family (Kv7 subfamily) that have been shown also to require PIP2 for proper function.
Drug design for ion channels

Ion channels are expressed ubiquitously in cells and are implicated in a variety of physiological processes. Thus, targeting them with small molecules that can act as drugs is particularly challenging (Bagal et al., 2013).

The use of computational methods can greatly enhance our understanding of drug-protein binding processes, provide estimates of drug binding affinities, and may ultimately contribute to the development of novel, potent protein inhibitors. Common methodologies used in estimating ligand-binding affinities include docking, continuum electrostatic methods such as the Molecular Mechanics/Poisson–Boltzmann Surface Area (MM/PBSA) and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) techniques, or Free Energy Perturbation (FEP) calculations. FEP calculations are an accurate quantitative link between experimental and computational studies (Erion et al., 2007; Lamb & Jorgensen, 1997; Reddy et al., 2007; Reddy & Erion, 2001; Zwanzig, 1954).

Relative binding free energies are calculated using the Thermodynamic Cycle Perturbation method. Using this thermodynamic cycle, the relative change in the free energy of binding between two ligands A and B is calculated through nonphysical paths connecting a reference and a terminal state (Zwanzig, 1954). Using Cycle 1 (Figure 4), the difference in binding free energy between ligand A and ligand B can be calculated, while the relative solvation free energy difference between two ligands A and B can be calculated based on Cycle 2 (Figure 4).

Calculations of relative free energies of ligand binding in membrane proteins and in particular in ion channels using the FEP methodology combined with MD simulations (FEP/MD) have been performed only for a limited number of systems and ligand alchemical transformations (Gkeka et al., 2013; Gonzalez et al., 2011; Zhao, Caplan & Noskov, 2010). We review below the usage of FEP/MD calculations to design novel drug candidates for A/M2TM.

The A/M2TM protein is a tetrameric pore of a proton channel that is activated by low pH in the viral endosome. The main class of known A/M2TM blockers is adamantane derivatives, such as amantadine and rimantadine (Figure 5, left), that can block the A/M2TM pore via a pore blocking mechanism (Cady et al., 2010; Cady et al., 2011; Chuang et al., 2009; Gu et al., 2011; Rosenberg & Casarotto, 2010; Stouffer et al., 2008). Based on this pore blocking mechanism, docking calculations and FEP/MD calculations with a restrained A/M2TM protein system were performed and resulted in poor or low correlation with experimental data ($R^2 \leq 0.20$) (Gkeka et al., 2013). This low correlation was attributed to the limited flexibility of the protein in the system, as the use of a “frozen” environment introduces a dependency of the free energies of binding on the selection of the protein structure, as discussed elsewhere (Michel, Foloppe & Essex, 2010). An improved protocol, where protein flexibility of the system was taken into account by embedding the protein inside a fully hydrated DPPC lipid bilayer, yielded an excellent correlation with the available experimental data ($R^2 = 0.85$, Figure 5, right), indicating that protein flexibility dramatically improves the accuracy of FEP calculations for this system. Comparison between the different MD trajectories indicated that protein flexibility greatly influences ligand location.
inside the pore, affecting their hydrogen bonding ability and overall interaction network. In addition, the calculated bundle tilt with respect to the membrane normal for the different ligands in the flexible protein system varied between 32° and 38° relative to the membrane plane in agreement with previous studies (Cady & Hong, 2008; Wang et al., 2001), while it was zero for the restrained system. This bundle tilt change reflects differences in tetramer flexibility. Though keeping the protein backbone rigid yields accurate results for certain binding pockets (Baggett et al., 2012), backbone flexibility appears essential in the case of A/M2TM.

In (Gkeka et al., 2013), ligand desolvation was used for the first time to provide an explanation for the binding preference of close aminoadamantine derivatives such as the 20-fold higher affinity of rimantadine 4 relative to amantadine 1 (Figure 5). It was shown that the desolvation free energy penalty is a major contributor to the A/M2TM binding affinity and should be considered for drug design purposes targeting membrane-associated systems. Decrease in the penalty of ligand desolvation facilitates its entrance into the binding site crevice and thus higher binding affinity can be obtained. However, a delicate balance is required between introducing hydrophobic elements on the ligands in order to decrease the ligand desolvation penalty and achieving optimal ligand-protein interactions (Gkeka et al., 2013).

Analysis of protein-ligand binding MD trajectories may reveal important interactions that affect the binding process. In the A/M2TM case, rimantadine 4 is implicated in favorable hydrophobic contacts due to its additional CHCH₃ bridge with residues closer to the C-terminal of the A/M2TM pore, which are missing in amantadine 1. Also, hydrogen bond analysis showed that amantadine forms on average two hydrogen bonds with neighboring water molecules, while rimantadine forms three hydrogen bonds with water molecules almost throughout the simulation time. The increased hydrogen bonding ability of 4, the additional hydrophobic interactions as well as the higher desolvation free energy compared to 1 may well explain the higher binding affinity of rimantadine to the A/M2TM channel. Overall, based on similar analysis performed on compounds 1–11 (Figure 5), it was demonstrated that a delicate balance of intermolecular interactions, hydrogen bonding, and free energy of solvation is required to achieve enhanced ligand-protein binding in membrane protein systems (Gkeka et al., 2013).

**Direct interactions of ligand-gated ion channels with cholesterol**

Ligand-gated ion channels of the nicotinic acetylcholine receptor (nAChR) family have long been known to be extremely sensitive to their lipid environment, particularly cholesterol. The first evidence for molecular interactions between the nAChR and cholesterol is a functional requirement: the receptor is not functional when purified and reconstituted in phospholipid mixtures, unless cholesterol or chemically related compounds are added (Criado, Eibl & Barrantes, 1984; Dalziel, Rollins & McNamee, 1980; Kilian et al., 1980; Ochoa, Dalziel & McNamee, 1983). Although interactions of GABA_A with cholesterol have received less attention, purification and reconstitution protocols for GABA_A receptors either include cholesterol (Bristow & Martin, 1987; Dunn et al., 1989), or yield non-functional and unstable receptors (Bristow & Martin, 1987). Subsequent experiments depleted native
membranes of cholesterol using liposomes (Lechleiter, Wells & Gruener, 1986; Leibel et al., 1987; Zabrecky & Raftery, 1985) or cyclodextrins (Baez-Pagan et al., 2008; Borroni et al., 2007; Santiago et al., 2001). An important result by Leibel et al. (Leibel et al., 1987) is that certain cholesterol molecules (about 36 molecules per receptor) remain in the system despite extensive depletion. Hence, this procedure may primarily modulate the cholesterol content of the bulk membrane. The theory that depletion/enrichment experiments access a different pool of cholesterol than reconstitution experiments is further supported by the opposing trend they yield: in depletion/enrichment experiments increased cholesterol content is associated with loss of nAChR function (Borroni et al., 2007; Santiago et al., 2001).

Functional modulation of GABA_A receptors by cholesterol has been demonstrated by Sooksawate and Simmonds (Sooksawate & Simmonds, 2001a). In this system, both enrichment and depletion of cholesterol reduce the potency of the receptor, but while the enrichment effect can be reproduced using epicholesterol, introducing epocholesterol did not reverse the depletion effect. Thus, it was suggested that the depletion effect reflects direct interactions, whereas the enrichment effect is mediated by physical properties of the membrane.

Functional studies as summarized above are complemented by biophysical results demonstrating the existence of strong association between the nAChR and a pool of proximal cholesterol. The nAChR has a higher affinity for cholesterol than other lipids, as evidenced from early lipid monolayer experiments by Popot et al. (Popot et al., 1978). Electron paramagnetic resonance (EPR) measurements by Marsh et al. (Marsh & Barrantes, 1978; Marsh, Watts & Barrantes, 1981) using spin-labeled lipid probes found that some lipids are immobilized in the presence of the receptor – in particular, the steroid androstanol exhibits hindered rotation around its long axis. High affinity of the receptor for cholesterol or its analogues was further demonstrated by EPR (Ellena, Blazing & McNamee, 1983), photolabeling (Middlemas & Raftery, 1987), and fluorescence studies (Arias et al., 1990; Barrantes et al., 2000).

Cholesterol also has significant effects on the nAChR function at the cellular level, which have been reviewed recently in the case of the murine muscle-type nAChR (Barrantes, 2007). Here we focus on molecular-level interactions and their basis in the protein structures. For a review of cholesterol modulation of membrane receptors (nAChR and G protein-coupled receptors), see refs. (Barrantes, 2004; Burger, Gimpl & Fahrenholz, 2000).

**Embedded cholesterol in a nicotinic receptor**

Medium resolution cryo-electron microscopy images of the nAChR observed within a native lipid environment (Unwin, 2005) provided the earliest atomic-level insights into pLGIC structure (Figure 6, PDB ID:2BG9). Surprisingly, the structure indicated several significantly sized gaps in protein density within the transmembrane domain. Similar gaps have not been observed in more recent crystal structures of other (primarily prokaryotic) pLGICs.

Using both automated and interactive docking, it was found that the gaps could contain cholesterol molecules embedded within the TM domain (Brannigan et al., 2008).
Furthermore, the TM domain quickly collapsed to fill the gaps during MD simulations; such collapse was correlated with the loss of critical contacts in the interface between the transmembrane and extracellular ligand binding domains. Collapse and the subsequent decoupling effect were not observed when cholesterol was bound to internal sites revealed by docking, and the structure was most stable when all gaps were filled with cholesterol (15 molecules per pentameric assembly, see Figure 6). It was therefore proposed that the gaps in the structure (PDB ID:2BG9 (Unwin, 2005)) structure may represent regions filled with unresolved cholesterol molecules.

Specific binding of cholesterol to GABA<sub>A</sub> receptors

Despite the scarcity of direct information, a significant set of indirect evidence points to the intersubunit interfaces of the GABA<sub>A</sub> transmembrane domain as an interaction site for cholesterol. GABA<sub>A</sub> receptors are positively modulated by cholesterol derivatives known as neurosteroids (Belelli & Lambert, 2005). The interface between α and β subunits in the transmembrane domain of the GABA<sub>A</sub>R, in the same region as the ivermectin site in GluCl, has been identified through mutagenesis by (Hosie et al., 2006; Hosie, Wilkins & Smart, 2007) as the likely site of activation by the positively modulating neurosteroids.

Modulation by neurosteroids depends on cholesterol concentration (Sooksawate & Simmonds, 1998; Sooksawate & Simmonds, 2001b). Furthermore, the antagonistic effect upon cholesterol enrichment is selective of cholesterol versus epicholesterol (Sooksawate & Simmonds, 2001a), and may therefore be ascribed to specific cholesterol binding, possibly competitive binding to the intersubunit sites.

GABA<sub>A</sub> receptors are also known to be potentiated by ivermectin (Dawson et al., 2000; Krusek & Zemkova, 1994), a compound which is a potentiator and agonist of the glutamate-gated chloride channel (GluCl) of the nematode C. elegans. The X-ray structure of GluCl in complex with ivermectin (PDB ID:3RHW(Hibbs & Gouaux, 2011) indicates ivermectin bound to the five subunit interfaces of the transmembrane domain.

Hénin et al (Hénin et al., 2014) built homology models of a GABA<sub>A</sub> receptor type α<sub>1</sub>β<sub>1</sub>γ<sub>2</sub> based on the crystal structure of GluCl, and investigated the possibility of cholesterol binding through three complementary approaches: alignment of cholesterol with crystallographic coordinates of ivermectin, automated docking with AutoDock, and explicit, atomistic MD simulations. Docking results confirmed the possibility of cholesterol binding to the five intersubunit clefts in a largely symmetric fashion, despite the asymmetric nature of this heteropentameric receptor (Hénin et al., 2014). The hydroxyl group of cholesterol mimics an ivermectin moiety in that it forms a hydrogen bond with a serine amino acid residue from the pore-forming helix M2. MD simulations indicated that cholesterol is stable within the cavities, but may reorient on a time scale of 100 ns or more. Two instances were observed of cholesterol unbinding, then spontaneously re-binding within the course of 200 ns trajectories, thus providing a picture of the spontaneous binding process in atomic detail (Figure 7) (Hénin et al., 2014).
G-PROTEIN COUPLED RECEPTORS

Direct and indirect lipid effects modulate GPCR functionality

GPCRs are TM proteins that recognize and respond to a variety of molecules such as hormones and neurotransmitters, thus being involved in many aspects of cell physiology. The mechanisms by which GPCRs couple to their lipid membrane environment (Sadiq et al., 2013) remains elusive but, as discussed below, it appears to involve both direct lipid-GPCR interactions and indirect effects.

Direct lipid-GPCR interactions—One of the best-studied GPCRs in terms of its membrane environment is bovine rhodopsin. On the one hand, it has been experimentally demonstrated that a high content of polyunsaturated lipid chains destabilizes the native receptor state and enhances the kinetics of the photocycle (Mitchell, Niu & Litman, 2001). On the other hand, cholesterol seems to stabilize the native state, thus slowing down its kinetics (Niu, Mitchell & Litman, 2001). As reviewed in ref. (Sadiq et al., 2013), several computational works have investigated the structural basis of lipid coupling to rhodopsin at an all-atom level. Feller et al. observed a preference for unsaturated lipid chains of 1-stearoyl-2-docosahexaenoyl-sn-glycero-3-phosphatidylcholine (SDPC) to interact with the protein, thereby excluding contacts with the saturated chains of SDPC (Feller, Gawrisch & Woolf, 2003). This finding was supported by a study of Pitman et al. (Pitman et al., 2005), who used a ternary membrane mixture of SDPC, 1-stearoyl-2-docosahexaenoyl phosphatidylethanolamine (SDPE) and cholesterol (2:2:1). Importantly, they found that the unsaturated docosahexaenoic acid (DHA) establishes contacts deep inside the protein, and hypothesized that this type of contacts could be responsible for receptor destabilization when placed into a DHA-rich environment (Grossfield, Feller & Pitman, 2006).

Recent X-ray crystal structures have captured direct lipid-receptor interaction also for other types of GPCRs. For example, a specific cholesterol binding site was observed in the case of the human beta2-adrenergic receptor (PDB ID:3D4S) (Hanson et al., 2008). This specific sterol binding site is suggested to be important for enhanced stability of the beta2-adrenergic receptor (Hanson et al., 2008). A different cholesterol-binding site, in helix 6, was detected in the high-resolution structure of the adenosine 2A receptor (PDB ID:4EIY) (Liu et al., 2012). Remarkably, this site had been previously anticipated from all-atom MD simulation (Lee & Lyman, 2012), highlighting the predictive potential of computational approaches.

Indirect effects of lipids on GPCRs—Indirect effects of the lipid membrane are mainly mediated via bulk properties of the membrane such as membrane thickness and fluidity. Mondal et al. (Mondal et al., 2011) showed that deformations around the rhodopsin receptor are radially asymmetric and involve local thickening and thinning. This asymmetric membrane adjustment allows partial alleviation of the hydrophobic mismatch between the protein and the membrane. In spite of the deformation of the surrounding lipid membrane, there remained a substantial residual hydrophobic exposure at specific regions of the protein (Mondal et al., 2011). Such residual exposure may contribute to the oligomerization of the receptors (Mondal et al., 2011). The hydrophobic mismatch between the protein and the surrounding lipid bilayer could also lead to conformational adaptation of the protein. Such a
conformational adaptation was observed in MD simulations of the metabotropic Glu2 GPCR (Bruno et al., 2012).

AMYLOIDS AND MEMBRANES

Membrane interactions in amyloid-related diseases

Amyloid-related diseases comprise a group of devastating human and animal illnesses. The most common among the human amyloid-related diseases are Alzheimer’s disease, Parkinson’s disease and type II diabetes mellitus. These diseases are characterized by abnormal protein aggregation into β-sheet rich molecular fibrils. Amyloid aggregation is a multi-step process. Initially, peptides that are natively or pathologically unfolded cluster together to form soluble bodies comprised of dimers, trimers or other low-order oligomers. Subsequently, they lead to the formation of higher order oligomers, filaments and fibrils. It is widely believed today that, with few exceptions, prefibrillar aggregates lead to cellular toxicity and mature fibrils are relatively harmless except of being a reservoir of monomers that are ready to form aggregates (Conway et al., 2000; Janson et al., 1996; Lambert et al., 1998).

Interactions between oligomers of amyloidogenic peptides and membranes or lipids appear to be one of the main reasons for the cellular toxicity of prefibrillar intermediates. For example, it has been shown that prefibrillar structures made by islet amyloid polypeptide (IAPP) monomers, whose aggregation is a pathological aspect of type 2 diabetes mellitus, can insert into membranes and lead to membrane damage (Engel, 2009). Interactions between amyloidogenic (poly)peptides and membranes, fatty-acids or other surfactants induce the formation of oligomers that appear to be involved in disease progression, through various mechanisms (see e.g., (Barghorn et al., 2005; Hasegawa et al., 2008; Vestergaard et al., 2013).

Challenges in experimental studies of protein oligomerization

Prefibrillar amyloidogenic structures are, by their nature, transient and dynamic, ranging in size from dimers to high order multimers and in structure from unordered, globular structures to filaments rich in β-sheet forming peptides. The monomers that aggregate typically lack ordered structures, and the native structures of disease-related amyloidogenic proteins such as amyloid-β (Aβ) or islet amyloid polypeptide are disordered or only partially structured (Danielsson et al., 2005; Vivekanandan et al., 2011).

The difficulty of studying amyloid aggregation experimentally is alleviated by the use of molecular simulations as an additional tool for the biophysical characterization of amyloid aggregates (Caflisch, 2006; Friedman, 2011a; Lemkul & Bevan, 2012; Shea & Urbanc, 2012). Simulations can shed light on times and length scales that may not be easily probed otherwise, providing insights into experimental findings (Engel, 2009) and leading to new testable hypotheses. Simulations suggest, for example, that amyloid fibril growth may be directional – e.g., for the HET-s prions, (Baiesi, Seno & Trovato, 2011; Friedman, 2013), and for Aβ1–40 (Buchete & Hummer, 2007; Friedman & Caflisch, 2014), but this has yet to be verified with experiments.
The sizes of many molecular oligomers, and the timescale for their formation are several orders of magnitude larger than those that can be followed by atomistic simulations. As a result, coarse-grained models are very common in the field (Derreumaux, 2013; Hung & Yarovsky, 2011; Liguori, Nerenberg & Head-Gordon, 2013; Seo et al., 2012). Fairly accurate coarse-grained force fields such as MARTINI (Marrink et al., 2007) or REACH (Moritsugu & Smith, 2007) are unfortunately not suitable to follow on amyloid aggregation where the conformational changes involve secondary structure. The approach used by Caflisch and co-workers has therefore been to develop phenomenological coarse-grained models for amyloid aggregation. In phenomenological models, a top-down approach is used to model a process. Instead of coarse-graining a collection of atoms to model membranes and peptides of different structures, lipids are represented by three beads having a single conformation (Friedman, Pellarin & Caflisch, 2009) and peptides by ten beads with two conformations: amyloid-prone $\beta$ and amyloid-protected $\pi$ (Pellarin & Caflisch, 2006). Such models are, in principle, simple and limited. Protein sequences, for example, are not modeled, and it is not possible to study features that are unique to a certain protein. Yet despite their simplicity, the models can be tailored to study biologically-relevant amyloid aggregation. The free energy difference between the $\beta$ and $\pi$ conformers can be tuned to represent peptides that form fibrils rapidly, more slowly, or not at all. Similarly, by changing the lipid’s head group size and the strength of the lipid/lipid interactions, the lipids can generate bilayered liposomes, micelles, worm-like structures and unordered aggregates, all of which are observed in experiments. The strength of the peptide-lipid interactions can be tuned, which can be used to emulate the effect of different types of biomolecular membranes (i.e., membranes whose interactions with peptides are weak, intermediate or strong).

The phenomenological coarse-grained model developed initially by Pellarin and Caflisch (Pellarin & Caflisch, 2006) and extended by Friedman et al. to include membranes (Friedman et al., 2009) have yield some interesting findings. Coarse-grained simulations were used to explain why growing filaments, but not mature fibrils, interfere with the membrane structure. An intriguing finding, that some membranes induce fibril degradation (Martins et al., 2008) that lead to the formation of smaller toxic aggregates (the fibrils themselves are not particularly toxic), was explained by molecular dynamics simulations (Friedman, Pellarin & Caflisch, 2010). Apparently, if the peptide/lipid interactions are very strong, this may lead to the disaggregation of fibrils and formation of oligomers. Finally, the same model has been used to simulate the formation of globulomers, i.e. of globular oligomers made of peptides and lipids (Friedman, 2011b). Their structures were shown to depend on the peptide-lipid ratio and the peptide’s amyloidogenicity.

A recurrent finding coming from these studies is that the amyloidogenicity of the peptides is important for the outcome of aggregation in the presence of lipids. For example, the kinetics of fibril formation is faster in the presence of membrane when highly amyloidogenic peptides are involved but slower when the peptides are non-amyloidogenic. The explanation was that unlike highly amyloidogenic peptides, peptides of low amyloidogenicity form
fibrils only in solution. This has indeed been later observed experimentally with IAPP (high amyloidogenicity) and pro-IAPP (low amyloidogenicity) (Khemtemourian et al., 2009).

**Probing the interactions of Alzheimer’s Aβ peptides with lipid membranes**

As discussed above, the toxicity of Alzheimer’s disease-related amyloid β (Aβ) peptides depends on their ability to disrupt adjacent cellular membranes. Fibrillar and globular-like Aβ oligomers can cause both non-specific and specific ion leakage, either by membrane thinning or by forming pore-like trans-membrane structures that may conduct ions. Tofoleanu and Buchete, probed the molecular interactions between preformed fibrillar Aβ oligomers and lipid bilayers, in the presence of explicit water molecules, using all-atom MD simulations (see Figure 8) (Tofoleanu & Buchete, 2012b). These interactions play an important role in the stability and function of both Aβ fibrils, and of the contiguous cellular membrane (Tofoleanu & Buchete, 2012a; Tofoleanu & Buchete, 2012b). The MD simulations revealed the relative contributions of different structural elements to the dynamics and stability of Aβ protofilament segments near membranes, and the first steps in the mechanism of fibril–membrane interaction. The simulations identified the electrostatic attraction between the charged side chains of Aβ peptides and the lipid head groups as a major driving force. In addition, hydrogen bonds were observed between specific residues in the Aβ protofilaments and the lipid head groups. Taken as a whole, these interactions facilitate the permeation of the hydrophobic C-terminus amino acids of Aβ peptides through the lipid head group region, subsequently leading to further loss of the β-sheet-rich fibril structure, on one hand, and to local membrane-thinning effects, on the other hand (Tofoleanu & Buchete, 2012b). Recent atomistic MD simulations also show that the chemical composition of the lipid head groups can control in a specific manner both the type and magnitude of interactions between Aβ protofilaments and membranes (Tofoleanu, Brooks & Buchete, 2015).

The findings summarized above advance our understanding of the detailed molecular mechanisms of Aβ-membrane interactions, and of their effects on both parties, and suggest a polymorphic structural character of amyloid ion channels embedded in lipid bilayers (Tofoleanu & Buchete, 2012a). Inter-peptide hydrogen bonds leading to the formation of long-range ordered β-strands could play a stabilizing role in the case of amyloid channel structures. However, these channels may also present a significant helical content in peptide regions (e.g., the N- and C-peptide termini) that are subject to direct interactions with lipids rather than with neighboring Aβ peptides. As currently there is no high-resolution structural information available on the amyloid channels, several models of Aβ pore-like structures traversing lipid bilayers were constructed, including helical, β-sheet and combinations of these two types of secondary structure (Tofoleanu & Buchete, 2012a). These results may guide new experiments that could test the assembly and structural features of membrane-formed amyloid channels.

**COUPLING OF LIPID AND WATER DYNAMICS IN BACTERIAL MEMBRANES**

Protein modulation by lipids also occurs in bacterial membranes. While bacteria are simple organisms, their cell envelopes are remarkably complex. The cell envelope of Gram-negative bacteria has a double membrane arrangement in contrast to the single membrane of
Gram-positive bacteria. In both cases, the membranes provide a formidable barrier to permeation of solutes both into and out of the cell. The outer membranes of Gram-negative bacteria e.g. *E. coli* and *P. aeruginosa*, are particularly complex. They are asymmetric in nature; the outer leaflet is composed of lipopolysaccharide (LPS) while the inner membrane is composed of a mixture of approximately 25 mixture of phospholipids, when all the different head group and tail combinations are included. This presents something of a challenge for MD simulations, which traditionally have treated bacterial membranes as a bilayer composed of one type of phospholipid (Bond, Derrick & Sansom, 2007; Khalid et al., 2006; Robertson & Tieleman, 2002). The reasons for the simplification of these membranes have largely been due to (i) simplicity of simulation setup and (ii) lack of parameters for LPS, (iii) the need for increased simulation lengths to allow for adequate sampling when considering lateral diffusion of multiple lipid species (ref: Disaccharides impact the lateral organization of lipid membranes (Moiset et al., 2014). While this usually an accurate representation of *in vitro* electrophysiology experiments, it clearly is far too simple to represent the heterogeneity of the *in vivo* environment.

Encouragingly, models of LPS for *P. aeruginosa* and *E.coli* have been reported in the literature in recent years. By pure serendipity models have been reported for each of the three most widely used families of force fields; Charmm, Amber and GROMOS thereby giving most membrane simulations groups an off-the-shelf solution (Lins & Straatsma, 2001; Piggot, Holdbrook & Khalid, 2011; Wu et al., 2013). The process of membrane electroporation has been studied via MD simulation (Delemotte & Tarek, 2012; Piggot et al., 2011; Tieleman, 2004).

Our recent simulation study has revealed the intricate interdependency of water and lipids in the process of electroporation (Piggot et al., 2011). The pore-forming process in a Gram-positive (*S. aureus*) and Gram-negative bacterium (*E.coli*) were compared. Our model of the outer membrane of *E.coli* comprised LPS molecules in the outer leaflet and a combination of phospholipids in the inner leaflet (phosphatidylethanolamine, phosphatidylglycerol and cardiolipin), whereas the model *S. aureus* membranes was modeled as a symmetric bilayer composed of lysyl-phosphatidylglycerol, phosphatidylglycerol and cardiolipin lipids. Briefly, our results showed two very different mechanisms of electroporation. In the symmetric membrane of *S. aureus*, the mechanism proceeds as follows: Initially, between 3–5 lipid headgroups move slightly towards the hydrophobic core of the bilayer. This is followed by water permeation into the core through the resulting ‘defect’ in the headgroup region of the bilayer. The molecules of water quickly extend through the bilayer core region and into the headgroups of the opposing leaflet, producing a continuous column or channel of water. This channel of water remained intact even when the magnitude of the external electric field was drastically reduced.

In contrast, in *E.coli*, the water channel was only observed to form, once a phospholipid had flip-flopped from the inner leaflet into the outer, LPS-containing leaflet. Furthermore, upon reduction of the applied electric field, the water was observed to exit the bilayer core, after which the headgroup region resealed, giving a defect-free bilayer. These results begin to reveal some of the molecular basis for the highly impermeable nature of Gram-negative bacteria. The LPS leaflet is tightly cross-linked by ions and extended LPS-LPS hydrogen
bonds that the individual molecules are extremely immobile and therefore difficult to penetrate. Conversely, the phospholipid molecules in the inner leaflet diffuse an order of magnitude faster enabling them to expel water and quickly reseal if defects do form.

Atomistic MD simulations of bacterial outer membrane proteins embedded within realistic membrane models have also been reported recently, building upon the earlier work of Lins & Straatsma who simulated a homology model of OprF from *P. aeruginosa* in an asymmetric bilayer composed of LPS and phospholipids (Holdbrook et al., 2013; Piggot, Holdbrook & Khalid, 2013; Straatsma & Soares, 2009; Wu et al., 2014). More recently we have shown how the conformational behavior of the TonB dependent transporter, FecA from *E. coli*, and the autotransporter Hia from *N. meningitides* (see Figure 9) differ in phospholipid bilayers compared with more representative outer membrane models (Holdbrook et al., 2013; Piggot et al., 2013).

Given the recent availability of LPS models, the outlook for the future of bacterial membrane simulations is extremely promising. In addition to the complexity of the lipids, one of the next key steps will be to incorporate realistic copy numbers of the major proteins native to these membranes, to an extent this is already being addressed with coarse-grained MD simulations (Goose & Sansom, 2013). Thus, it is easy to envisage that simulations will play a major role in understanding the fine interplay between proteins, lipids and water in bacterial membranes.

**NEW THEORETICAL CONCEPTS**

**Molecular transporters and transport processes – Membrane pores**

Upon passage through membrane pores, peptides undergo conformational transitions and sample intermediates that block the transmembrane current (Dekker, 2013; Mohammad & Movileanu, 2008; Movileanu et al., 2000; Oukhaled et al., 2007; Pastoriza-Gallego et al., 2011; Payet et al., 2012; Stefureac et al., 2008) that would otherwise flow in an open pore under a potential drop. These intermediate states can be considered jammed states. On the other hand, the phenomenology of jamming has been widely studied, separately in the context of dynamical arrest of macroscopic granular matter and in that of macroscopic glasses (Freedman et al., 2013; Head, 2009; Heussinger & Barrat, 2009; Olsson & Teitel, 2007; Silbert, 2010; Snoeijer et al., 2004; Vagberg et al., 2011). Jamming can be identified by signature characteristics in the force distributions, $p(f)$ – the force profile of a jammed system will exhibit a peak in $p(f)$ that deviates from the exponential decay in the forces of an unjammed state (Dijksman et al., 2011; Reis, Ingale & Shattuck, 2007; Snoeijer et al., 2004; Tighe et al., 2010; Tighe et al., 2005). Such jamming analyses have been applied to study protein folding in aqueous solution (Jose & Andricioaei, 2012). Accelerated MD simulations captured the atomistic details of these jammed states of a single unstructured peptide molecule, CAMA (Asandei et al., 2011; Mereuta et al., 2014; Mereuta et al., 2012), as it traverses the pore during a translocation event, jamming signatures of which were further compared with that of the same peptide when in free state without the physical constriction of the pore.
Translocation of the above mentioned peptide, Cecropin A – MA gainin(CAMA) with a sequence of 20 amino acids KWKLFKIGIKFLQSAKKF-NH$_2$, through biological nanopore α-HL (Song et al., 1996), has been studied using single molecule electrophysiology, inserted from the trans-end (Krasilnikov et al., 2000; Mereuta et al., 2014). Molecular dynamics simulations of the above process at complete atomistic resolution (Aksimentiev, 2010; Aksimentiev & Schulten, 2005; Butler, Gundlach & Troll, 2006; Butler, Gundlach & Troll, 2007; Kantor & Kardar, 2004; Tian & Andricioaei, 2005), was successful in capturing the sub-states, which induce partial or complete blockage of the pore, reflected as spikes in the current blockage signature.

The peptide in its free state assumes a heterogeneous kinked shape due to the presence of two alternate glycines (G) at positions 9 and 11 separated by one Isoleucine (I) at position 10, showing force distribution profile most reminiscent of an un-jammed peptide. When the peptide is unfolded within the pore constriction (observed through biased MD simulations) and translocated through the beta-barrel of α-HL in a linear fashion it exhibits characteristics akin to a less jammed state whereas the signatures point towards a jammed peptide when CAMA traverses the pore in a wound, condensed state. Representative snapshots of the conformational ensemble assumed at each state are depicted with their respective non-bonded force profiles in Figure 10. All three states exhibit jamming signatures, as marked by the pronounced peak above the mean force, though this peak is shifted more to the right the more unjammed the peptide, e.g. the free CAMA, which is shifted farthest. Thus, it may be surmised that the position of the peaks depends upon the relative degree of jamming between the configurations. It is noteworthy that the simulation for the free CAMA was carried out at room temperature where jamming is shown, from our previous study, to be extant in peptides. A simulation at a much higher temperature will most likely wash out the jamming traits in the force profile, which may be the focus of investigation of a future study.

The Energy Landscapes Framework

Studying the potential energy landscape of membrane proteins provides both conceptual and computational tools for understanding a wide range of observable properties in molecular science (Wales, 2003; Wales, 2010; Wales, 2005; Wales & Bogdan, 2006). In particular, we can exploit stationary points (minima and transition states) for structure prediction and analysis of global thermodynamic and kinetic properties.

Basin-hopping global optimization (Wales & Doye, 1997) represents a powerful tool for structure prediction, while basin-sampling/parallel tempering (Bogdan, Wales & Calvo, 2006; Wales, 2013) and discrete path sampling (Wales, 2002; Wales, 2006; Wales & Dewsbury, 2004) enable us to address broken ergodicity and rare event dynamics. Basin-hopping and basin-sampling require only local minimization; discrete path sampling involves location of transition states between local minima. This coarse-graining in terms of stationary points is founded upon efficient geometry optimization procedures; the corresponding software and energy landscape databases are available at URL http://www-wales.ch.cam.ac.uk.

Applications have been presented ranging from prediction and interpretation of high resolution spectra for small water clusters, (Wales & Walsh, 1997; Walsh & Wales, 1996)
folding and misfolding of biomolecules, (Evans & Wales, 2003; Mortenson, Evans & Wales, 2002; Strodel, Whittleston & Wales, 2007) to the analysis of structural glass-formers (de Souza & Wales, 2008; Wales & Doye, 2003) and condensed soft matter systems (Chakrabarti & Wales, 2009; Fejer & Wales, 2007). Some recent results for small oligomers of the Aβ₁₋₄₂ peptide are illustrated in Figure 11 (Strodel et al., 2010).

The structures illustrated in Figure 11 were obtained using CHARMM19 (Neria, Fischer & Karplus, 1996) with the implicit membrane potential IMM1 (Lazaridis, 2003). A basin-hopping/parallel tempering scheme with exchanges between basin-hopping runs at different temperatures was used, (Strodel et al., 2010) together with intra-and intermolecular coordinate moves for the peptides. The most favorable monomer transmembrane structure has residues 17 to 42 inserted in the membrane. The most stable octamer structures can be viewed as displaced tetramers composed of two or three β-sheets (Strodel et al., 2010).

Coarse-grained modeling has been used to explain the underlying principles that lead to the emergence of shells, tubes, helices and spirals (Chakrabarti & Wales, 2009; Fejer & Wales, 2007). Recently we have discovered design principles that connect the morphology of amyloid fibers to seedpods and macroscopic helices formed from elliptical magnets (Figure 12) (Forman et al., 2013). Experimentally, the introduction of a cytochrome domain into an amyloid fiber permits dynamic adjustment of the fiber morphology via heme binding (Forman et al., 2012). The interlocking of fiber filaments introduces systematic kinking, (Forman et al., 2012) and a transition from a twisted ribbon to a spiral ribbon morphology was also observed (Baldwin et al., 2006). By constructing a rigid link between two anisotropic interacting units we have now shown that the precise helical morphology of aggregates formed from such a composite building block primarily depends on the internal geometry, rather than the nature of the interaction or the anisotropy of the interacting units. This framework can be viewed as a discrete version of the bilayered frustration principle, which drives the morphological transitions of the Bauhinia seedpod.

**Biomolecular Visualization**

Biomolecular visualization has played an important role since the beginning of computer-assisted biophysical and biochemical analysis. In the following, we present some visualization techniques that bear a great potential for the analysis of structure, dynamics and function of membrane proteins. Some of these techniques have been specifically developed for membrane proteins and others were designed with different molecular structures in mind. Nevertheless, we believe that all of them are readily applicable to support the analysis of membrane substructures, single membrane proteins or even whole membranes.

**Molecular conformations**—During the past two decades methods have been developed to identify metastable conformations of biomolecules and to reveal the dynamics of conformations (Schütte & Sarich, 2013). Here, a conformation is considered as a connected subset in state space. In order to better understand molecular systems, it is useful to depict molecular conformations, making the differences between them clearly visible. In references (Rheingans & Joshi, 1999; Schmidt-Ehrenberg, Baum & Hege, 2002), an approach has been
proposed to represent sets of molecular configurations as spatial densities; these can be depicted using iso-surfacing or volume rendering; and with color-coding, different parts of the molecules can be distinguished or distinct conformations can be rendered in superposition (Figure 13).

Molecular configurations that can be transformed into each other by a rigid transformation are considered physically equivalent. Therefore, configurations are usually defined without reference to an external coordinate system. For a graphical representation, however, they must be placed in a given external coordinate system. While this is no problem when displaying a single molecular system, especially when one can rotate and move it interactively, the optimal alignment for sets of molecular systems is less clear; to create informative depictions of such sets, the molecules must be aligned relative to each other in a specific way. For illustration, in Figure 14 four different types of alignments of a set of molecules are shown using the simple molecule pentane. For sufficiently similar geometries, generalized partial Procrustes analysis provides an objective function (sum of squared distances between corresponding atoms, or, equivalently sum of atom position variances) and an alignment algorithm. If groups of substantially different forms are to be compared, however a more sophisticated alignment strategy is necessary. In Ref. (Schmidt-Ehrenberg, 2008) it is proposed to perform a hierarchical decomposition of the molecule into maximally stiff and less stiff structures and to apply an alignment strategy adapted to this decomposition. The resulting visualization provides a clear impression of the stiff structure, while the rest of the molecule, exhibiting interesting form differences between the metastable sets, are depicted fuzzy.

**Extraction and visualization of molecular paths**—Before molecular interactions can take place, the interacting molecules must come close enough to each other. For molecules like nanotransporters, which are often complex and endowed with specific binding sites, this is of particular interest. Identification of ligand binding sites via molecular dynamics simulations, however, is too expensive. In Ref. (Lindow, Baum & Hege, 2011) it was suggested to restrict the search space by first computing the paths that are geometrically possible. These paths can be found by computing the 3D Voronoi diagram of the atom spheres and selecting a subset of the Voronoi edges (i.e. all points that locally are furthest away from the van der Waals spheres of the molecule). Utilizing computer graphics techniques like deferred shading, ray-casting and screen space ambient occlusion, and by placing light sources along the paths, one obtains very informative depictions (Figure 15). All of these techniques are directly applicable to, for example, ion channels. In contrast to static channels, which can be observed in a single time step of a molecular dynamics trajectory, the evolution of molecular cavities needs to be analyzed over time to identify dynamic channels. This aspect is addressed below.

**Interactive exploration of cavity dynamics in proteins**—The internal cavities of proteins are dynamic structures and their dynamics is associated with conformational changes. In Refs. (Lindow et al., 2012a; Lindow et al., 2013), a tool was presented that allows rapid identification of internal cavities and an interactive assessment of their time-dependent shapes. Given a MD trajectory, in a pre-processing step the structure of the
cavities is computed from the Voronoi diagram of the van der Waals spheres. Then the user can interactively select, visualize and analyze the time-dependent cavity components and the dynamic molecular paths with a user-defined minimum constriction size. The cavity dynamics can be depicted in a single image that shows the cavity surface colored according to its occurrence in time. As an example, in Figure 16 the dynamic structure of internal cavities in the bacteriorhodopsin proton pump is depicted and in Figure 17 the related splits and merges of the cavities as well as their position along a user-defined axis through the proton pump are shown. An alternative technique, based on object-space ambient occlusion for the detection of cavities, channels and pockets has been proposed in Refs. (Krone et al., 2014; Krone et al., 2013).

Multi-scale visualization of biological structures—Properties of inorganic and organic materials can often be understood by analyzing them simultaneously on different scales, e.g. on the atomic and nanoscopic scale. This raises the problem of rendering images with millions to billions of atoms. In Ref. (Lindow, Baum & Hege, 2012b), a simple yet efficient rendering method was proposed for visualizing interactively large sets of atomic data, bridging 5 orders of magnitude in length scale. The method exploits the recurrence of molecular structures and the fact that these are typically rendered opaquely, such that only a fraction of the atoms is visible. The method has been demonstrated by visualizing large molecular structures that where fitted to subcellular objects which had been imaged with electron tomography. On a commodity PC, 150 million atoms could be visualized with interactive speed; and for scenes with up to 10 billion atoms, still rates of 3 frames per second could still be achieved. An example is shown in Figure 18, depicting microtubules reconstructed from electron tomography data with fitted molecular structure. This method was further optimized in Ref. (Falk, Krone & Ertl, 2013) and extended to render triangular objects, such that arbitrary molecular models can be displayed.

Since membranes also consist of a large number of recurring molecules, the described methods should be directly applicable to data from membrane simulations. A great challenge here is to visualize membrane dynamics at an interactive speed.

CONCLUSIONS

The above considerations of recent progress in membrane protein structural biology present a unifying theme: even in atomic-detail considerations one cannot escape at least elements of a holistic approach. The protein, lipid and solvent components of a biological membrane system have been shown to be highly interdependent, such that consideration, say, of a simulation model of a protein in a single model phospholipid, such as DOPC, is becoming rather anachronistic. Increased computer power, improved computational methods and experimental advances in techniques such as, for example, the crystallography of large complexes are revealing the workings of integrated membrane systems of hundreds of thousands of atoms. Undoubtedly in the future it will be necessary to integrate these approaches still further, with “systems-level” imaging and simulation. Possibly only in the context of non-equilibrium cellular fluxes will the mechanics of membrane function be able to be fully comprehended.
Acknowledgments

ZC, ANB, JCS would like to acknowledge funding from Centre Européen de Calcul Atomique et Moléculaire (CECAM) to host the Workshop “Coupling between protein, water, and lipid dynamics in complex biological systems: Theory and Experiments” that took place in September 2013, Lausanne, Switzerland. JTD, IA, and MR used the computational resources of the Modeling Facility of the Department of Chemistry, University of California Irvine funded by NSF Grant CHE-0840513 for this work. A-NB is supported in part by the Marie Curie International Reintegration Award IRG-26920. TWA was supported by ARC DP120103548, NSF MCB1052477, DE Shaw Anton (PSCA00061P; NRBSC, through NIH RC2GM093307), VLSCI (VR0200) and NCI (dd7). BA and SV acknowledge the support by ERC advanced grant #268888. ZC and PG would like to acknowledge Reference Framework (NSRF) 2011 – 2013, National Action “Cooperation”, under grant entitled “Magnetic Nanoparticles for targeted MRI therapy (NANOTHER)”, with code “11ΣYN-1-1799”. The programme is cofunded by the European Regional Development Fund and national resources. Part of the calculations presented herein were performed using resources of the LinkSCEEM-2 project, funded by the EC under FP7 through Capacities Research Infrastructure, INFRA-2010-1.2.3 Virtual Research Communities, Combination of Collaborative Project and Coordination and Support Actions (CP-CSA) under grant agreement no RI-261600. GB is supported in part by NSF grant MCB1330728 from the National Science Foundation, and grant PO1GM55876-14A1 from the National Institutes of Health. LD receives funding from EU FP7 (PIOF-GA-2012-329534). LD and MLK used the computational resources of Temple University, supported by the National Science Foundation through major research instrumentation grant number CNS-09-58854. FT and NVB are grateful for financial support from the Irish Research Council, and for using the computational facilities of the Biowulf Linux cluster at the National Institutes of Health, USA and the Irish Centre for High-End Computing (ICHEC). JS acknowledges support from the Instituto de Salud Carlos III FEDER (CP12/03139) and the GLISTEN European Research Network.

References


Friedman, R. Wild type and mutants of the HET-s(218–289) prion show different flexibility at fibrillar ends: A simulation study. Proteins. 2013 In Press: 0.


Goddard, AD.; Dijkman, PM.; Adamson, RJ.; Watts, A. Methods Cell Biol. Elsevier Inc; United States: 2013. Lipid-dependent GPCR dimerization; p. 341-57. copyright 2013


Krasilnikov OV, Merzlyak PG, Yuldasheva LN, Rodrigues CG, Bhakdi S, Valeva A. Electrophysiological evidence for heptameric stoichiometry of ion channels formed by


Middlemas DS, Raftery MA. Identification of subunits of acetylcholine receptor that interact with a cholesterol photoaffinity probe. Biochemistry. 1987; 26:1219–23. [PubMed: 3567168]


Fig 1.
Hypothetical model for full-length *E. coli* GlpG dimerization. A cartoon illustration is shown in rainbow coloring with N-terminus in blue. The membrane domain of *E. coli* GlpG is dimeric. In this model, the dimer interface occurs through interaction of loop 1 that extends partially in the lipid bilayer (light blue). The cytoplasmic domain is separated from the membrane domain by a flexible linker. The structure of the cytoplasmic domain is from Ref. (Lazareno-Saez et al., 2013).
Fig 2.
(a) Fully atomistic KcsA ion channel system (protein as yellow ribbons, lipid tails as chains, water as sticks and K$^+$ and Cl$^-$ ions as red and grey balls, respectively). (b) Differing multi-ion barriers for K$^+$ and Na$^+$. The low barrier to K$^+$ outward conduction results from the ability of a K$^+$ ion (red) to bind to the entrance of the selectivity filter. (c) The shift of binding site for Na$^+$ (green), eliminates this process and requires a substantial amount of energy. Based on (Nimigean & Allen, 2011).
Fig 3.
The transmembrane Kv1.2 voltage-gated cation channel.
Fig 4.
Thermodynamic cycle used to compute free energy differences for estimating the difference in the free energy of binding of a ligand to a protein (Cycle 1) or for the relative solvation free energy difference (Cycle 2).
Fig 5.
Structures of M2TM inhibitors and their binding constants. On the left side, note the excellent correlation between the FEP/MD calculations with a flexible protein and the experimental data. The image is adapted from (Gkeka et al., 2013).
Fig 6.
A view of the transmembrane domain of the nicotinic acetylcholine receptor from Torpedo marmorata, with proposed coordinates for embedded cholesterol molecules. Each subunit contains four membrane-spanning helices (M1 purple, M2 green, M3 blue, M4 cyan) and three bound cholesterol molecules (yellow, orange and red). The image is based on the crystal structure PDB ID: 2BG9 from Ref. (Unwin, 2005).
Fig 7.
Spontaneous re-binding of cholesterol to an intersubunit cleft in the GABA$_A$ receptor. Two frames from a simulated MD trajectory are shown (Hénin et al., 2014). TM helices M2 (right) and M3 (left) forming the back of the pocket are shown as white cartoon, and cholesterol as orange bonds. Other specific groups are depicted as bonds using the following color code: yellow – phospholipid, orange – membrane cholesterol, green – protein amino acid residues from the back receptor subunit, and cyan – protein groups from the front receptor subunit. Data from (Hénin et al., 2014).
**Fig 8.**
Cartoon representation of an Aβ protofilament segment (4 two-peptide layers) interacting with a lipid membrane, lateral (left) and top (right) views. The N-terminus is represented in red, the turn region in green and the C-terminus in blue. The lipid heavy atoms are represented as spheres, the headgroups are colored in yellow and the lipid tails in grey, respectively. The image is based on Ref. (Tofoleanu & Buchete, 2012b).
Fig 9.
Electroporation in the cell membrane of the Gram-positive bacterium, S. aureus. Deformation of the lipid bilayer is clearly visible. Phospholipids that have been pulled in closer to the membrane core are shown in green and red, water molecules are shown in blue and the remaining lipid molecules are shown in pink (Piggo, Piñeiro & Khalid, 2012).
Fig 10.
The non-bonded, normalized force distributions for three different states of CAMA, ranging from jammed (wound) to marginally jammed to relatively un-jammed (free). All three exhibit jamming signatures, namely a pronounced peak about the exponential line (dashed), though this is shifted to higher values the more apparently unjammed the peptide. Additionally, the tails of those that are less jammed are closer to the exponential line.
Conformational ensembles of CAMA P6 in (b) wound state (c) linear and (d) free state. (e) Translocation of peptide CAMA – P6 through lumen of α-hemolysin. Panels b–e are adapted from Ref. (Mereuta et al., 2014).
Fig 11.
Predicted structures for Aβ_{1-42} monomer (transmembrane and surface structures, left), dimer (middle) and octamer (right) in contact with an implicit membrane. The molecular graphics images are adapted from Ref. (Strodel et al., 2010).
Fig 12.
Left: introduction of a cytochrome domain into an amyloid fibre can change the morphology from twisted to spiral ribbons and induce systematic kinking. Centre: rigid building blocks consisting of two ellipsoids can reproduce these structures. Right: the structure depends mostly on the internal geometry of the building blocks. The molecular graphics images are based on Ref. (Forman et al., 2013).
Fig 13.
Direct volume rendering of the BSI molecule 2-(biphenyl-4-sulfonyl)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid based on 241,760 molecular configurations in 192 metastable sets, comprising 12 chemically different groups that are distinguished by color. On the left the spatial alignment is based on a fixed substructure, on the right on Procrustes analysis.
Fig 14.
Illustration of the alignment problem by aligning 150 forms of Pentane by minimizing (a) the distances between the centers of gravity of each form, furthermore, the distances between corresponding atoms summed over (b) all atoms, (c) three selected atoms, and (d) atoms in a hierarchy of sets, defined by stiff substructures and an interactively selected anchor structure.
Fig 15.
Geometric molecular paths (blue) in a nano-transporter, computed by the Voronoi diagram of the atom spheres and combined with a path-dependent illumination of the molecular surface. Based on Ref. (Lindow et al., 2011).
Fig 16.
Cavity dynamics in a bacteriorhodopsin monomer. From left to right: the cavities traced at time step $t = 63$, a channel created by the dynamics from $t = 50$ till $t = 64$, and the overall spatial cavity probability (depicted with maximum intensity projection). The images are based on Refs. (Lindow et al., 2012a; Lindow et al., 2013)
Fig 17.
Temporal development of cavities in a bacteriorhodopsin monomer (for the simulation and the time steps also used in Figure 16); upper diagram: topology graph showing splits and merges of cavities; lower diagram: penetration graph depicting the cavities’ location along a user-defined axis. The image is based on Refs. (Lindow et al., 2012a; Lindow et al., 2013)
Fig 18.
Visualization of microtubules (reconstructed from electron tomography data) containing about 10 billion atoms; rendered on commodity PC with subsecond framerate. The image is based on ref. (Lindow et al., 2012b).