Biochimica et Biophysica Acta 1859 (2017) 2021–2039

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Review

Membrane proteins structures: A review on computational modeling tools

Jose G. Almeida [a,1], Antonio J. Preto [a,1], Panagiotis I. Koukos [b], Alexandre M.J.J. Bonvin [b], Irina S. Moreira [a,b,*]

[a] CNC - Center for Neuroscience and Cell Biology, Rua Larga, FMUC, Polo I, 1°andar, Universidade de Coimbra, 3004-517, Coimbra, Portugal
[b] Bijvoet Center for Biomolecular Research, Faculty of Science - Chemistry, Utrecht University, Utrecht, Padualaan 8, 3584CH, The Netherlands

ARTICLE INFO

Article history:
Received 23 March 2017
Received in revised form 4 July 2017
Accepted 13 July 2017
Available online 15 July 2017

Keywords:
Membrane proteins
Computational modeling
Machine-learning
Transporters
GPCRs

ABSTRACT

Background: Membrane proteins (MPs) play diverse and important functions in living organisms. They constitute 20% to 30% of the known bacterial, archaean and eukaryotic organisms' genomes. In humans, their importance is emphasized as they represent 50% of all known drug targets. Nevertheless, experimental determination of their three-dimensional (3D) structure has proven to be both time consuming and rather expensive, which has led to the development of computational algorithms to complement the available experimental methods and provide valuable insights.
Scope of review: This review highlights the importance of membrane proteins and how computational methods are capable of overcoming challenges associated with their experimental characterization. It covers various MP structural aspects, such as lipid interactions, allostery, and structure prediction, based on methods such as Molecular Dynamics (MD) and Machine-Learning (ML).
Major conclusions: Recent developments in algorithms, tools and hybrid approaches, together with the increase in both computational resources and the amount of available data have resulted in increasingly powerful and trustworthy approaches to model MPs.
General significance: Even though MPs are elementary and important in nature, the determination of their 3D structure has proven to be a challenging endeavor. Computational methods provide a reliable alternative to experimental methods. In this review, we focus on computational techniques to determine the 3D structure of MP and characterize their binding interfaces. We also summarize the most relevant databases and software programs available for the study of MPs.

Contents

* Corresponding author at: CNC - Center for Neuroscience and Cell Biology, Rua Larga, FMUC, Polo I, 1°andar, Universidade de Coimbra, 3004-517, Coimbra, Portugal.
E-mail address: irina.moreira@cnc.uc.pt (I.S. Moreira).
[1] Equal contribution.

http://dx.doi.org/10.1016/j.bbamem.2017.07.008
0005-2736/

1. Introduction

Membrane proteins (MPs) have diverse functional roles, featuring important functions such as ion and molecule transport, immune system molecule recognition and energy transduction [1]. It is therefore fundamental to comprehensively understand their structure and structure-function relationships. 3D structures of various MPs have been characterized in recent years by several experimental methods, such as Nuclear Magnetic Resonance (NMR), X-ray crystallography and cryo-electron microscopy [2]. MPs, unlike soluble proteins, are difficult to analyze in their native environment, due to their insertion in the lipidic membrane [2–3]. They are affected by the membrane and various specific factors, such as cholesterol content [4] and hydrophobic thickness of the lipid bilayer [5], but also influence the membrane structure itself [5a]. All these aspects contribute to the technical experimental difficulties in the structural characterization of MPs, which explains their relatively low number in the Protein Data Bank (PDB) [6], despite their high proportion in the human proteome [7].

The computational prediction of soluble protein structure can be considered a particularly advanced field, both in terms of variety of approaches and the accuracy they can achieve [8]. However, computational prediction of MPs and their interfaces, especially when studying dimers or high-order oligomers, is still in its early days [9]. Current approaches are usually based on a combination of homology modeling [10] or de novo protein structure determination [11] with ML algorithms [12] to predict binding interfaces and/or intermolecular contacts, and MD simulations to refine the models and study their dynamical properties [13].

Some MPs are of particular interest for therapy assessment and drug targeting given their role in physiological processes and biochemical pathways. Among them are G-protein Coupled Receptors (GPCRs), ion channels and transporters. All these cover a wide array of functions while maintaining some common traits among their respective (super)families. Here, we aim at giving a brief overview of MP and the experimental methods for determining their structure, followed by a comprehensive assessment of known computational methods for the prediction of MP structure and structure-related characteristics, such as topology and binding interface prediction. Lastly, we highlight some recent computational studies on key MPs and their main features.

2. Membrane proteins

MPs have been defined as proteins associated to lipid domains, which are involved in communication, regulation and structural coherence. In fact, proteins that entirely or partially span the membrane (intrinsic/Trans membrane (TM) proteins), as well as proteins that are peripherally membrane-bound (peripheral MPs – PMPs), can carry out these functions. Due to the high amount of information and computational methods for MPs, we focused on TM proteins, which will be referred to as MPs. For readers interested in PMPs, specialized reviews can be found covering this class of membrane proteins [14], their interaction with the membrane [15] and the experimental and computational methods for their study [16].

Only a detailed understanding of MP structure-function relationships will allow the understanding of common pathologies at a molecular level and the development of improved pharmacological procedures [17,18,19]. The most functionally relevant intrinsic MPs are typically split into ion channels, membrane receptors and transporters [1a,20]. Ion channels facilitate the diffusion of ions across membranes, bridging the intra- and extracellular environments across the hydrophobic lipid bilayer by allowing hydrophilic molecules and ions to pass through the membrane. Ion channels are structurally modulated by the TM electrochemical potential, the binding of ligands, and mechanical stress and/or changes in the local lipid environment [21]. In some cases, this modulation is required for biological function [22]. Membrane receptors, comprising GPCRs as well as olfactory receptors (ORs) and nuclear receptors [23], play roles in different biochemical and signaling pathways, and in triggering environment, immune, hormonal and neurological responses, which makes them highly interesting targets for therapeutical investigation. They often share common structural traits, allowing for their classification into protein families or superfamilies. Transporters span the cell membrane with recurring specific membrane topologies, energy coupling mechanisms and substrate specificities. They are capable of transporting molecules and ions across the membrane, triggering environment-driven responses, delivering essential nutrients and disposing cellular waste.

MPs as defined in this review consist typically of several domains: extracellular (typically involved in cell-cell signaling and/or interactions), intracellular (performing a wide range of functions such as activating signaling pathways and anchoring cytoskeletal proteins) and intramembrane (such as pores and channels) [24]. TM proteins in general are amphipathic, meaning that they have different electronegativity and hydrophobicity profiles along their structure, allowing them to be both in contact with water (hydrophilic environment) and the membrane (hydrophobic environment). The structure and function of many TM proteins depend on Post Translational Modifications (PTM) such as phosphorylation and glycosylation. The two major recurrent protein structure motifs in MPs are TM α-helices [25], repeatedly crossing the membranes in α-helical bundles and β-strands arranged into super-secondary structures known as β-barrels [26].

3. Experimental structural determination of membrane proteins

Despite their functional importance, only 4.193 structures of membrane proteins (or rather of sub-domains) can be found among the 131.485 determined protein structures deposited at the PDB [7] (statistics from June 29th 2017) (Fig. 1). This means that <1% of all determined protein structures belong to MP families. This number includes multiple submissions of the same protein under a variety of experimental conditions. In contrast to the limited number of available MP 3D structures, there are 199.322 MP sequence clusters according to UniProt's UniRef (June 29th 2017).

Two major factors can explain this discrepancy: i) difficulties in both expression, which can be done in several organisms [27] but mostly in *Escherichia coli* (*E. coli*) [28] and purification processes; ii) challenges associated with the actual determination of the 3D structure of the

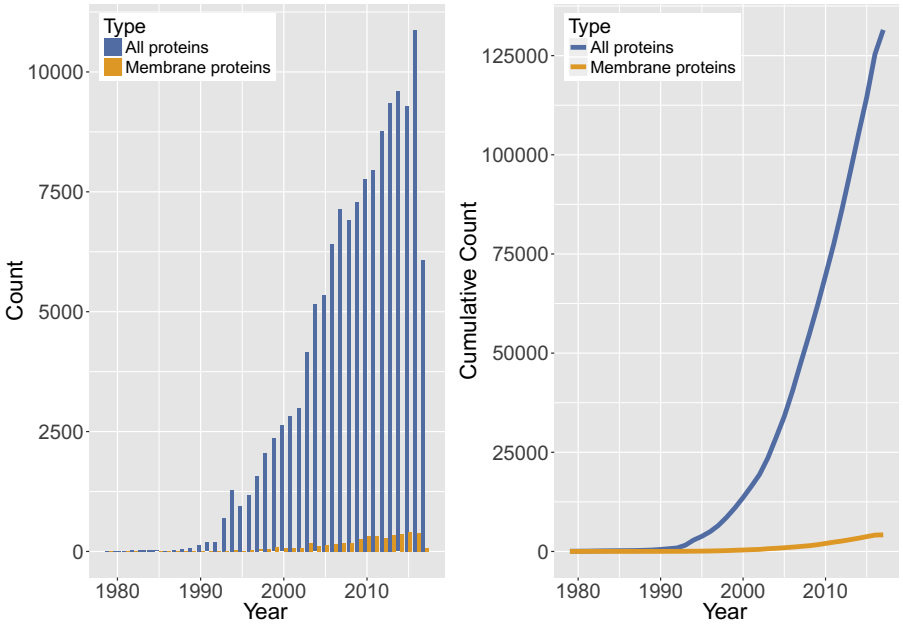

Fig. 1. PDB entries by year of deposition (left) and cumulative number (right). The total number of 3D structures is 131.485 as of June 29th 2017. As is evident in the left plot the number of MP 3D structures (yellow) being deposited every year has plateaued whereas the total number of entries has been steadily increasing since the early 1990s. Source: http://www.rcsb.org/pdb/statistics/contentGrowthChart.do?content=total&seqid=100

purified MPs through NMR, X-ray crystallography and cryo-electron microscopy (the three most common methods for protein structure determination). Concerning the first point, overexpression of MPs usually leads to cytoplasmic aggregates and changes in the cell metabolism [29]. A few methods have been devised to avoid cytotoxicity, such as using and tuning *E. coli* strains that are not as affected by the protein overexpression (a well-known example being "Walker strains") [30]. Protein extraction and purification can be troublesome as well, since different conditions can lead to different outcomes when it comes to protein stability, state and viability for structure determination [31] (these conditions may come down to something as apparently simple as choosing the right detergent for MP isolation [32]).

3.1. Structure determination

Structures solved by X-ray crystallography are often the result of a high amount of time invested in fine tuning the best possible experimental conditions. After establishing good initial crystallization conditions, further optimization is required [33], namely detergent addition, use of different 3D continuous lipid phases (allowing the protein to freely flow) [34] or antibody fragments to stabilize the protein structure [35]. The latter often results in more stable crystals, but the MP conformation might differ from its native state due to the additional interactions with the antibody fragments. Data collection and analysis can also be problematic, as the variability of crystals and their conditions (i. e. hydrophobic protein regions camouflaged by hydrophobic solvent, making it difficult to assess the transmembrane MP structure) might prevent automated and stable data acquisition and processing [33]. Three noteworthy examples to illustrate the challenges associated with the experimental structure determination of MPs using X-ray crystallography are the aspartate protease, which required 160.000 different conditions in order to achieve good, analyzable crystals [36], an engineered human β_2-adrenergic GPCR, which took 15 years to be solved [37], and the 13 year-long structure determination of the membrane-integral diacylglycerol kinase [38], as noted by Leman et al. in their 2015 review paper [9].

Structure determination by solution NMR spectroscopy has come a long way as well, but some major drawbacks can still be identified: The low sensitivity, the size limitation and the intrinsic motions of the system under investigation. When it comes specifically to MPs, more problems can be identified, such as sample preparation and spectral crowding [39]. Nonetheless, NMR has proven useful to study the dynamics (e.g. relative population and conformation of different states, exchange rates, internal motions) of MPs undergoing conformational changes, such as channels, transporters and receptors [40]. Recently, new techniques such as solid state NMR (ssNMR) have provided much better results when compared to liquid phase NMR, as there is no molecular weight cap. However, this does not prevent spectral crowding. Compared to X-ray crystallography, NMR, and in particular ssNMR, has the great advantage that is allows to study MP in an actual membrane environment and not in a "detergent simulation" of a membrane [41]. MP structure determination has also been conducted using paramagnetic tags, a technique focused on labelling MPs so that they can later be analyzed with NMR [42] and/or Electron Paramagnetic resonance (EPR) [43]. Recently, it has even been demonstrated that MPs can be studied by ss NMR in their native cellular environment [44].

Cryo-electron microscopy (Cryo-EM) is a technique that has recently gained a lot of popularity among structural biologists. Its main aspect is the imaging of radiation-sensitive entities – cells, viruses and macromolecules – under cryogenic conditions using a transmission electron microscope [45]. It offers great advantages over X-ray crystallography as it does not require crystallization. Its main drawback is the relatively low resolution for membrane proteins when compared to X-ray structures. Some recent examples of MP structures solved by cryo-EM are the transient receptor potential channel 1 at 3.4 Å [46] and the chloride conducting (CLC) ion channel at 3.7 Å [47]. A unified database for protein structures determined through cryo-EM – EMDataBank – is publicly available at http://emdatabank.org/index.html [48].

3.2. Interaction with the lipid environment

When considering MPs, the lipid environment is essential in defining their structure and function, often significantly changing the proteins' properties. While most MP structures are not easily determined, it is useful to note that some MPs can retain their structure and function while in soluble form, which can be tested through their expression as fusion proteins [49]. When this is not possible, detergents can be used to solubilize the expressed proteins [50] by extracting them from the

membrane, ideally without affecting their structure. Distinct detergents, with different hydrophobicity properties, can be used depending on the protein's properties [51]. The choice of the detergents can be time and resource consuming, with no guaranteed results [51–52]. The use of detergents leads to micelle-like structure formation, which is not an accurate representation of the bilayer environment and can result into deformations in structure. Some approaches to overcome these problems include the inclusion of MPs in nanodiscs – detergent free membrane-like structures stabilized by polymers or proteins, which allow for liquid-state NMR studies [53] – and the lipid cubic phase method [54]. The latter works by isolating a biological membrane with the target protein and solubilizing it with detergent. The resulting micelle is purified and homogenized with monoacylglycerol, and contains a bilayer with the target protein [55].

Membrane domains, such as lipid rafts, can change significantly the structure and function of some proteins as these domains have different properties (i.e. high glycosphingolipids content) [5b,56]. In lipid rafts, solvent extraction can be less effective, since these are more effective at retaining MPs than other lipid membrane domains. This works either by surrounding the protein with a tighter and more ordered lipid packing, or by other mechanisms, such as anchoring [5c]. Furthermore, even when not considering lipid rafts or lipid raft-like domains, other lipid structures and molecular organizations (depending on factors such as temperature, pressure, lipid composition and other proteins) can influence the membrane structure, which, in turn, can affect membrane-inserted proteins. This is usually referred to as lipid polymorphism, to which distinct lipid phases are associated, and which has been observed to play a role in G-protein structure and function [5d].

Some intrinsic protein properties such as hydrophobicity, van der Waals (vdW) interactions, prosthetic groups, among others, can play a major role in the interaction between the protein and the membrane. Hydrophobic mismatch, for instance, occurs when the thickness of the bilayer's hydrophobic section does not correlate with the length of the hydrophobic residues of the membrane, generating a mismatch, as characterized for example by calorimetry [57], NMR [58] and fluorimetry techniques [5a,59]. Further changes in the membrane can occur upon insertion and formation of dimers or even high-order oligomers, for example, which contributes toward the complexity of MP-membrane interactions. Other relevant changes are the insertion of peripheral groups (adding a step to the usual two step model considered for MPs' inclusion and dimerization/oligomerization) such as prosthetic groups, more elaborate protein folding, generation of new binding surfaces or portioning of space away from the lipid. This can be studied through a combination of kinetic analysis and NMR [5a,59b,60].

3.3. Oligomerization

Membrane proteins tend to form dimers (homo- or hetero-dimers) or higher-order oligomers by establishing contacts between specific TM domains. TM Helix Association has been proposed to occur in two steps (although, as mentioned above, other models have been proposed that consider additional steps [60]): first the insertion of helices in the membrane, then their association [61]. Methods like Analytical Ultracentrifugation (AUC) [59a], Thiol-Sulfide Exchange [62], Forster Resonance Energy Transfer (FRET) [5h], Steric Trapping [5e] and Genetic Systems [63] can be used to study this phenomenon. Analytical Ultracentrifugation (AUC) [59a] provides information on protein association strength and stability, by measuring sedimentation velocity and equilibrium (SV and SE, respectively) upon centrifugation. Thiol-Sulfide Exchange Method [62] enables the study of TM helix-helix association in a membrane-mimicking environment. This method promotes TM association by disulphide crosslinking, and allows the detection of dimer interfaces [64]. Since cysteine residues are labelled, and the experiment is conducted in a redox environment, it provides information on TM helices orientation in membrane-mimicking environments. FRET [5h] can be used to report the energetics of a system by following the energy transfer between a donor and an acceptor molecule, which in turns provide distance information. When TM helices have donor and acceptor molecules, their association distances can be monitored, in membrane mimicking environments such as liposomes, providing information about the conformational changes they induce in each other. Steric Trapping [5e] tests helix interaction strength by competitive binding with monovalent streptavidin (mSA). It can probe different levels of binding strength by using mutated mSA with weakened interaction propensity. Genetic Systems [63] can inform about TM helix association by introducing reporter genes. When a dimer is formed, a DNA-binding domain activates a reporter gene, thus reporting on the dimerization. As an alternative and to complement experimental and lengthy structure determination methods, new, less time-consuming computational approaches to MP structure predictions have been developed over the years, which will be reviewed in the next section.

4. Computational structure prediction of MPs

4.1. Membrane protein structure prediction

Even though various computational methods have been developed for the prediction of soluble protein structures, most of these cannot be directly applied to MPs as the latter are incorporated in a very different environment. All available tools and algorithms had to be adapted, in particular the “solvent” representation has to be changed to create or mimic a hydrophobic layer within a hydrophilic environment. The protocols were developed and tested on well-defined MP databases. To predict MP structure, it can be helpful to consider cytosolic/extracellular partner interactions, since their interacting motifs are easier to study than those responsible for MP-MP interaction. By doing so, part of the cytosolic/extracellular regions of a protein can be determined, making it easier to identify membrane spans. Experimentally determined MP structures in different conformations may also help in achieving more accurate predictions, as some predictions might recreate only one conformation, while important interactions are also taking place in other possible conformations. Furthermore, membrane lipid composition should be considered when attempting to predict the structure of a MP as it affects the conformation and activation state of membrane-embedded proteins [65].

For the determination of tertiary structures, the most common methods are de novo methods and homology modeling [66]. De novo methods can also make use of already determined features such as secondary structure, or TM spans and topology. Incorporating that information can drastically reduce the computational costs, which nonetheless remain too high for these methods to be routinely applied on most cases of interest. Sequence-based homology modeling provides the best results within a reasonable time-frame, but depends entirely on the availability of homologous proteins with resolved 3D structures. The availability (or rather scarcity) of homologues is particularly relevant for MPs since, as was previously discussed, the number of unique MP 3D structures is significantly lower than that of soluble proteins. Some methods have been developed specifically for membrane protein modeling, namely MEMOIR (Membrane protein modeling pipeline), [67] which can model the 3D structure of a protein of known sequence provided there are available homologous MPs with determined 3D structures, and MEDELLER [68], which has provided interesting results thanks to its tailor-made MP structure prediction – a sequential prediction of protein core and loops. MEDELLER will not generate 3D coordinates for regions for which the prediction is uncertain. This has the advantage of rendering the models more accurate but also slightly more incomplete. Structural homology modeling (threading) can overcome the lack of homologues for given sequences, however, as already mentioned, the small number of experimentally available MP structures can lead to insufficient sampling. An example of a pipeline using threading is TMFoldWeb [69], a web implementation of TMFoldRec [70]. Upon topology prediction, systematic sequence to structure

alignment is performed, resulting in the selection of several templates which are ordered according to energy and reliability. Rosetta has also been widely applied to MP prediction [71]. The main improvement over soluble protein prediction was the implementation of a new membrane-specific version of the original Rosetta energy function, which considers the membrane environment as an additional variable next to amino acid identity, inter-residue distances and density [71]. Rosetta has been used to reveal important structural details in voltage sensor MPs, namely the K(v)1.2 and KvAP channels [72], and gain insight into voltage-dependent gating [73]. Recently, RosettaMP was developed as a general framework for membrane protein modeling, featuring modeling tools developed in the past few years [74].

ML methods are becoming highly popular in biological fields [75], and TM structure prediction is no exception. ML is a general-purpose approach defined as the automatic extraction of information from large amounts of data by efficient algorithms, to discover patterns and correlations and build predictive models. ML involves the creation of algorithms that improve their own performance when undertaking a certain task based on their own experience. They should be fully automatic and *off-the-shelf* methods that process the available data and maximize a problem-dependent performance criterion [76]. They aim to be statistically consistent, computationally efficient, and simple to implement and interpret. The choice of a ML algorithm for a specific problem should be made in light of its characteristics, deep familiarity with the theoretical foundations of the field, data source and prediction performance [77].

PsiPred [78] is a broadly utilized platform for secondary structure prediction that utilizes Position Specific Scoring Matrix (PSSMs) as inputs to an Artificial Neural Network (ANN) approach. However, this is hardly specific for MPs. Adding hydrophobicity scales to the prediction of secondary structures, should yield better results [79]. Initially, the utilized scales were focused on ranking single amino acids or small peptides [9]. More recent advances in hydrophobicity scales include the energy of amino acids in fully folded proteins, such as the hydrophobicity scale developed by White and von Heijne [80], which was shown to deliver the best results along with scales such as the Unified Hydrophobicity Scale [9]. Other possible features to take into account are the regions of the protein that actually face the membrane, cytosolic or extracellular sides, and which are the motifs responsible for interactions, whether they are membrane-protein interactions or secondary structure-secondary structure interactions [81].

MP topology prediction by ML techniques has benefited from the recent Big Data explosion. When applied to the study of MP, these techniques progressively increase in complexity and difficulty, beginning with the prediction of secondary structure, followed by that of 3D domains, including super-secondary structures. They are also used to predict the TM protein segments, nowadays often making use of direct residue coevolution features, which are then translated into residue-residue contacts [82,83], or to characterize proteins as MPs from genomic sequences – such as the work of Gromiha and Suwa for outer membrane proteins (OMPs) [84]. Predicting which MP regions are α-helixes has also benefited from ML methods [85]. Martelli et al. [86] developed in 2003 ENSEMBLE, a topological predictor for all-α MPs. Their method can determine how far is an α-helix in the membrane through a combination of two Hidden Markov Models (HMM) and one ANN in what is known as an ensemble approach that combines both methods. Although the individual methods were also tested separately on a dataset of 59 known, well resolved MPs, the final approach is an average of the scores coming from the different methods. With a slight improvement in predictive power, TMSEG [87] was developed in 2016. It first characterizes the protein as soluble or transmembrane, and in the latter case predicts its topology. Based on sequences and PSSM scores, this method utilizes two Random Forests (RFs) and one Neural Network (NN) in a four step process consisting of: i) a per-residue analysis; ii) a per-protein analysis; iii) a refinement method and, finally, iv) a topology prediction step. The first three steps are considered by the authors as filters that categorize the protein as soluble or TM, while the last is the actual topology predictor. Whereas the previously addressed methods focus on α-helices, *BETAWARE* [88] has a bolder task: Detection of TM β-barrel proteins and topology assignment. It combines an extreme learning machine, a type of NN with a single hidden layer that assigns the weights between input and hidden layers in a single training step, as one would do with a linear model, [89] to decide whether a protein is a TM β-barrel or not. If this is the case, a Grammatical Restrained Hidden Conditional Random Field [90] (GRHCRF) model predicts the topology of the protein. This uses a type of Conditional Random Field [91] (CRF) model. Such models have become popular in biological sequence analysis as methods that can make stepwise predictions of MP structure, just as HMMs. Other tools have also been developed for the prediction of β-barrels and their membrane span, such as the one by Ou et al. [92], which uses sequence information and a Radial Basis Function (RBF) network – an artificial neural network that uses radial basis functions as activating functions for individual neurons [93]. TOPCONS [94], another method for the identification of MP topology works both as a meta-server and as a ML approach. It combines the topology profiles from several other MP prediction platforms and feeds these outputs to a HMM, which creates a final output. Its key aspect, compared to other methods, is its ability to distinguish MPs from signaling peptides. Merging several of the previously mentioned methods, MemPype [95] is a Python pipeline for the prediction of signal peptides, based on homology and other computational methods.

A few methods managed to combine various sources of information to predict TM α-helices and α-helical bundles, as well as β-barrels. OCTOPUS [96] may be one of the most complex ML approaches for TM α-helical spans, as it combines four different ANNs – membrane, interface, loops and globular residues – through a HMM. HMMs consist of a set of sequential states, whose progress is dependent on the confirmation of the current state [97]. TMs were also predicted using Support-Vector-Machine (SVM) - Memsat-SVM [98]. BOCTOPUS [99], developed by the same group as OCTOPUS, allows to predict β-barrels. BOCTOPUS combines local predictions through SVMs and a HMM to combine all local SVM predictions.

Another determining factor when predicting MP structure is the prediction of membrane-facing regions and interaction motifs. LIPS [100] for example takes into account hydrophobicity scales and is able to predict the orientation of helices and amino acids. PRIMSIPLR [101] is another method that operates differently, namely through the training of the algorithm with several helical MPs of known structure. To predict interaction motifs, some approaches for helical MPs have been created, such as MemBrain [102] – a ML trained server that is, so far, possibly one of the best ranking systems of its kind.

Some peculiar MP secondary structures are still not correctly predicted when applying these methods. According to Leman et al. [9], these are usually re-entrant helices (sometimes mentioned as "P-loops"), half-helices (α-helices that do not span the entire membrane), amphipathic helices (α-helices that lie on the surface of the membrane, with both hydrophobic and hydrophilic regions), TM helix kinks and β-barrels whose composition is made up by more than one chain. A good example of software able to deal with more complex proteins is TMkink [103], whose main aim is to determine helix kinks.

4.2. Membrane proteins databases

Genome-wide annotation, relating gene expression with protein expression and activity, is a useful resource that underlies many computational studies. Regarding the original human genome sequencing project, it was estimated that around 20% of the genes coded for membrane proteins [104]. For MPs it becomes relevant to annotate not only the overall protein associated-genome, but also the secondary and tertiary structures formed, so that this information can be further used for application of other bioinformatics methods [105]. THGS (Transmembrane Helices in Genome Sequences) [106] is an example of

structural annotation for membrane proteins, focusing on transmembrane helices. The Membrane Protein Hub, for instance, used sequence annotation to predict human α-helical transmembrane proteome [107]. Regarding β-barrels MPs, OMPdb (Outer Membrane Protein database) [108] focuses on outer-membrane proteins from Gram-negative bacteria, registering information on both structure and function, further organizing the entries by families, regarding evolutionary information.

As previously discussed, both the experimental and computational structural determination of MPs present unique challenges that require specialized approaches. Additionally, the community has a need for databases that specialize in the collection, validation, analysis and publication of 3D MP structures. Some of these databases are of general interest as their purpose is curation and storage of 3D structures. Examples include MPs of known 3D STRUCture (mpstruc) and MPNMR (MPs of known structure determined by NMR). MPs Molecular Dynamics (MemProtMD) [109] is a meta-database which emphasizes analysis and presents the results of MD simulations of some of the proteins in mpstruc. Finally, there are even databases of specific interest such as the G-Protein Coupled Receptor Database (GPCRdb) [110], which stores and analyses data pertinent to GPCRs.

Mpstruc is the largest repository of unique MP structures. As of June 29th 2017 it contains 702 unique entries. It is a manually curated database and its entries serve as the foundation for the identification of MPs in the PDB. Entries are categorized by their biological features, their novelty and experimental conditions under which the structure they represent were determined. More specifically, biological classification takes into account he structural elements of each protein and the way they are inserted in the membrane, in order to categorize the protein into one of the following categories: i) Monotopic MPs; ii) β-Barrel TM proteins and iii) α-Helical TM proteins. As implied by its the name, proteins in categories ii) and iii) differ from those in category i) since ii) and iii) are embedded in the membrane whereas i) has only one domain inserted in the membrane. All proteins are grouped in protein families regardless of the category they belong to. As for the novelty and experimental conditions of an mpstruc entry, they are distributed across three categories: i) unique/primary entries, ii) redundant/secondary entries, and iii) related entries. A unique entry represents a protein that has been identified as a MP by the curators of the database and no single structure of this MP is available. The unique entries can contain duplicates when multiple structures for the same protein from different organisms are available. If, on the other hand, this is not the first structure that is solved for a protein family and organism, the entry will be characterized as secondary relative to the unique structure. These entries always represent proteins from the same organism as an unique entry, and they often showcase the structure of the unique protein under different experimental conditions or with different binding partners. Finally, the related entries are always associated with a unique or secondary entry and represent the structure of the associated entry under different experimental conditions, often originating from the same publication. Fig. 2 illustrates the breakdown of the various types of entries and of the proteins they refer to. Regardless of the entry type, helical proteins are the most popular type. As mentioned, the unique entries set consists of 702 proteins, 516 of which are α-Helical, 139 β-Barrel and 47 Monotopic.

PDBTM [111] is an automatically constructed MP database. Several protein structures were retrieved from the PDB and subjected to the TMDET algorithm [112] for membrane prediction. This information was used to build PDBTM, which comprises to date 3.227 proteins – 2.848 α-proteins and 366 β-proteins (accessed in June 29th of 2017) – and acts as a basic repository of membrane proteins [111]. ExTopoDB [113] is a comprehensive database with various information on several trans-MPs topology. The topology of a protein is the simplified description of the primary structure of a given secondary structure, as well as the latter's relative spatial position and orientations [114]. Another related database is the Orientations of Proteins in Membrane (OPM) database, featuring all unique MPs or in the case of MP families with multiple solved structures, one representative structure per family, some peripheral proteins and membrane-active peptides, with adjustable membrane thickness [115]. The MemProtMD database is an analytical pipeline which embeds MPs in lipid bilayer simulations and stores and publishes the results along with instructions that allow independent researchers to set up their own simulations. The pipeline consists of identifying MPs in the PDB, converting the coordinates into a coarse-grained model (see the Molecular Dynamics of Membrane Proteins subsection) adding membrane lipids and solvent to the system and simulating it for

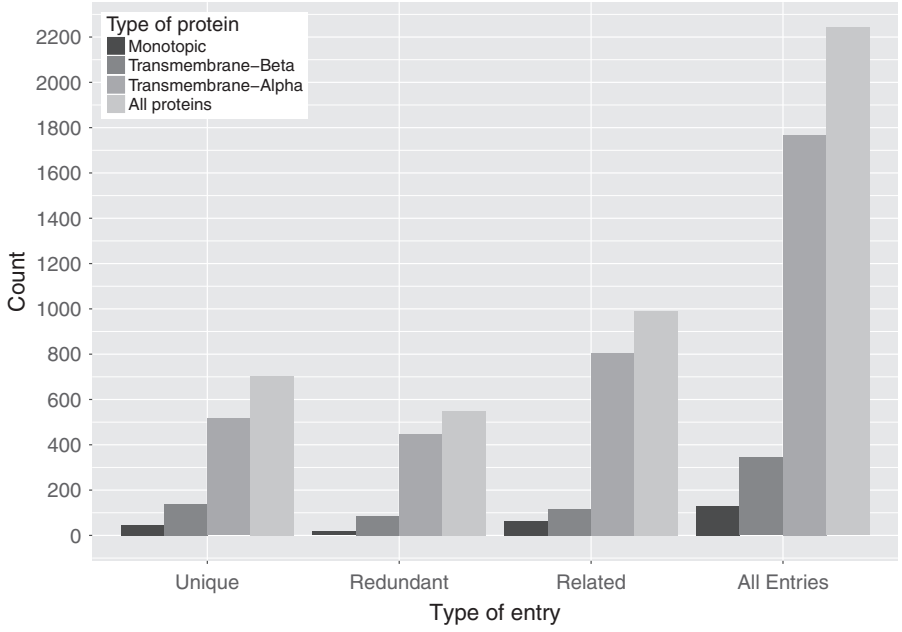

Fig. 2. Number of entries in mpstruc (http://blanco.biomol.uci.edu/mpstruc/) broken down by protein type and grouped by entry type. The α-Helical proteins are, by far, the most populated group (~78%), followed by the β-Barrel ones (~16%) and the Monotopic ones at the bottom (~6%). The most common type of entry are the related ones (~44%), followed by the uniques (~31%) and the secondary (~25%).

100 ns allowing the membrane to self-assemble in a bilayer, and followed by a 900 ns production run in order to study the membrane dynamics for a total simulation time of 1 μs. The systems are then converted back into atomistic detail and are analyzed. All TM proteins found in mpstruc can be found in MemProtMD as well. There are also specialized databases that gather information regarding important biological targets such as GPCRs and Transporters. GPCRdb [110a], GPCR Consortium and GPCR Network have made important contributions to the cataloguing of known data on GPCRs and their understanding. In particular, GPCR Consortium aggregates researchers and data on the matter, in partnerships with both GPCRdb and GCPR Network. GPCRdb provides information on specific GPCRs, their structures, known mutations, homologues, ligands and phylogenetic relationships. It contains 46 unique crystalized entries (statistics of June 29th 2017), varying in resolution between 1.7 and 7.7 Å, covering various receptor types as shown in Fig. 3. Table SI-1 contains all the aforementioned databases with a short description and their website.

GPCRdb also offers other services:

i) GPCRM: Generation of GPCR models using the available template structures;
ii) scPDB: using structure-based approaches, PDB structures are used to identify binding sites on the GPCR suitable for drug-like ligand docking;
iii) GPCR-SSFE: A GPCR – Sequence-Structure-Feature-Extractor – provides template suggestions and homology models of the helical regions of >5000 family A GPCRs;
iv) GOMoDo: Automatic homology modeling and ligand docking of GPCR receptors;
v) GPCR-ModSim: Generation of homology-based 3D models using GPCR sequences and further model equilibration using MD simulations.

Linked to GPCRdb, GPCR-OKB [116](G-Protein Coupled Receptor Oligomerization Knowledge Base) provides detailed information on GPCR oligomerization. GPCR-I-TASSER [117] predicts GPCR structure with the aid of a software program (LOMETS [118]) responsible for identifying a suitable template. It assembles full-length models through a template-based fragment approach.

The Transporter Classification Database (TCDB) is another important database that gathers functional and phylogenetic information on transporters. It features 943 transporter structures from well-defined biological systems (Fig. 4, accessed in June 29th of 2017). TransportDB [119] is another example of a comprehensive transporter database; however, it has not been updated since 2013.

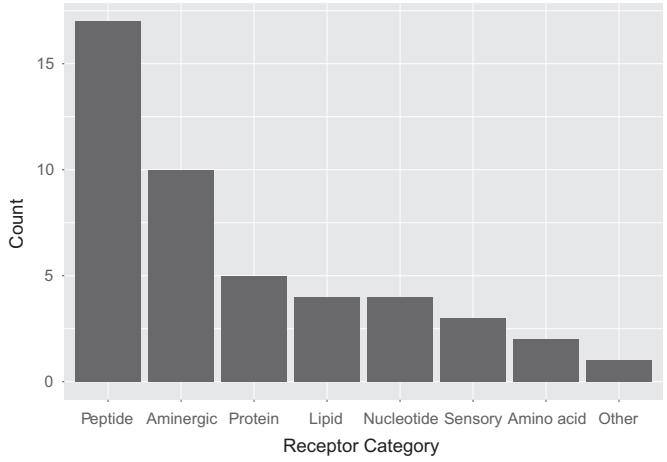

Fig. 3. GPCRs available at GPCRdb (at June 29th 2017).

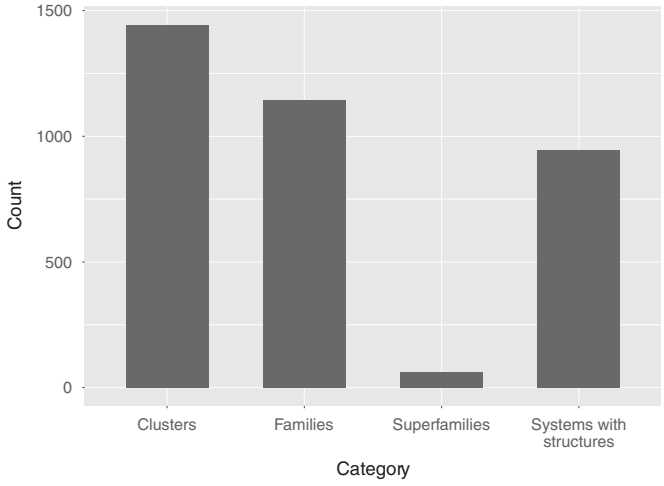

Fig. 4. TCDB information, regarding transporters (at February 29th of 2017), in http://www.tcdb.org/.

4.3. Molecular dynamics of membrane proteins

Molecular dynamics (MD) simulations have benefited greatly from advances in recent years and proven a valuable resource for the study of MPs. Various force fields have been developed for membranes such as: CHARMM36 [120] (implemented in the Chemistry and HARvard Macromolecular Mechanics – CHARMM – software) and AMBER Lipid14 (implemented the in Assisted Model Building with Energy Refinement – AMBER – software) [121]. The main difference when compared to other force fields is the number of supported acyl chains, head groups and lipids. For example, Lipid14 features 12 acyl chains, 9 head groups and 8 lipids, which enables the creation of a wide array of molecules. Although constructing these systems might seem challenging, some platforms provide simple solutions, such as the CHARMM-GUI [122], QwikMD [123] and High Throughput Molecular Dynamics (HTMD) [75f]. CHARMM-GUI is an online webserver (http://www.charmm-gui.org/), which takes input from a number of MD software programs such as AMBER [121] or GROMACS [124] and constructs a simulation-ready system with a membrane-embedded protein. QwikMD and HTMD are standalone software programs which offer accessible interfaces for both beginners and experts to perform several MD tasks, such as the insertion of a protein in the membrane, among others. When performing MD simulations of membrane-inserted proteins, protein-lipid interactions are an important factor to consider as they can be greatly affected and should be carefully modeled. For example, hydrophobic mismatch [125] may produce profound modifications in the system: When the bilayer is thicker than the hydrophobic region of a TM protein, hydrophobic residues exposure to water leads to an energy penalty, which causes an increase in the membrane thickness and a slight tilt of the protein. When the opposite occurs, the membrane is compressed to prevent interaction between hydrophilic residues and the membrane.

The membrane can be represented in MD in two ways: i) explicitly (all-atom or coarse-grained (CG)) and ii) implicitly. All-atom models represent the membrane explicitly with as little approximations as possible. One of the first models was gramicidin in a DMPC bilayer back in 1994 [126]. The system comprised 4390 atoms and the simulation was run for 0.5 ns. The access to powerful computational resources and advanced sampling techniques allow now longer MD simulations for membrane protein systems (reaching the microsecond timescale in some cases) [75l,127]. Such long simulations allow observing and characterizing interesting mechanistic phenomena. Just to name a few, Ogata et al. [128] were able to determine that transfer of oxygen,

water and protons across the thylakoid membrane by the photosystem II is actually done through different pathways [129]. Allostery is an important phenomenon which gives rise to complex intramolecular mechanisms and is often used as an information transfer mechanism across the membrane. MD can also be used to study relevant phenomena and aspects of some techniques, such as x-ray crystallography: For example, a 2016 study showed that small detergents at high concentration can bind to important protein sites, something which does not happen with larger detergents [130]. This was made possible by making use of a particular technique, accelerated MD, which allows for better search of the conformational space [131]. A similar advanced sampling method was applied to determine binding site characteristics and the differential interaction of agonists, partial agonists and antagonists in the M3 muscarinic receptor [132]. To reduce the computational expense associated with all-atom models, Markov models can be integrated. Such models assume that only the current state is important for future states. This requires an initial dimensionality reduction (which eliminates redundant data on atom coordinates) and grouping similar kinetic modes [133]. The Markov model itself is then able to project longer timescales from shorter simulations [134]. This combination allowed Razavi et al. to determine the most likely sodium release pathways for the human dopamine transporter (hDAT) triggered by hydration of the Na2 (at sodium binding sites of hDAT) [135]. As advanced sampling techniques in MD have become very relevant for the study of membrane proteins we will review a few here (for a more thorough description refer to Mori et al. [75l]). Replica exchange MD (REMD) is an advanced increased sampling technique in which several replicas of the system are sampled in parallel, each with a different temperature or Hamiltonian. Exchange between different replicas can take place if the Metropolis-Hasting criterion is satisfied [136]. REMD has been applied to study the insertion mechanism of α-helical peptides, suggesting that folding is more likely to happen inside the membrane than on the membrane/water interface as previously thought [137]. Steered MD (SMD), a technique which uses external forces to drive the simulation, was used to determine the gating mechanism of MscL – a mechanosensitive channel of large conductance – by applying external forces of 35–70 pN to residues near the membrane-water interface [138]. Metadynamics is another advanced sampling technique used to calculate several state functions – such as free energy – and it works essentially by adding positive Gaussian potential to already searched spaces. By doing so, it prevents the system from sampling previously sampled conformations, allowing a faster and more efficient description of the full energetic landscape [139]. One of its application to MP systems was the characterization of ion permeation through ion channels [140].

Despite the improvement observed in both harware and software for explicit MD simulations of MPs, the time- and length-scales are still prohibitive for a large number of systems. One approach that has been gathering interest in recent years is the use of CG models. These are able to represent lipids (and also proteins) by approximating ensembles of atoms as single particles. MARTINI [141] is one of the best known CG force fields. It uses a "four-to-one mapping", which considers four heavy atoms and their respective hydrogen atoms as a single particle. By doing so, it greatly reduces the computational cost, allowing for longer MD simulations. However, MARTINI fixes protein secondary structure, which prevents this method from being able to predict protein structure or detect conformational changes. As such, it can preferenciably be used as a tool for the prediction of the oligomerization and interactions of MPs. An interesting case study on the effect of cell membrane's curvature on protein and lipid dynamics for the F_1F_o-ATP synthase revealed that the energetic cost of membrane deformation can be reduced through side-by-side association of multiple dimers [142]. Another example was the study of β-barrel outer MPs turnover in *E. coli*, which led to the conclusion that an accumulation of proteins occurs in the *E. coli* poles due to a slowdown of their diffusion through the formation of homologous and heterologous assemblies, mostly mediated by aromatic residues [143]. An interesting study on the formation of Ras nanoclusters in extensive CG simulations revealed that these nanoclusters influence the membrane curvature and that their clustering depends on palmitoylation and farnesylation [144]. CG simulations were also performed to observe Ras clusters and their effect on the formation of cholesterol-rich domains in the membrane [145]. All-atom and CG models can also be combined as, for example, in the study of the binding site of phosphatidylinositol 4,5-biphosphate (PIP_2) in Kir2.2: PIP_2 was found to have conserved binding sites in receptor tyrosine kinases, which are constituted by basic amino acids responsible for the clustering of PIP_2 in juxtamembrane regions [146].

Next to the explicit models discussed above, the membrane can also be represented implicitly, by some continuum model as a mathematical function, with the remaining of the system – the proteins – represented at an all–atom level. This allows to greatly increase the computational efficiency. When using an implicit solvation model, the membrane is modeled based on its solvent free energy, which depends on the free energies of (i) the electrostatic solute-solute and solute–solvent interactions, (ii) the cost of cavity formation for protein insertion in the membrane, and (iii) the solute-solvent vdW interactions. Simulating membranes implicitly can be done through equations such as the Poisson-Boltzmann (PB) eq. [147]. This equation is able to describe the electrochemical potential of a solution perpendicular to a charged surface (a membrane), which can be hard to implement in MD simulations due to its high computational costs [148]. As such, approximations such as the Generalized Born (GB) method have been considered to reproduce the PB model [149]. GB with a simple SWitching (GBSW) [150] is an implementation of a GB method in the CHARMM software [120, 151]. It considers identical dielectric constants for the membrane and the protein (represented as the solute), and a smoothing function that models membrane-water and solute-water interactions. The membrane is represented as a solvent-inaccessible infinite planar low-dielectric slab. A similar example of a GB model for membranes is the Heterogenous Dielectric Generalized Born (HDGB) method, which considers the membrane as several layers with different dielectric constants [152].

Initially these models represented the protein as cylinders or point particles inside the membrane [153], which is a rather poor representation of a MP. Since then several other aspects have been incorporated, namely considering individual α-helices, the hydrophobic effect, the large free energy penalty associated with peptide desolvation [154], hydrodynamic effects (as reviewed on [155]), or the macroscopic strain energy of the membrane [156]. Different mathematical models have been implemented, including the already mentioned PB model [157], with a wide range of applications, such as studying the interaction of peptides and ions with the membrane [158]. For example, Argudo et al. [153c] have used a model that takes into account surface area, mean curvature, Gaussian curvature, preferred curvature, and bending and Gaussian moduli [159] of the membrane as well as its compression [160] and lipid tilt. The latter considers that not all lipids are perfectly oriented according to the bilayer's normal vector [161]. The validity of implicit models has been assessed by comparing them with atomistic models. One study that looked into the epidermal growth factor receptor dimerization in the membrane concluded that atomistic Monte Carlo simulations yield dimerization rates that could differ by two orders of magnitude compared to simple partial differential equations [162]. Grand Cannonical Monte Carlo (GCMC), a technique yielding promising results when used in MP simulations, allows to simulate systems at constant chemical potential, volume and temperature (μVT). A 2014 study combining GCMC with other MD techniques shed light on the interaction between single-stranded DNA homopolymers and the α-hemolysin pore, concluding that cations act as primary charge carriers through α-hemolysin pore, and that polymer passage through the channel might lead to a flickering gate behaviour [163]. Similar techniques have been used to study ion permeation through an α-hemolysin channel in a comparative study with a continuum model showing that GCMC seemed to give better results than a continuum model

[164]. Other studies have compared MD simulations to continuum models for electropores [165] and CG models to continuum models for lipid bilayer fusion pores [166], both with not so favourable results for the continuum models, suggesting that hybrid approaches (combining both methods) might lead to better results.

4.4. Prediction of interactions between a membrane protein and its soluble partner

Membrane proteins play an essential role in key cellular functions by interacting with several different molecules and particles, ranging from photons to macromolecules such as other proteins. Typically, MPs interact with soluble partners. The modeling of those interactions becomes easier if the involved binding interfaces can be identified. Many approaches have been developed over the years for this purpose. Sequence-based methods are gaining an increasing importance for MPs. They compare sequence related proteins and identify the surface residues and even contacts that are significantly conserved, as conservation is overall correlated with functional importance [167]. Other approaches such as PS-HomPPI and NPS-HomPPI [168] predict interfaces using data on interfacial residues belonging to homologous proteins. ML methods are also used and can predict protein interfaces based on either structure, sequence or both. There are meta-servers as well, which, instead of implementing their own method, provide an interface that gathers structural and/or sequence features from several other servers and software programs and analyzes them to identify protein interfaces. An example is CPORT [169], which gathers information from six different servers and combines this information, provides results better than those ones presented by the individual servers. It is important to mention that most methods have been developed for soluble proteins and might have limitation when it comes to binding site prediction for MPs.

ProMate [170] is an interesting example of a structure-based method, which uses several features such as secondary structure, length of non-secondary structure protein regions and pairwise amino acid residues distribution to calculate an interface propensity value for each residue. Part of the development of ProMate involved the elimination of redundant or highly correlated features, which reduces computation and search space. PPIPP [171] is a good example of a sequence based method by using propensity scores based on the presence of a given residue compared to any other residue at the interface. To solve the lack of partner information, the model was trained by comparing residues in intermolecular protein-protein interface with intra-protein contacts. PPIPP is built on 24 ANN and returns the average score as final score, using PSSMs as one of its main features. PAIRpred [172] is an hybrid approach, using both sequence and structure-based features: the structure-based features consist of relative Surface Accessible Surface Area (SASA), residue depth, half sphere amino acid composition and a protrusion index, while the sequence-based features are based on PSSMs and predicted relative accessible surface area. All these are combined through a SVM to predict protein-protein interactions.

An important note made by Xue et al. in their 2015 review [173] is that partner information is very valuable for protein interface prediction, which is often overlooked. A comparison of the results obtained through PPIPP and PAIRpred – with partner information – with the ones from PSIVER [174] (sequence-based) and SPPIDER [175] (structure-based) proved that partner information greatly improves the predictions made. Evolutionary conservation of residues, and co-evolution [176], is also a growingly utilized feature in protein interface prediction [177]. It is based on the conservation of amino acids at the interfaces, as it relies on scoring residues or residues pairs, depending many times on Multiple Sequence Alignment (MSA) methods [82–83,178].

4.5. Prediction of interactions between membrane proteins

It is well known that membrane proteins can associate. Several methods have been developed to predict their interfaces. In the case of β-barrel MPs, the Transmembrane Strand Interaction Potential (TMSIP) has been recently used to predict oligomer interfaces [179]. By analyzing each strand independently, information can be gathered that hints at the location of an interface – whether it is with other β-barrels or with other structural elements. A higher TMSIP in a β-barrel OMP region indicates a high chance that there lies a protein interface that needs to be stabilized in order to lower the TMSIP for that region [180].

BTMX (β-barrel TransMembrane eXposure) [181] is a server dedicated to the prediction of β-barrel interfaces by predicting the exposure of known TM residues. It uses a simple feature based on the orientation of hydrophobic residues. Table 1 summarizes some of the most commonly used features in the literature for contact prediction of proteins.

4.6. Modeling of membrane protein complexes by docking

The determination of the 3D structure of protein complexes is of the utmost importance for molecular biology and drug discovery. Docking, a computational method that assembles the biologically relevant complex from its known constituents, is the method of choice for this task [197]. It involves “search and scoring”, both of which will be addressed in the following. First, during the search phase, a large number of candidate association models must be generated. Any available biological information on the location of the interface can in principle be used to direct this step, which should generate a set of poses as close as possible to the native structure. Search methods can be split into three different approaches [198]: i) rigid docking, where both proteins are considered as rigid structures; ii) semi-flexible docking, where only one of the components (usually the smallest) or limited regions in both components are flexible; and iii) flexible docking, where both proteins are allowed to be flexible. As one would expect, the latter is more demanding when it comes to computational resources, but will probably yield the best results as it will be prepared to deal with possible conformational changes upon complex formation.

Once models (poses) have been generated, the next stage is scoring, which aims at ranking the poses and identifying the native-like models. Various metrics have been used for this purpose such as geometric complementarity, exclusion of solvent from the interface and associated entropy changes (desolvation), electrostatic and van der Waals interactions, and hydrogen bonds. A current limitation of existing scoring functions is that they do not provide accurate identification of native-like solutions. This is particularly true for MPs, as most docking software programs were developed for quaternary structure prediction of primarily water-soluble proteins. Membrane proteins are surrounded by the lipidic environment of the membrane, and consequently commonly used scoring functions are not suited to deal with these systems. Moreover, search algorithms typically do not consider the membrane, which does provide additional restraints on the possible orientations of the components of a complex.

Despite this, some methods have been specifically developed for MP docking. For single-spanning TMHs, TMDOCK [75d] is one of the most recent one, using an all-atom model for TMHs and inserting them into the membrane following the folding of membrane-associated peptides (FMAP) [75m] protocol [199]. By using templates for right- and left-handed TMH, several models are created by parallel translation of the helices, followed by energy minimization of both single TMHs and the TMH dimer. The scoring function estimates the free energy of helix association based on a combination of van der Waals interactions, H-bonds, solvation energy, changes in entropy upon helix association and helix dipole electrostatic interactions. PREDDIMER [200] is one of the oldest TMH docking software and can be described as a three-step process consisting of: (i) TMH structure prediction requiring user input for the protein sequence, pH and relative orientation of the helices in a dimer (parallel/antiparallel), (ii) TMH docking and, (iii) characterization of the TMH dimer with hydrophobic properties and contact regions. CATM [201] is another method which, based on observed recurring patterns in TMH dimer interfaces, predicts homodimeric

Table 1
Features used on several servers and methods, with the respective biological reasoning and the ML algorithm applied to program and/or train the predictor. The target for each feature is mentioned between parentheses.

Feature	Remark	Method	Reference
Amino acid occurrence in Trans Membrane Helices (TMH) segments and their amino acid frequency in proteins with different size windows (predicting TMH).	Using solely the amino acid occurrence leads to a lack of prediction accuracy. The variation of windows in order to find the perfect size and incorporating the amino acid frequency improves it.	Modified genetic algorithm	[182]
Propensity values for individual amino acids according to the Beuming and Weinstein (BW) scale and evolutionary conservation (EC) of amino acids with different window sizes (predicting the burial status of TMH proteins).	BW scales had been already used, however by adding EC to the function, a better scoring function can be made.	Support vector classifier; Fisher's indexes were utilized to assess which of the windows elements were more significant for what particular features (amino acid presence or EC).	[183]
Amino acid physicochemical properties (charge, polarity, aromaticity, size, electronic character). Each amino acid was filed in a category for each property. Compositional index (amino acid composition for TMH) (predicting TMH).	The physicochemical properties and composition index are input to different support vector decomposition machines and these were trained together in a weighted random forest. The hybrid method proved to be more efficient.	Singular Value Decomposition (reduces dimensions with strong correlation and reduced signal to noise ratio).	[184]
Normalized accessible surface area (ASA values for residues from MSMS/reference ASA values for Gly-residue-Gly)(solvent accessibility of TM residues).	ASA from membrane and non-membrane segments are different.	PSI-BLAST profiles were used on the non-redundant NCBI DB. ε-insensitive support regression vector (defining the loss function that ignores errors – values over a given distance from the SVR).	[185]
PSSM for a 9-element window centered on the target residue. 6-letter exchange group representation (20 amino acids are filed into 6 categories – letters) (identification of TMH segments).	With PSSM, patterns are recognized through a 20*n matrix display in a n-element window; amino-acids can be classified into 6 classes (which reduces the dimension, making it easier to compute).	Particle swarm optimization to optimize the weights of the different residues and 6-letters; Fuzzy SVM for machine learning.	[186]
Relative solvent accessibility, secondary structure, relative solvent accessibility, multiple alignment, KD and WW hydropathy profiles (TM domains).	Hydropathy profiles led to the confusion of MPs with globular proteins and were later excluded.	Neural networks. The main changes done were with the number of hidden layer nodes, sliding window, training protocols. In the training set globular proteins and signal peptides were also utilized (reduces confusion with these moieties in the final classifier).	[187]
Windowed PSSM profiles, position of the residue in the TMH (cytosolic side, hydrophobic, extracellular side), orientation of the sidechain (with LIPS [100]), sequence distance between residues and residue coevolution, protein length and number of helices (helix-helix contacts and interacting helices).	These features were considered as out-of-context features (apparently would not affect the predictions). As such, several features for globular and soluble proteins were taken into account (PSSM profiles, sequence distance, coevolution and protein length) and some particular MP features were added. The protein length was considered to be solely the length of the TMHs.	ANN with several Boolean vectors describing all of the features.	[188]
Solvent inaccessible residues (buried polar residues) - entirely conserved within protein families and superfamilies - and hydrogen bonded to an equivalent main chain atom in each family member (helix capping).	Conservation of polar residues and the hydrogen-bond interactions that they form play an important role in maintaining protein structure, by promoting strong restraints on amino acid substitutions during divergent protein families and superfamilies evolution.	The residue propensity to form hydrogen bonds to main chain atoms was calculated as: $P_{arch(x)} = \frac{\left(\frac{n_{arch(x)}}{N_{(x)}}\right)}{\left(\frac{n_{arch(total)}}{N_{(total)}}\right)}$ Where P_{arch} is the particular architectural context, $n_{arch}(x)$ is the number of residues forming hydrogen bonds to main chain atoms in a P_{arch}, $N(x)$ is the number of type x residues in the total dataset and N (total) is the total number of residues in the dataset of 131 families.	[189]
Salt Bridges under varying ion (Ca(II)) concentrations with proteins, namely membrane-anchored proteins (SNAP25) – concentration-dependent ion-induced protein oligomerization. Preferred interaction partners of Ca(II) ions.	Ion-selective effects concerning channel permeability, enzyme activity and protein oligomerization have shown to be ubiquitous, concerning ion-protein interactions, however, the underlying molecular ion-binding patterns and the effect of ions on proteins in cellular multicomponent environment have not been resolved.	Radial distribution functions (RDFs) between Ca(II) and the carbon atoms of the carboxylate groups of Asp, Glu and the C-terminus; oxygen atoms of the hydroxyl groups of Ser and Thr and nitrogen atoms of the side chain amino groups of Gln, Arg and Lys. The RDFs yielded increased probabilities for locations ranging from 0.25 (carbonyl oxygens) to 0.5 nm (amino nitrogen).	[190]
Mean value of burial propensity within a 19 residues sliding windows (B-value of the residue) – defined as the average fraction of the buried, solvent-accessible surface area relative to the total solvent-accessible surface area (proposed as hydrophobicity scales).	Topology prediction of Transmembrane Helices (TMHs); most TM proteins are TMH, strongly associated to the membrane by hydrophobic interactions.	THUMBUP – TMH proteins' topology predictor using Hidden-Markov-model based in UMDHMM (University of Maryland HMM) – UMDHMM$^{TM^{HP}}$ (UMDHMM for topology prediction of TMH proteins) assigning one of five states to each residue	[191]
Surface Propensity (SP) – defined by the inside/outside propensity of amino acids, Surface Fraction (SF) – reflects the probability of finding a residue on the surface of the TM protein; (also, conservation, hydrophobicity, etc.).	The method was most successful in predicting residue orientation in TMHs by combining conservation and knowledge-based features such as surface propensity.	Jack-knife approach was used to determine the SFs for the SP scale; MSA (Multiple Sequence Alignment) for the prediction of the lipid-facing probability of residues in a protein sequence, assigning a cut-off value for the Probability of the residue being inside.	[192]
One dimensional Equilibrium Structural Features (ESFs) – eight one dimensional structural features: SASA (three features), Helix-helix Contact (HCC – one feature) and Backbone conformational angles (four features).	There is little explanation on why some residues are predicted as belonging to a wrong structural class or with large errors in the absolute values of these features.	Multilayer feed forward neural network, which input is the position specific scoring matrix using Leave-one-out (LOO) method to avoid overfitting.	[193]

Table 1 (*continued*)

Feature	Remark	Method	Reference
Prediction of mean secondary structure, solvent accessibility and coiled-coil regions from multiple sequence alignments or amino acid sequences.	The use of different multiple sequence alignments of the same protein proves to improve greatly the results.	Neural network-based prediction. Training was done with different multiple sequence alignments for the same proteins for several different proteins.	[194]
Functional impact of amino acid substitution from proteins in the Uniprot or NCBI RefSeq databases and the residue substitution.	Functional impact is based on evolutionary conservation of the amino acid being replaced in close homologues.	Several numerical estimates for amino acid residue substitution. These are calculated through the difference in entropy, which is use as a measure of the impact of a mutation.	[195]
Evolutionary information encoded in multiple sequence alignments.	Traditional sequence-based prediction methods might not be sufficient for structure and function prediction.	Mutual information, chi-square statistic, Pearson Correlation, Joint Entropy.	[83a]
Evolutionary residues couplings.	3D complex structure determination.	Interacting proteins' sequence pairing and comparison.	[196]

interfaces – it considers important amino acid residues motifs, namely Gly-Ala-Ser (GAS_{right}) [202]. Its surprisingly simple scoring function mainly consists of hydrogen bonding and van der Waals interactions. Evolutionary-based TMH docking can also be performed by EFDOCK-TM that combines several approaches and features, such as EVFold [203] (evolutionary features), LIPS [100] (selection of interacting surface region for TMHs) and OCTOPUS [96] (determining intramembrane protein segments).

Most methods described in Sections 4.4 and 4.5 are summarized in Table SI-2, which contains a short description of the method as well as the easiest way to access it.

5. Case studies

In this section, we will limit ourselves to GPCRs, transporters and ion channels which constitute the main target of many computational/experimental drug targeting strategies.

5.1. G-protein-coupled receptors

GPCRs belong to one of the largest superfamilies of membrane associated proteins with the most diverse functions [204]. GPCRs share a typical pattern consisting of seven TM helixes (TM1–7) and similar intracellular binding partners. Three High Variability Regions (HVRs) have been identified between TM5 and TM6 and at the N- and C-terminal regions [205] (Fig. 5 A) [206]. Even though GPCRs share high structural similarity, their ligands can range from a photon to a protein [207]. GPCRs can receive distinct stimuli, having roles on metabolic, neuronal, hormonal and immunological functions, as well as in cell growth and cell death [208]. Apart from their ligands (e.g. G-proteins, arrestins and GPCR-interacting proteins (GIPs), membrane-inserted GPCR-binding proteins [209]) the lipid membrane environment also has an active role in modulating GPCR structure and function. For example, interaction with cholesterol significantly changes GPCRs conformational flexibility [210] and modulates their interactions. As such, it was suggested that rather than "binding sites", GPCRs, many times, have "high occupancy sites", when associated to these cholesterol "hot-spots" in the membrane. Constitutive internalization of GPCRs, a crucial cellular function responsible for receptor regulation, is regulated by GPCR interactions and can be clathrin-dependent or clathrin-independent, stressing the large array of interactions and the versatility of GPCRs [211]. Trafficking of GPCRs, which can be agonist dependent or independent, commonly displays an important role on the signaling routes these receptors are involved in [211].

GPCRs play a central role in an enormous variety of cellular mechanisms in human physiology and disease and are the targets of 40% of

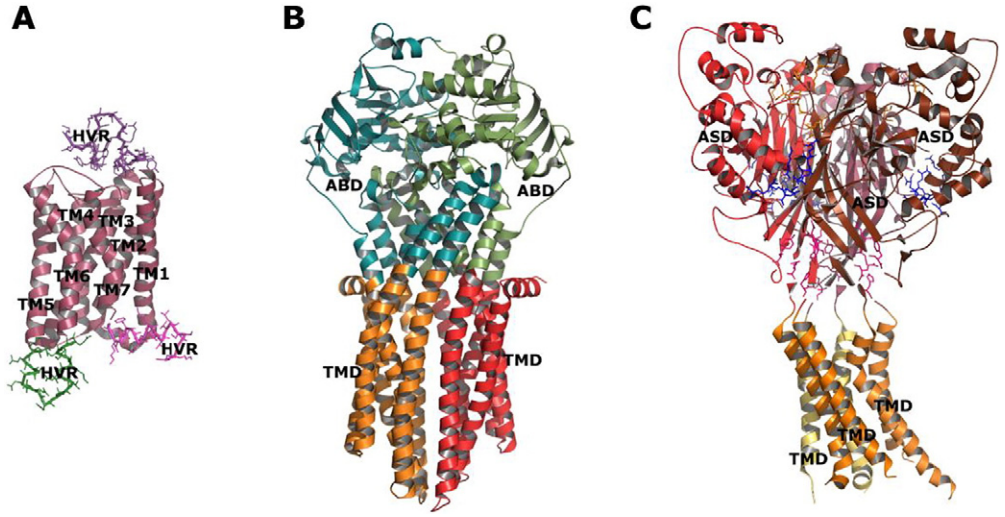

Fig. 5. Examples of membrane protein structures. (A) 2.65 Å resolution crystal structure of bull opsin (PDBID: 4J4Q [206]). The 7 transmembrane α-helical domains (TM1–7) – a key aspect of GPCRs – and the three high variability regions (HVR) are indicated. The latter correspond to the C-terminus region (pink), the extracellular loop between TM5 and TM6 (green) and the N-terminus region (purple) [206]. (B) 3.42 Å resolution crystal structure of the *E. coli* ATP-Binding Cassette (ABC) Transporter McjD (PDBid: 5EG1 [212]). Its key features are its TM and ATP-binding domains (ABD). The latter are responsible for binding and degrading ATP in order to power the drug-pumping function typical of ABC transporters. McjD is a good example of how the lipidic environment affects the stability of the MP. Through MD simulations, Mehmood et al. [212] were able to determine the synergistic effect of zwitterionic lipids on its stability, besides from McjD negative lipid-dependent function and the impact of both on the structure of McjD. (C) 1.9 Å resolution crystal structure of the chicken acid-sensing ion channel (ASIC) 1 in its closed state (PDBid: 2QTS [213]). The key aspects of this homotrimeric ion channel are its acid sensing domains (ASD) which, upon activation, lead to the opening of the channel, formed by its transmembrane domains (TMD). The ASD where named as such because of their ability to bind both protons and chloride. Each ASD has 3 regions of amino acids (colored in pink, blue and orange) which provide essential interactions for the structure of the ion channel [213].

all commercialized drug targets. As such, they are the subject of major efforts toward understanding their function and signaling selectivity [214]. New insights have been provided by recent GPCR structures in selected conformations, stabilized by a variety of ligands with pharmacologically distinct properties (agonists, inverse agonists, etc.), by nanobodies mimicking signal transducers [215], and in some cases by full heterotrimeric G-proteins (GTP-binding protein) [216,217]. Still, the molecular mechanisms connecting GPCR structure to these states and how these states are related to GPCR functional mechanisms remain undetermined.

Computational methods can be useful in understanding structure-activity relationships [218], conformational changes upon ligand binding [219], and complex mechanisms such as allostery [220]. Such methods have also been used to study of GPCR dimerization and oligomerization [221], explaining more complex GPCR signaling pathways [222], and to clarify molecular mechanisms of signaling selectivity by exploring the dynamic properties of the systems [75i,223]. Using evolutionary conservation, a GPCR-independent mechanism of trimeric G protein activation [224] was elucidated, proving the significance of common regions across these proteins and illustrating how co-evolution can elucidate complex GPCR-related mechanisms.

GPCRs comprise a large number of different receptors. Dopamine receptors (DR), for instance, have been studied under the scope of determining structural relevant elements, such as intracellular and extracellular loops (ICLs and ECLs, respectively, which have been deemed critical for their interaction with many other effectors) as well as agonist and antagonist binding [225]. When considering drug design against GPCRs, studying the differences between the active and inactive state [226] is important, as is the case for Dopamine receptors 2 and 4 (D2R and D4R, respectively) [227]. Computational studies have also been able to assess how post translational modifications (PTMs), such as phosphorylation, influence GPCR-arrestin interaction [228]. The interactions GPCRs with G-proteins is another relevant field of study since their complexes can generate several biological effects. Many of the published studies have concentrated on the rhodopsin-like family (class A) [229]. The effect of dimerization and allosteric communication on GPCR activation has also been studied computationally [230], particularly for D2R. In particular Helix 8 (HX8), a perimembrane substructure, has been identified as a key player in D2R-PDZ domain interaction [231]. Another GPCR, Serotonin-2A, a serotonin receptor, has also been studied through computational methods: A hybrid approach using experimental data (NMR and mass spectrometry, for example) and computations (MD) revealed the importance of the ICL2 substructure when considering conformational changes upon ligand binding [232]. Metabotropic Glutamate Receptors (mGluR) are yet another thoroughly studied GPCR family, since they are involved in a large array of neurologic disorders. For example, it was found that orthosteric modulation by ligands is apparently not as relevant as allosteric modulation. This led to the design of allosteric drug (both negative and positive) against mGluR [233].

5.2. Transporters

Transporter proteins are one of the most interesting and diverse groups of TM proteins. The TCDB database [234] contains about 10,000 transporter protein sequences distributed across > 1110 families. Such a high number of families has given rise to algorithms for the prediction of transporter targets using RBF networks [235], which simplifies the assignment of families to individual transporters. The ATP-binding cassette (ABC) transporters superfamily is one of the largest and most important transporter superfamilies [236], with 7 different gene families [237]. 48 human ABC transporters have been identified, with several different targets and cellular locations [238]. Their involvement in diseases such as Alzheimer's, Tangier's, Harlequin Ichthyosis and Stargardt's [239] makes them highly relevant and, as such, common drug targets [240].

ABC transporters are structurally organized in at least 4 domains: 2 nucleotide-binding domains (NBDs), also known as ATP-binding cassettes, responsible for the binding and hydrolysis of ATP, and two TMs, responsible for substrate transport and typically consisting of 6 helices, across 2 different chains [241], as shown in Fig. 5 B for the *E. coli* ABC transporter McjD [212]. Even though ABC transporters are expressed in both prokaryotes and eukaryotes, few structures are available in the PDB [242]. Pinto et al. [241] have reviewed how structure relates with drug interaction in ABC transporters and how computational methods, namely web platforms, can be used to determine drug inhibition and transport mechanisms. The major drawback seems to be the absence of a generalized method applicable to all ABC transporters since each method is custom made for a single receptor or for a small subset of receptors. Interesting strategies to further understand ABC transporter-drug interactions could be the identification of recurring features through different methods and their combination to build models to predict these interactions, or the study of transporters according to their targets, as was done in a recent study by Gromiha et al. [243], leading to the identification of mutations responsible for activity and drug response alterations.

Glutamate transporters [244] are an interesting case when discussing transporters, since various computational methods were used to understand their structure and mechanism of action, such as homology modeling and MD. The major setback is the lack of homologous structures, since the only available structure was of the archaeal glutamate transporter Glt_{Ph}, which was resolved in 2007 [245]. Functionally, it is highly similar to its prokaryote and eukaryote homologues [246]. Glt_{Ph} is a trimer, each monomer being composed by eight primary α-helices (TM1–8) and two helical hairpins (HP1 and HP2). TM3, TM6–8 and both hairpins in each monomer are responsible for the co-transport of 3 Na^+ and 1 aspartate. Most MD simulations have been conducted to describe this transport mechanism [247]. Homology modeling studies were able to show that, despite the low sequence identity between Glt_{Ph} and Excitatory Amino Acid Transporters (EAATs), the mammalian counterparts of Glt_{Ph} and LeuT, their binding pockets are 60% homologous, with important amino acids being present in both Glt_{Ph} and EAATs, which should explain their different mechanisms [247] (EAATs co-transport 1 glutamate, 3 Na^+, $1H^+$ to the inside of the cell, while exporting 1 K^+ [248]). MD simulations of EAAT3, based on a homology model of the Glt_{Ph} homologue, have highlighted the necessity of protonating E374 to stabilize the binding of glutamate [249], revealed the location of the potassium binding site and shed light on the actual opening and closing mechanism.

Glt_{Ph} belongs to a rather large group of membrane transporters called sodium symporters. As already mentioned, their main driving force for transport is a sodium ion (and sometimes chloride) gradient, whose dissipation enables the transport of other substances, such as leucine and aspartate, besides glutamate. LeVine et al. [129] recently reviewed the information from different techniques, including MD, and compared different symporters, namely Glt_{Ph} and LeuT. They showed that these transporters share common features in their allosteric regulation: Regardless of sequence and 3D structure, they share common functional motifs that allow them to bind specifically both ions and substrate for their transport. Such findings lead to a better understanding of a protein class that also encloses sMATs (MATs will be further discussed in the following paragraph) and EAATs.

Another transporter protein, MonoAmine Transporters (MATs), is responsible for the reuptake of monoamine neurotransmitters (serotonin, dopamine and norepinephrine) in presynaptic neurons [250]. MATs are named accordingly to their substrate, hSERT (human Serotonin Transporter), hDAT (human Dopamine Transporter) and hNET (human Norepinephrine Transporter). The transport is energetically driven by the co-transport of sodium and chloride ions [251]. Unfortunately, no high-resolution structure is available for MAT. However, in 2015, Koldso et al. [252] developed models for each of the human MATs using *Drosophila melanogaster* DAT (dDAT) and an engineered

LeuT (LeuBAT) as templates. Their binding to substrate and to different psychostimulants, antidepressants and Mazindol, an anorectic, was assessed, revealing important parts of the mechanism of all human MATs.

Several studies throughout the years have revealed recurring structural motifs on different transporter proteins. Shi [253] gathered all recurring folds and secondary structures motifs presented by secondary transport proteins. These were considered the Major Facilitator Superfamily (MFS) fold, the LeuT fold and the Na^+/H^+ antiporter (NhaA) fold. They mostly consist of different rearrangements of α-helices into structurally identical monomers. Upon dimerization, these monomers forming the secondary transport protein. Another key feature of transporters is how monomers orient in an anti-parallel fashion upon dimerization.

5.3. Ion channels

Ion channels enable the crossing of hydrophilic molecules (ions) through a hydrophobic environment (lipid bilayer) [254]. These MPs are key at regulating the concentration of ions and are central to several biological functions, such as the generation and maintenance of electrochemical gradients [255]. Ion channels connecting intracellular and extracellular environments are generally narrow and very sensitive, while ion channels connecting intracellular environments are larger and allow for a steadier flux of ions. It was shown that these channels are ion selective and have "open" and "closed" states, allowing for the maintenance of electrochemical gradients. To better understand their role in neuronal communication, it is necessary to cross-reference ion channels with membrane carriers and pumps [256].

Ion pumps are essential for the maintenance of intracellular ion concentrations, being able to generate gradients at the expense of energy. Ion channels, on the other hand, feature sensitivity and high ion conductivity, allowing them to rapidly disrupt gradients generated by pumps by allowing ions to move according to their electrochemical gradient. This makes them decisive, for example, in the generation of electrical impulses, which can propagate and shape a well-functioning nervous system. In some cases, ions can also act as secondary messengers (e.g. zinc, calcium magnesium divalent cations), which is relevant in innate and adaptive immunity, where ion channels, pumps and carriers play a role in this process [257]. The membrane electrical potential is largely influenced by ion concentration and fast leakage, making these ion channels highly relevant in therapy, particularly in neurodegenerative diseases [258].

The voltage-gated ion channel superfamily, including the K^+-, Na^+- and Ca^{2+}-channels, with four homologous functional units consisting of six TM domains are key examples due to their important therapeutic and functional role. The fourth domain has been shown to work as the voltage sensor, while the linker between the fifth and sixth domains forms the pore structure for the passage of ions. The ligand operated receptor channel superfamily, associated to neurotransmitter receptors such as $GABA_A$ receptor, glutamate receptor and nicotinic acetylcholine receptor are MPs composed of five homologous subunits, each containing a large hydrophilic amino-terminal domain which binds ligands and/or competitive blockers, followed by four TM segments [259]. Computational algorithms have been developed for ion channel classification. VGIchan [260], makes use of amino acid sequences and ML-based tools to classify channels into sodium, potassium, calcium or chloride channels, while the Voltage-Gated K+ Channel Database (VKGDB) [261], analyses and compiles data regarding channels using phylogenetic analysis and HMMs. Due to the relevance and increasing information available regarding ligand-gated ion channels, the Ligand-Gated Ion Channel Database (LGICdb) was created [262], containing 554 entries (accessed in June 29th of 2017). Channelpedia is another database with information on most ion channels [263].

A few studies on structural prediction on ion channels were published recently. Ion channels usually undergo conformational changes to perform their designated biological functions. A structural model of the "down state" of a potassium channel voltage sensing domain (VSD) was predicted through homology modeling from a supposedly up-state – the gating state corresponding to the activated receptor [264]. The up and down states of VSD are important to understand conformational changes happening upon activation as they represent the states for positive and negative voltages, respectively. Acid-Sensing Ion Channels (ASICs) are voltage-insensitive cation channels sensitive to extracellular photons whose activity can be altered by several agents [265]. These channels are believed to be composed by 3 chains, each containing one large, bulky acid-sensing domain (ASD) and one TM domain (TMD). The ASDs of the Acid-Sensing Ion Channel 1 (ASIC1), besides binding both protons and chloride ions, are also responsible for the key interactions that maintain the homotrimeric structure of ASIC1, across three different regions [213] (Fig. 5C). Since ion channels typically visit a large ensemble of possible conformations, the combination of different computational techniques can lead to a better understanding of their structure-function relationship.

Homology modeling and evolutionary couplings were used in the study of Cyclic nucleotide-gated ion channels, involved in visual and olfactory sensory transduction [266]. The results support a modular model of allosteric gating, according to which protein domains can move independently, but are coupled to each other. Despite the lack of a proven global approach to the prediction of ion channel structure, they remain targets for drug design with computational methods such as homology modeling, ligand-based methods and structure-based methods being used on a case-by-case basis [267].

An important research focus for ion channel function are lipid-protein interactions, which have been shown to be an important factor affecting the channel structure and function [22]. To name a few examples, nonanular lipid binding to the pore domain of a mutant KcsA-Kv1.3 channel – a prokaryotic potassium channel and a human potassium channel, respectively – was shown to increase its activity [268] and, through an hybrid experimental-computational approach, Kir2.1 was shown to bind cholesterol in two nonannular hydrophobic regions that affect the hinge motion responsible for channel gating [269].

6. Conclusions

In this review, we have presented an overview of both theoretical and experimental methods used for the structural characterization of membrane proteins and their interactions. We have shown that the combination of computational tools and MP-associated data offers approaches that are complementary to the more expensive and time-consuming experimental studies. Even though computational methods for MP study depend on available data, the rising trend for experimental methods to perform better at characterizing MPs draws an exciting future for novel integrative approaches to study these highly relevant proteins. The many developments in computational methods for predicting MP topology, structure and interactions, together with the ever-increasing amount of experimental data, computational power and new user-friendly interfaces to access these methods, should boost research in this field and enable researchers with little training in computational biology to further advance the field of MP structure.

Funding

Irina S. Moreira acknowledges support by the Fundação para a Ciência e a Tecnologia (FCT) Investigator programme - IF/00578/2014 (co-financed by European Social Fund and Programa Operacional Potencial Humano), and a Marie Skłodowska-Curie Individual Fellowship MSCA-IF-2015 [MEMBRANEPROT 659826]. This work was also financed by the European Regional Development Fund (ERDF), through the Centro 2020 Regional Operational Programme under project CENTRO-01-0145-FEDER-000008: BrainHealth 2020, and through the

COMPETE 2020 - Operational Programme for Competitiveness and Internationalisation and Portuguese national funds via FCT, under project POCI-01-0145-FEDER-007440.

Panagiotis Koukos and Alexandre Bonvin acknowledge financial support from the Dutch Foundation for Scientific Research (NWO) (TOP-PUNT grant BC.000664.1).

Transparency document

The Transparency document associated with this article can be found, in the online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbamem.2017.07.008.

References

[1] (a) M.M. Gromiha, Y.Y. Ou, Bioinformatics approaches for functional annotation of membrane proteins, Brief. Bioinform. 15 (2) (2014) 155–168;
(b) E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, Protein Sci. 7 (4) (1998) 1029–1038;
(c) M.A. Yildirim, K.I. Goh, M.E. Cusick, A.L. Barabási, M. Vidal, Drug-target network, Nat. Biotechnol. 25 (10) (2007) 1119–1126.
[2] A. Doerr, Membrane protein structures, Nat. Methods 6 (1) (2009) 35.
[3] I. Moraes, G. Evans, J. Sanchez-Weatherby, S. Newstead, P.D.S. Stewart, Membrane protein structure determination — the next generation, Biochim. Biophys. Acta Biomembr. 1838 (1, Part A) (2014) 78–87.
[4] (a) J. Grouleff, S.J. Irudayam, K.K. Skeby, B. Schiøtt, The influence of cholesterol on membrane protein structure, function, and dynamics studied by molecular dynamics simulations, Biochim. Biophys. Acta Biomembr. 1848 (9) (2015) 1783–1795;
(b) J. Grouleff, S.J. Irudayam, K.K. Skeby, B. Schiott, The influence of cholesterol on membrane protein structure, function, and dynamics studied by molecular dynamics simulations, Biochim. Biophys. Acta 1848 (9) (2015) 1783–1795.
[5] (a) A.G. Lee, How lipids affect the activities of integral membrane proteins, Biochim. Biophys. Acta Biomembr. 1666 (1–2) (2004) 62–87;
(b) M.A. Alonso, J. Millán, The role of lipid rafts in signalling and membrane trafficking in T lymphocytes, J. Cell Sci. 114 (22) (2001) 3957–3965;
(c) D.A. Brown, E. London, Functions of lipid rafts in biological membranes, Annu. Rev. Cell Dev. Biol. 14 (1998) 111–136;
(d) P.V. Escribá, A. Ozaita, C. Ribas, A. Miralles, E. Fodor, T. Farkas, J.A. García-Sevilla, Role of lipid polymorphism in G protein-membrane interactions: nonlamellar-prone phospholipids and peripheral protein binding to membranes, Proc. Natl. Acad. Sci. 94 (21) (1997) 11375–11380;
(e) H. Hong, Y.C. Chang, J.U. Bowie, Measuring transmembrane helix interaction strengths in lipid bilayers using steric trapping, Methods Mol. Biol. (Clifton, N.J.) 1063 (2013) 37–56;
(f) Thomas A. Hopf, Lucy J. Colwell, R. Sheridan, B. Rost, C. Sander, Debora S. Marks, Three-dimensional structures of membrane proteins from genomic sequencing, Cell 149 (7) (2012) 1607–1621;
(g) L. Kall, A. Krogh, E.L. Sonnhammer, A combined transmembrane topology and signal peptide prediction method, J. Mol. Biol. 338 (5) (2004) 1027–1036;
(h) A. Khadria, A. Senes, Measurement of transmembrane peptide interactions in liposomes using Forster resonance energy transfer (FRET), Methods Mol. Biol. (Clifton, N.J.) 1063 (2013) 19–36.
[6] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 234–242.
[7] UniProt: a hub for protein information, Nucleic Acids Res. 43 (Database issue) (2015) D204–12.
[8] (a) J.C. Whisstock, A.M. Lesk, Prediction of protein function from protein sequence and structure, Q. Rev. Biophys. 36 (3) (2003) 307–340;
(b) T.N. Petersen, C. Lundegaard, M. Nielsen, H. Bohr, J. Bohr, S. Brunak, G.P. Gippert, O. Lund, Prediction of protein secondary structure at 80% accuracy, Proteins 41 (1) (2000) 17–20;
(c) B. Rost, C. Sander, Prediction of protein secondary structure at better than 70% accuracy, J. Mol. Biol. 232 (2) (1993) 584–599;
(d) R.H. Kretsinger, R.E. Ison, S. Hovmoller, Prediction of protein structure, Methods Enzymol. 383 (2004) 1–27;
(e) D. Shortle, Prediction of protein structure, Curr. Biol. 10 (2) (2000) R49–51;
(f) P. Argos, J.K. Rao, Prediction of protein structure, Methods Enzymol. 130 (1986) 185–207;
(g) Y.J. Edwards, A. Cottage, Prediction of protein structure and function by using bioinformatics, Methods Mol. Biol. 175 (2001) 341–375;
(h) L. Nanni, S. Brahnam, A. Lumini, Prediction of protein structure classes by incorporating different protein descriptors into general Chou's pseudo amino acid composition, J. Theor. Biol. 360 (2014) 109–116;
(i) C. Hartlmuller, C. Gobl, T. Madl, Prediction of protein structure using surface accessibility data, Angew. Chem. (2016);
(j) B. Al-Lazikani, E.E. Hill, V. Morea, Protein structure prediction, Methods Mol. Biol. 453 (2008) 33–85;
(k) D.R. Westhead, J.M. Thornton, Protein structure prediction, Curr. Opin. Biotechnol. 9 (4) (1998) 383–389;
(l) S.A. Benner, D.L. Geroff, J.D. Rozzell, Protein structure prediction, Science 274 (5292) (1996) 1448b-9b;
(m) G.J. Barton, R.B. Russell, Protein structure prediction, Nature 361 (6412) (1993) 1447–1451;
(n) B. Robson, J. Garnier, Protein structure prediction, Nature 361 (6412) (1993) 506;
(o) J. Garnier, Protein structure prediction, Biochimie 72 (8) (1990) 513–524.
[9] J. Koehler Leman, M.B. Ulmschneider, J.J. Gray, Computational modeling of membrane proteins, Proteins 83 (1) (2015) 1–24.
[10] V.K. Vyas, R.D. Ukawala, M. Ghate, C. Chintha, Homology modeling a fast tool for drug discovery: current perspectives, Indian J. Pharm. Sci. 74 (1) (2012) 1–17.
[11] G.A. Khoury, J. Smadbeck, C.A. Kieslich, C.A. Floudas, Protein folding and de novo protein design for biotechnological applications, Trends Biotechnol. 32 (2) (2014) 99–109.
[12] (a) T. Nugent, De novo membrane protein structure prediction, Methods Mol. Biol. 1215 (2015) 331–350;
(b) T. Nugent, D.T. Jones, Accurate de novo structure prediction of large transmembrane protein domains using fragment-assembly and correlated mutation analysis, Proc. Natl. Acad. Sci. U. S. A. 109 (24) (2012) E1540–7.
[13] (a) W.Q. Hui, Q. Cheng, T.Y. Liu, Q. Ouyang, Homology modeling, docking, and molecular dynamics simulation of the receptor GALR2 and its interactions with galanin and a positive allosteric modulator, J. Mol. Model. 22 (4) (2016) 90;
(b) P.J. Stansfeld, Computational studies of membrane proteins: from sequence to structure to simulation, Curr. Opin. Struct. Biol. 45 (2017) 133–141.
[14] B.A. Seaton, M.F. Roberts, Peripheral membrane proteins, in: K.M. Merz, B. Roux (Eds.), Biological Membranes: A Molecular Perspective from Computation and Experiment, Birkhäuser Boston, Boston, MA 1996, pp. 355–403.
[15] A.M. Whited, A. Johs, The interactions of peripheral membrane proteins with biological membranes, Chem. Phys. Lipids 192 (2015) 51–59.
[16] V. Monje-Galvan, J.B. Klauda, Peripheral membrane proteins: tying the knot between experiment and computation, Biochim. Biophys. Acta 1858 (7 Pt B) (2016) 1584–1593.
[17] S. London, O. Gurdal, C. Gall, Automatic export of PubMed citations to EndNote, Med. Ref. Serv. Q. 29 (2) (2010) 146–153.
[18] Y. Arinaminpathy, E. Khurana, D.M. Engelman, M.B. Gerstein, Computational analysis of membrane proteins: the largest class of drug targets, Drug Discov. Today 14 (23–24) (2009) 1130–1135.
[19] R.F. Alford, J. Koehler Leman, B.D. Weitzner, A.M. Duran, D.C. Tilley, A. Elazar, J.J. Gray, An integrated framework advancing membrane protein modeling and design, PLoS Comput. Biol. 11 (9) (2015), e1004398. .
[20] R. Peyronnet, D. Tran, T. Girault, J.M. Frachisse, Mechanosensitive channels: feeling tension in a world under pressure, Front. Plant Sci. 5 (2014).
[21] E.S. Haswell, R. Phillips, D.C. Rees, Mechanosensitive channels: what can they do and how do they do it? Structure 19 (10) (2011) 1356–1369.
[22] T.S. Tillman, M. Cascio, Effects of membrane lipids on ion channel structure and function, Cell Biochem. Biophys. 38 (2) (2003) 161–190.
[23] C.M. Tice, Y.J. Zheng, Non-canonical modulators of nuclear receptors, Bioorg. Med. Chem. Lett. 26 (17) (2016) 4157–4164.
[24] H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, Membrane Proteins, 4th ed W. H. Freeman, New York, 2000 (Section 3.4-Section 3.4).
[25] S.H. White, W.C. Wimley, Membrane protein folding and stability: physical principles, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 319–365.
[26] G.E. Schulz, Transmembrane beta-barrel proteins, Adv. Protein Chem. 63 (2003) 47–70.
[27] F. Bernaudat, A. Frelet-Barrand, N. Pochon, S. Dementin, P. Hivin, S. Boutigny, J.B. Rioux, D. Salvi, D. Seigneurin-Berny, P. Richaud, J. Joyard, D. Pignol, M. Sabaty, T. Desnos, E. Pebay-Peyrola, E. Darrouzet, T. Vernet, N. Rolland, Heterologous Expression of Membrane Proteins: Choosing the Appropriate Host, 2011.
[28] M. Zoonens, B. Miroux, Expression of membrane proteins at the *Escherichia coli* membrane for structural studies, Methods Mol. Biol. 601 (2010) 49–66.
[29] S. Wagner, L. Baars, A.J. Ytterberg, A. Klussmeier, C.S. Wagner, O. Nord, P.-A. Nygren, K.J. van Wijk, J.-W. de Gier, Consequences of membrane protein overexpression in *Escherichia coli*, Mol. Cell. Proteomics 6 (9) (2007) 1527–1550.
[30] S. Wagner, M.M. Klepsch, S. Schlegel, A. Appel, R. Draheim, M. Tarry, M. Högbom, K.J. van Wijk, D.J. Slotboom, J.O. Persson, J.-W. de Gier, Tuning *Escherichia coli* for membrane protein overexpression, Proc. Natl. Acad. Sci. U. S. A. 105 (38) (2008) 14371–14376.
[31] J.P. Rosenbusch, A. Lustig, M. Grabo, M. Zulauf, M. Regenass, Approaches to determining membrane protein structures to high resolution: do selections of subpopulations occur? Micron 32 (1) (2001) 75–90.
[32] G.G. Privé, Detergents for the stabilization and crystallization of membrane proteins, Methods 41 (4) (2007) 388–397.
[33] E.P. Carpenter, K. Beis, A.D. Cameron, S. Iwata, Overcoming the challenges of membrane protein crystallography, Curr. Opin. Struct. Biol. 18 (5) (2008) 581–586.
[34] V. Cherezov, A. Peddi, L. Muthusubramaniam, Y.F. Zheng, M. Caffrey, A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases, Acta Crystallogr. Sect. D 60 (2004) 1795–1807.
[35] C. Hunte, J. Koepke, C. Lange, T. Roßmanith, H. Michel, Structure at 2.3 Å resolution of the cytochrome bc1 complex from the yeast *Saccharomyces cerevisiae* co-crystallized with an antibody Fv fragment, Structure 8 (6) (2000) 669–684.

[36] X. Li, S. Dang, C. Yan, X. Gong, J. Wang, Y. Shi, Structure of a presenilin family intramembrane aspartate protease, Nature 493 (7430) (2013) 56–61.

[37] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G.F. Rasmussen, F.S. Thian, T.S. Kobilka, H.-J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, R.C. Stevens, High resolution crystal structure of an engineered human β(2)-adrenergic G protein-coupled receptor, Science (New York, N.Y.) 318 (5854) (2007) 1258–1265.

[38] W.D. Van Horn, H.-J. Kim, C.D. Ellis, A. Hadziselimovic, E.S. Sulistijo, M.D. Karra, C. Tian, F.D. Sönnichsen, C.R. Sanders, Solution NMR structure of membrane-integral diacylglycerol kinase, Science (New York, N.Y.) 324 (5935) (2009) 1726–1729.

[39] B. Liang, L. Tamm, NMR as a tool to investigate the structure, dynamics and function of membrane proteins, Nat. Struct. Mol. Biol. 23 (6) (2016) 468–474.

[40] K. Oxenoid, J.J. Chou, A functional NMR for membrane proteins: dynamics, ligand binding, and allosteric modulation, Protein Sci. 25 (5) (2016) 959–973.

[41] (a) D.T. Murray, N. Das, T.A. Cross, Solid state NMR strategy for characterizing native membrane protein structures, Acc. Chem. Res. 46 (9) (2013) 2172–2181;
(b) S.A. Shahid, B. Bardiaux, W.T. Franks, L. Krabben, M. Habeck, B.-J. van Rossum, D. Linke, Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals, Nat. Methods 9 (12) (2012) 1212–1217;
(c) A. Watts, S.K. Straus, S.L. Grage, M. Kamihira, Y.H. Lam, X. Zhao, in: A.K. Downing (Ed.), Membrane Protein Structure Determination Using Solid-State NMR BT - Protein NMR Techniques, Humana Press, Totowa, NJ 2004, pp. 403–473.

[42] S. Wang, R.A. Munro, S.Y. Kim, K.H. Jung, L.S. Brown, V. Ladizhansky, Paramagnetic relaxation enhancement reveals oligomerization interface of a membrane protein, J. Am. Chem. Soc. 134 (41) (2012) 16995–16998.

[43] S. Ganguly, B.E. Weiner, J. Meiler, Membrane protein structure determination using paramagnetic tags, Structure (London, England : 1993) 19 (4) (2011) 441–443.

[44] M. Kaplan, C. Pinto, K. Houben, M. Baldus, Nuclear magnetic resonance (NMR) applied to membrane-protein complexes, Q. Rev. Biophys. 49 (2016), e15.

[45] J.L. Milne, M.J. Borgnia, A. Bartesaghi, E.E. Tran, L.A. Earl, D.M. Schauder, J. Lengyel, J. Pierson, A. Patwardhan, S. Subramaniam, Cryo-electron microscopy—a primer for the non-microscopist, FEBS J. 280 (1) (2013) 28–45.

[46] M. Liao, E. Cao, D. Julius, Y. Cheng, Structure of the TRPV1 ion channel determined by electron cryo-microscopy, Nature 504 (7478) (2013) 107–112.

[47] E. Park, E.B. Campbell, R. MacKinnon, Structure of a CLC chloride ion channel by cryo-electron microscopy, Nature 541 (7638) (2017) 500–505.

[48] C.L. Lawson, A. Patwardhan, M.L. Baker, C. Hryc, E.S. Garcia, B.P. Hudson, I. Lagerstedt, S.J. Ludtke, G. Pintilie, R. Sala, J.D. Westbrook, H.M. Berman, G.J. Kleywegt, W. Chiu, EMDataBank unified data resource for 3DEM, Nucleic Acids Res. 44 (D1) (2016) D396–403.

[49] A.E. Rawlings, Membrane proteins: always an insoluble problem? Biochem. Soc. Trans. 44 (3) (2016) 790–795.

[50] I. Moraes, G. Evans, J. Sanchez-Weatherby, S. Newstead, P.D. Stewart, Membrane protein structure determination — the next generation, Biochim. Biophys. Acta 1838 (1 Pt A) (2014) 78–87.

[51] S. Lund, S. Orlowski, B. de Foresta, P. Champeil, M. le Maire, J.V. Moller, Detergent structure and associated lipid as determinants in the stabilization of solubilized Ca2+-ATPase from sarcoplasmic reticulum, J. Biol. Chem. 264 (9) (1989) 4907–4915.

[52] V.L. Postis, S.E. Deacon, P.C. Roach, G.S. Wright, X. Xia, J.C. Ingram, J.M. Hadden, P.J. Henderson, S.E. Phillips, M.J. McPherson, S.A. Baldwin, A high-throughput assay of membrane protein stability, Mol. Membr. Biol. 25 (8) (2008) 617–624.

[53] J.M. Gluck, M. Wittlich, S. Feuerstein, S. Hoffmann, D. Willbold, B.W. Koenig, Integral membrane proteins in nanodiscs can be studied by solution NMR spectroscopy, J. Am. Chem. Soc. 131 (34) (2009) 12060–12061.

[54] M. Caffrey, A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes, Acta Crystallogr. Sect. F Struct. Biol. Commun. 71 (Pt 1) (2015) 3–18.

[55] V. Cherezov, M. Caffrey, Membrane protein crystallization in lipidic mesophases. A mechanism study using X-ray microdiffraction, Faraday Discuss. 136 (2007) 195–212 (discussion 213-29).

[56] H.-B. Mao, G.-F. Li, D.-H. Li, Q.-Y. Wu, Y.-D. Gong, X.-F. Zhang, N.-M. Zhao, Effects of glycerol and high temperatures on structure and function of phycobilisomes in Synechocystis sp. PCC 6803, FEBS Lett. 553 (1–2) (2003) 68–72.

[57] Y.P. Zhang, R.N. Lewis, R.S. Hodges, R.N. McElhaney, Interaction of a peptide model of a hydrophobic transmembrane alpha-helical segment of a membrane protein with phosphatidylcholine bilayers: differential scanning calorimetric and FTIR spectroscopic studies, Biochemistry 31 (46) (1992) 11579–11588.

[58] A. Grau-Campistany, E. Strandberg, P. Wadhwani, J. Reichert, J. Burck, F. Rabanal, A.S. Ulrich, Hydrophobic mismatch demonstrated for membranolytic peptides, and their use as molecular rulers to measure bilayer thickness in native cells, Sci Rep 5 (2015) 9388.

[59] (a) J. Lebowitz, M.S. Lewis, P. Schuck, Modern analytical ultracentrifugation in protein science: a tutorial review, Protein Science : A Publication of the Protein Society 11 (9) (2002) 2067–2079;
(b) T.M. Weiss, P.C.A. van der Wel, J.A. Killian, R.E. Koeppe, H.W. Huang, Hydrophobic mismatch between helices and lipid bilayers, Biophys. J. 84 (1) (2003) 379–385;
(c) R.J. Webb, J.M. East, R.P. Sharma, A.G. Lee, Hydrophobic mismatch and the incorporation of peptides into lipid bilayers: a possible mechanism for retention in the Golgi, Biochemistry 37 (2) (1998) 673–679.

[60] D.M. Engelman, Y. Chen, C.-N. Chin, A.R. Curran, A.M. Dixon, A.D. Dupuy, A.S. Lee, U. Lehnert, E.E. Matthews, Y.K. Reshetnyak, A. Senes, J.-L. Popot, Membrane protein folding: beyond the two stage model, FEBS Lett. 555 (1) (2003) 122–125.

[61] T. Stockner, W.L. Ash, J.L. MacCallum, D.P. Tieleman, Direct simulation of transmembrane helix association: role of Asparagines, Biophys. J. 87 (3) (2004) 1650–1656.

[62] L. Cristian, Y. Zhang, Use of thiol-disulfide exchange method to study transmembrane peptide association in membrane environments, Methods Mol. Biol. (Clifton, N.J.) 1063 (2013) 3–18.

[63] L. Tome, D. Steindorf, D. Schneider, Genetic systems for monitoring interactions of transmembrane domains in bacterial membranes, Methods Mol. Biol. (Clifton, N.J.) 1063 (2013) 57–91.

[64] W. Guo, L. Shi, M. Filizola, H. Weinstein, J.A. Javitch, Crosstalk in G protein-coupled receptors: changes at the transmembrane homodimer interface determine activation, Proc. Natl. Acad. Sci. U. S. A. 102 (48) (2005) 17495–17500.

[65] A.G. Lee, How lipids affect the activities of integral membrane proteins, Biochim. Biophys. Acta 1666 (1–2) (2004) 62–87.

[66] J. Moult, K. Fidelis, A. Kryshtafovych, T. Schwede, A. Tramontano, Critical assessment of methods of protein structure prediction (CASP) — round x, Proteins: Struct. Funct. Bioinform. 82 (2014) 1–6.

[67] J.P. Ebejer, J.R. Hill, S. Kelm, J. Shi, C.M. Deane, Memoir: template-based structure prediction for membrane proteins, Nucleic Acids Res. 41 (Web Server issue) (2013) W379–83.

[68] S. Kelm, J. Shi, C.M. Deane, MEDELLER: homology-based coordinate generation for membrane proteins, Bioinformatics (Oxford, England) 26 (22) (2010) 2833–2840.

[69] D. Kozma, G.E. Tusnady, TMFoldWeb: a web server for predicting transmembrane protein fold class, Biol. Direct 10 (2015) 54.

[70] D. Kozma, G.E. Tusnady, TMFoldRec: a statistical potential-based transmembrane protein fold recognition tool, BMC Bioinform. 16 (2015) 201.

[71] V. Yarov-Yarovoy, J. Schonbrun, D. Baker, Multipass membrane protein structure prediction using Rosetta, Proteins 62 (4) (2006) 1010–1025.

[72] V. Yarov-Yarovoy, D. Baker, W.A. Catterall, Voltage sensor conformations in the open and closed states in ROSETTA structural models of K(+) channels, Proc. Natl. Acad. Sci. U. S. A. 103 (19) (2006) 7292–7297.

[73] E. Vargas, V. Yarov-Yarovoy, F. Khalili-Araghi, W.A. Catterall, M.L. Klein, M. Tarek, E. Lindahl, K. Schulten, E. Perozo, F. Bezanilla, B. Roux, An emerging consensus on voltage-dependent gating from computational modeling and molecular dynamics simulations, J. Gen. Physiol. 140 (6) (2012) 587–594.

[74] J. Koehler Leman, B.K. Mueller, J.J. Gray, Expanding the toolkit for membrane protein modeling in Rosetta, Bioinformatics (Oxford, England) 33 (5) (2017) 754–756.

[75] (a) I.S. Moreira, P. Koukos, R. Melo, J.G. Almeida, A.J. Preto, J. Schaarschmidt, M. Treller, Z.H. Gumus, J. Costa, A.M.J.J Bonvin, SpotOn: A Web Server for Protein-Protein Binding Hot-Spots, Sci. Rep., 2017 (accepted);
(b) C.R. Munteanu, A.C. Pimenta, C. Fernandez-Lozano, A. Melo, M.N. Cordeiro, I.S. Moreira, Solvent accessible surface area-based hot-spot detection methods for protein-protein and protein-nucleic acid interfaces, J. Chem. Inf. Model. 55 (5) (2015) 1077–1086;
(c) R. Melo, R. Fieldhouse, A. Melo, J.D. Correia, M.N. Cordeiro, Z.H. Gumus, J. Costa, A.M. Bonvin, I.S. Moreira, A machine learning approach for hot-spot detection at protein-protein interfaces, Int. J. Mol. Sci. 17 (8) (2016);
(d) A.L. Lomize, I.D. Pogozheva, TMDOCK: an energy-based method for modeling alpha-helical dimers in membranes, J. Mol. Biol. (2016);
(e) F.N. Baker, A. Porollo, CoeViz: a web-based tool for coevolution analysis of protein residues, BMC Bioinform. 17 (2016) 119;
(f) S. Doerr, M.J. Harvey, F. Noe, G. De Fabritiis, HTMD: high-throughput molecular dynamics for molecular discovery, J. Chem. Theory Comput. 12 (4) (2016) 1845–1852;
(g) L. Skjaerven, S. Jariwala, X.Q. Yao, B.J. Grant, Online interactive analysis of protein structure ensembles with Bio3D-web, Bioinformatics (Oxford, England) 32 (22) (2016) 3510–3512;
(h) V.K. Garg, H. Avashthi, A. Tiwari, P.A. Jain, P.W. Ramkete, A.M. Kayastha, V.K. Singh, MFPPI - multi FASTA ProtParam Interface, Bioinformation 12 (2) (2016) 74–77;
(i) A.A. Kaczor, E. Rutkowska, D. Bartuzi, K.M. Targowska-Duda, D. Matosiuk, J. Selent, Computational methods for studying G protein-coupled receptors (GPCRs), Methods Cell Biol. 132 (2016) 359–399;
(j) D.L. Parton, P.B. Grinaway, S.M. Hanson, K.A. Beauchamp, J.D. Chodera, Ensembler: enabling high-throughput molecular simulations at the superfamily scale, PLoS Comput. Biol. 12 (6) (2016), e1004728. ;
(k) C. Angermueller, T. Parnamaa, L. Parts, O. Stegle, Deep learning for computational biology, Mol. Syst. Biol. 12 (7) (2016) 878;
(l) T. Mori, N. Miyashita, W. Im, M. Feig, Y. Sugita, Molecular dynamics simulations of biological membranes and membrane proteins using enhanced conformational sampling algorithms, Biochim. Biophys. Acta 1858 (7 Pt B) (2016) 1635–1651;
(m) A.L. Lomize, M.A. Lomize, S.R. Krolicki, I.D. Pogozheva, Membranome: a database for proteome-wide analysis of single-pass membrane proteins, Nucleic Acids Res. 45 (D1) (2017) D250–D255;
(n) J.A. Morrone, A. Perez, J. MacCallum, K.A. Dill, Computed binding of peptides to proteins with MELD-accelerated molecular dynamics, J. Chem. Theory Comput. 13 (2) (2017) 870–876;
(o) K.M. Gayvert, O. Aly, J. Platt, M.W. Bosenberg, D.F. Stern, O. Elemento, A computational approach for identifying synergistic drug combinations, PLoS Comput. Biol. 13 (1) (2017), e1005308. ;
(p) P. Krishna, S. Sarvagalla, B. Madhuri, S. Pajaniradje, V. Baskaran, M.S. Coumar, R. Baskaran, Identification of natural inhibitors of Bcr-Abl for the treatment of chronic myeloid leukemia, Chem. Biol. Drug Des. (2017).

[76] Y. Baştanlar, M. Özuysal, Introduction to machine learning, in: M. Yousef, J. Allmer (Eds.), miRNomics: MicroRNA Biology and Computational Analysis, Humana Press, Totowa, NJ 2014, pp. 105–128.

[77] A. Lavecchia, Machine-learning approaches in drug discovery: methods and applications, Drug Discov. Today 20 (3) (2015) 318–331.

[78] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics (Oxford, England) 16 (4) (2000) 404–405.

[79] H. Cid, M. Bunster, E. Arriagada, M. Campos, Prediction of secondary structure of proteins by means of hydrophobicity profiles, FEBS Lett. 150 (1) (1982) 247–254.

[80] T. Hessa, H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S.H. White, G. von Heijne, Recognition of transmembrane helices by the endoplasmic reticulum translocon, Nature 433 (7024) (2005) 377–381.

[81] G. von Heijne, Membrane-protein topology, Nat. Rev. Mol. Cell Biol. 7 (12) (2006) 909–918.

[82] D.S. Marks, L.J. Colwell, R. Sheridan, T.A. Hopf, A. Pagnani, R. Zecchina, C. Sander, Protein 3D structure computed from evolutionary sequence variation, PLoS One 6 (12) (2011).

[83] (a) F.N. Baker, A. Porollo, CoeViz: a web-based tool for coevolution analysis of protein residues, BMC Bioinform. 17 (1) (2016) 119;
(b) J. Andreani, G. Faure, R. Guerois, InterEvScore: a novel coarse-grained interface scoring function using a multi-body statistical potential coupled to evolution, Bioinformatics (Oxford, England) 29 (14) (2013) 1742–1749.

[84] M.M. Gromiha, M. Suwa, Discrimination of outer membrane proteins using machine learning algorithms, Proteins 63 (4) (2006) 1031–1037.

[85] M.M. Gromiha, A simple method for predicting transmembrane alpha helices with better accuracy, Protein Eng. 12 (7) (1999) 557–561.

[86] P.L. Martelli, P. Fariselli, R. Casadio, An ENSEMBLE machine learning approach for the prediction of all-alpha membrane proteins, Bioinformatics (Oxford, England) (19 Suppl. 1) (2003) i205–11.

[87] M. Bernhofer, E. Kloppmann, J. Reeb, B. Rost, TMSEG: novel prediction of transmembrane helices, Proteins (2016).

[88] C. Savojardo, P. Fariselli, R. Casadio, BETAWARE: a machine-learning tool to detect and predict transmembrane beta-barrel proteins in prokaryotes, Bioinformatics (Oxford, England) 29 (4) (2013) 504–505.

[89] (a) G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (1–3) (2006) 489–501;
(b) L.P. Wang, C.R. Wan, Comments on "The extreme learning machine", IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council 19 (8) (2008) 1494–1495 (author reply 1495-6).

[90] P. Fariselli, C. Savojardo, P.L. Martelli, R. Casadio, Grammatical-restrained hidden conditional random fields for bioinformatics applications, Algorithms Mol. Biol. 4 (1) (2009) 13.

[91] J. Lafferty, A. McCallum, F. Pereira, Conditional random fields: probabilistic models for segmenting and labeling sequence data, Proceedings of ICML01, 2001.

[92] Y.Y. Ou, S.A. Chen, M.M. Gromiha, Prediction of membrane spanning segments and topology in beta-barrel membrane proteins at better accuracy, J. Comput. Chem. 31 (1) (2010) 217–223.

[93] K.-L. Du, M.N.S. Swamy, Radial basis function networks, Neural Networks and Statistical Learning, London, Springer London 2014, pp. 299–335.

[94] (a) A. Bernsel, H. Viklund, A. Hennerdal, A. Elofsson, TOPCONS: consensus prediction of membrane protein topology, Nucleic Acids Res. 37 (Web Server issue) (2009) W465–8;
(b) K.D. Tsirigos, C. Peters, N. Shu, L. Kall, A. Elofsson, The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides, Nucleic Acids Res. 43 (W1) (2015) W401–7.

[95] A. Pierleoni, V. Indio, C. Savojardo, P. Fariselli, P.L. Martelli, R. Casadio, MemPype: a pipeline for the annotation of eukaryotic membrane proteins, Nucleic Acids Res. 39 (Web Server issue) (2011) W375–80.

[96] H. Viklund, A. Elofsson, OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar, Bioinformatics (Oxford, England) 24 (15) (2008) 1662–1668.

[97] S.R. Eddy, What is a hidden Markov model? Nat. Biotechnol. 22 (10) (2004) 1315–1316.

[98] (a) D.T. Jones, W.R. Taylor, J.M. Thornton, A model recognition approach to the prediction of all-helical membrane protein structure and topology, Biochemistry 33 (10) (1994) 3038–3049;
(b) D.T. Jones, Improving the accuracy of transmembrane protein topology prediction using evolutionary information, Bioinformatics (Oxford, England) 23 (5) (2007) 538–544;
(c) T. Nugent, D.T. Jones, Transmembrane protein topology prediction using support vector machines, BMC Bioinform. 10 (2009) 159.

[99] S. Hayat, C. Peters, N. Shu, K.D. Tsirigos, A. Elofsson, Inclusion of dyad-repeat pattern improves topology prediction of transmembrane beta-barrel proteins, Bioinformatics (Oxford, England) 32 (10) (2016) 1571–1573.

[100] V.E. Angarica, J. Sancho, Protein dynamics governed by interfaces of high polarity and low packing density, PLoS One 7 (10) (2012) e48212.

[101] D. Nguyen, V. Helms, P.-H. Lee, PRIMSIPLR: prediction of inner-membrane situated pore-lining residues for alpha-helical transmembrane proteins, Proteins 82 (7) (2014) 1503–1511.

[102] H. Shen, J.J. Chou, MemBrain: improving the accuracy of predicting transmembrane helices, PLoS One 3 (6) (2008) e2399.

[103] A.D. Meruelo, I. Samish, J.U. Bowie, TMKink: a method to predict transmembrane helix kinks, Protein science : a publication of the Protein Society 20 (7) (2011) 1256–1264.

[104] M.S. Almen, K.J. Nordstrom, R. Fredriksson, H.B. Schioth, Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin, BMC Biol. 7 (2009) 50.

[105] S. Piccoli, E. Suku, M. Garonzi, A. Giorgetti, Genome-wide membrane protein structure prediction, Curr. Genomics 14 (5) (2013) 324–329.

[106] S.A. Fernando, P. Selvarani, S. Das, C.K. Kumar, S. Mondal, S. Ramakumar, K. Sekar, THGS: a web-based database of transmembrane helices in genome sequences, Nucleic Acids Res. 32 (Database issue) (2004) D125–D128.

[107] U. Pieper, A. Schlessinger, E. Kloppmann, G.A. Chang, J.J. Chou, M.E. Dumont, B.G. Fox, P. Fromme, W.A. Hendrickson, M.G. Malkowski, D.C. Rees, D.L. Stokes, M.H.B. Stowell, M.C. Wiener, B. Rost, R.M. Stroud, R.C. Stevens, A. Sali, Coordinating the impact of structural genomics on the human α-helical transmembrane proteome, Nat. Struct. Mol. Biol. 20 (2) (2013) 135–138.

[108] K.D. Tsirigos, P.G. Bagos, S.J. Hamodrakas, OMPdb: a database of β-barrel outer membrane proteins from gram-negative bacteria, Nucleic Acids Res. 39 (Database issue) (2011) D324–D331.

[109] P.J. Stansfeld, J.E. Goose, M. Caffrey, E.P. Carpenter, J.L. Parker, S. Newstead, M.S. Sansom, MemProtMD: automated insertion of membrane protein structures into explicit lipid membranes, Structure 23 (7) (2015) 1350–1361.

[110] (a) C. Munk, V. Isberg, GPCRdb: the G protein-coupled receptor database — an introduction, 173 (14) (2016) 2195–2207;
(b) V. Isberg, S. Mordalski, C. Munk, K. Rataj, K. Harpsøe, A.S. Hauser, B. Vroling, A.J. Bojarski, G. Vriend, D.E. Gloriam, GPCRdb: an information system for G protein-coupled receptors, Nucleic Acids Res. 44 (D1) (2016) D356–D364.

[111] D. Kozma, I. Simon, G.E. Tusnady, PDBTM: protein data Bank of transmembrane proteins after 8 years, Nucleic Acids Res. 41 (Database issue) (2013) D524–9.

[112] G.E. Tusnady, Z. Dosztanyi, I. Simon, TMDET: web server for detecting transmembrane regions of proteins by using their 3D coordinates, Bioinformatics (Oxford, England) 21 (7) (2005) 1276–1277.

[113] G.N. Tsaousis, K.D. Tsirigos, X.D. Andrianou, T.D. Liakopoulos, P.G. Bagos, S.J. Hamodrakas, ExTopoDB: a database of experimentally derived topological models of transmembrane proteins, Bioinformatics (Oxford, England) 26 (19) (2010) 2490–2492.

[114] D.R. Westhead, T.W. Slidel, T.P. Flores, J.M. Thornton, Protein structural topology: automated analysis and diagrammatic representation, Protein Sci. Publ. Protein Soc. 8 (4) (1999) 897–904.

[115] M.A. Lomize, A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, OPM: orientations of proteins in membranes database, Bioinformatics (Oxford, England) 22 (5) (2006) 623–625.

[116] G. Khelashvili, K. Dorff, J. Shan, M. Camacho-Artacho, L. Skrabanek, B. Vroling, M. Bouvier, L.A. Devi, S.R. George, J.A. Javitch, M.J. Lohse, G. Milligan, R.R. Neubig, K. Palczewski, M. Parmentier, J.P. Pin, G. Vriend, F. Campagne, M. Filizola, GPCR-OKB: the G protein coupled receptor oligomer Knowledge Base, Bioinformatics 26 (14) (2010) 1804–1805.

[117] J. Zhang, J. Yang, R. Jang, Y. Zhang, GPCR-I-TASSER: a hybrid approach to G protein-coupled receptor structure modeling and the application to the human genome, Structure 23 (8) (2015) 1538–1549.

[118] S. Wu, Y. Zhang, LOMETS: a local meta-threading-server for protein structure prediction, Nucleic Acids Res. 35 (10) (2007) 3375–3382.

[119] (a) Q. Ren, K. Chen, I.T. Paulsen, TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels, Nucleic Acids Res. 35 (Database issue) (2007) D274–9;
(b) Q. Ren, K.H. Kang, I.T. Paulsen, TransportDB: a relational database of cellular membrane transport systems, Nucleic Acids Res. 32 (Database issue) (2004) D284–8.

[120] B.R. Brooks, C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, CHARMM: the biomolecular simulation program, J. Comput. Chem. 30 (10) (2009) 1545–1614.

[121] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, Comput. Phys. Commun. 91 (1–3) (1995) 1–41.

[122] (a) S. Jo, J.B. Lim, J.B. Klauda, Im, W., CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes, Biophys. J. 97 (1) (2008) 50–58;
(b) E.L. Wu, X. Cheng, S. Jo, H. Rui, K.C. Song, E.M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R.M. Venable, J.B. Klauda, W. Im, CHARMM-GUI membrane builder toward realistic biological membrane simulations, J. Comput. Chem. 35 (27) (2014) 1997–2004.

[123] J.V. Ribeiro, R.C. Bernardi, T. Rudack, J.E. Stone, J.C. Phillips, P.L. Freddolino, K. Schulten, QwikMD - integrative molecular dynamics toolkit for novices and experts, Sci Rep 6 (2016) 26536.

[124] (a) B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theory Comput. 4 (3) (2008) 435–447;
(b) H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1–3) (1995) 43–56.

[125] O.S. Andersen, R.E. Koeppe II, Bilayer thickness and membrane protein function: an energetic perspective, Annu. Rev. Biophys. Biomol. Struct. 36 (2007) 107–130.

[126] T.B. Woolf, B. Roux, Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer, Proc. Natl. Acad. Sci. U. S. A. 91 (24) (1994) 11631–11635.

[127] (a) G. Khelashvili, S.G. Schmidt, L. Shi, J.A. Javitch, U. Gether, C.J. Loland, H. Weinstein, Conformational dynamics on the extracellular side of LeuT controlled by Na+ and K+ ions and the protonation state of Glu290, J. Biol. Chem. 291 (38) (2016) 19786–19799;

(b) M. Chavent, A.L. Duncan, M.S. Sansom, Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale, Curr. Opin. Struct. Biol. 40 (2016) 8–16.
[128] E. Lindahl, M.S. Sansom, Membrane proteins: molecular dynamics simulations, Curr. Opin. Struct. Biol. 18 (4) (2008) 425–431.
[129] M.V. LeVine, M.A. Cuendet, G. Khelashvili, H. Weinstein, Allosteric mechanisms of molecular machines at the membrane: transport by sodium-coupled symporters, Chem. Rev. 116 (11) (2016) 6552–6587.
[130] M.V. LeVine, G. Khelashvili, L. Shi, M. Quick, J.A. Javitch, H. Weinstein, Role of annular lipids in the functional properties of leucine transporter LeuT Proteomicelles, Biochemistry 55 (6) (2016) 850–859.
[131] (a) D. Hamelberg, J. Mongan, J.A. McCammon, Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules, J. Chem. Phys. 120 (24) (2004) 11919–11929;
(b) U. Doshi, D. Hamelberg, Towards fast, rigorous and efficient conformational sampling of biomolecules: advances in accelerated molecular dynamics, Biochim. Biophys. Acta 1850 (5) (2015) 878–888.
[132] K. Kappel, Y. Miao, J.A. McCammon, Accelerated molecular dynamics simulations of ligand binding to a muscarinic G-protein-coupled receptor, Q. Rev. Biophys. 48 (4) (2015) 479–487.
[133] G. Perez-Hernandez, F. Paul, T. Giorgino, G. De Fabritiis, F. Noe, Identification of slow molecular order parameters for Markov model construction, J. Chem. Phys. 139 (1) (2013), 015102.
[134] J.D. Chodera, W.C. Swope, J.W. Pitera, K.A. Dill, Long-time protein folding dynamics from short-time molecular dynamics simulations, Multiscale Model. Simul. 5 (4) (2006) 1214–1226.
[135] A.M. Razavi, G. Khelashvili, H. Weinstein, A Markov state-based quantitative kinetic model of sodium release from the dopamine transporter, Sci Rep 7 (2017) 40076.
[136] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chem. Phys. Lett. 314 (1–2) (1999) 141–151.
[137] H. Nymeyer, T.B. Woolf, A.E. Garcia, Folding is not required for bilayer insertion: replica exchange simulations of an alpha-helical peptide with an explicit lipid bilayer, Proteins 59 (4) (2005) 783–790.
[138] J. Gullingsrud, K. Schulten, Gating of MscL studied by steered molecular dynamics, Biophys. J. 85 (4) (2003) 2087–2099.
[139] R.C. Bernardi, M.C. Melo, K. Schulten, Enhanced sampling techniques in molecular dynamics simulations of biological systems, Biochim. Biophys. Acta 1850 (5) (2015) 872–877.
[140] S. Furini, C. Domene, Computational studies of transport in ion channels using metadynamics, Biochim. Biophys. Acta 1858 (7 Pt B) (2016) 1733–1740.
[141] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, The MARTINI force field: coarse grained model for biomolecular simulations, J. Phys. Chem. B 111 (27) (2007) 7812–7824.
[142] K.M. Davies, C. Anselmi, I. Wittig, J.D. Faraldo-Gomez, W. Kuhlbrandt, Structure of the yeast F1Fo-ATP synthase dimer and its role in shaping the mitochondrial cristae, Proc. Natl. Acad. Sci. U. S. A. 109 (34) (2012) 13602–13607.
[143] P. Rassam, N.A. Copeland, O. Birkholz, C. Toth, M. Chavent, A.L. Duncan, S.J. Cross, N.G. Housden, R. Kaminska, U. Seger, D.M. Quinn, T.J. Garrod, M.S.P. Sansom, J. Piehler, C.G. Baumann, C. Kleanthous, Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria, Nature 523 (7560) (2015) 333–336.
[144] L. Janosi, Z. Li, J.F. Hancock, A.A. Gorfe, Organization, dynamics, and segregation of Ras nanoclusters in membrane domains, Proc. Natl. Acad. Sci. U. S. A. 109 (21) (2012) 8097–8102.
[145] E. Jefferys, M.S.P. Sansom, P.W. Fowler, NRas slows the rate at which a model lipid bilayer phase separates, Faraday Discuss. 169 (0) (2014) 209–223.
[146] G. Hedger, M.S.P. Sansom, H. Koldsø, The juxtamembrane regions of human receptor tyrosine kinases exhibit conserved interaction sites with anionic lipids, Sci Rep 5 (2015) 9198.
[147] (a) I. Oren, S.J. Fleishman, A. Kessel, N. Ben-Tal, Free diffusion of steroid hormones across biomembranes: a simplex search with implicit solvent model calculations, Biophys. J. 87 (2) (2004) 768–779;
(b) B. Roux, S. Bernèche, W. Im, Ion channels, permeation, and electrostatics: insight into the function of KcsA, Biochemistry 39 (44) (2000) 13295–13306.
[148] N.A. Baker, Improving implicit solvent simulations: a Poisson-centric view, Curr. Opin. Struct. Biol. 15 (2) (2005) 137–143.
[149] (a) A. Onufriev, D.A. Case, D. Bashford, Effective born radii in the generalized born approximation: the importance of being perfect, J. Comput. Chem. 23 (14) (2002) 1297–1304;
(b) M. Feig, A. Onufriev, M.S. Lee, W. Im, D.A. Case, C.L. Brooks, Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures, J. Comput. Chem. 25 (2) (2004) 265–284.
[150] (a) W. Im, M.S. Lee, C.L. Brooks III, Generalized born model with a simple smoothing function, J. Comput. Chem. 24 (14) (2003) 1691–1702;
(b) W. Im, M. Feig, C.L. Brooks III, An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins, Biophys. J. 85 (5) (2003) 2900–2918.
[151] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (2) (1983) 187–217.
[152] S. Tanizaki, M. Feig, Molecular dynamics simulations of large integral membrane proteins with an implicit membrane model, J. Phys. Chem. B 110 (1) (2006) 548–556.
[153] (a) K.S. Kim, G. Neu J Fau-Oster, G. Oster, Effect of Protein Shape on Multibody Interactions between Membrane Inclusions, 2000 (1063-651X (Print));
(b) K.S. Kim, J. Neu, G. Oster, Curvature-mediated interactions between membrane proteins, Biophys. J. 75 (5) (1998) 2274–2291;
(c) D. Argudo, N.P. Bethel, F.V. Marcoline, M. Grabe, Continuum descriptions of membranes and their interaction with proteins: towards chemically accurate models, Biochim. Biophys. Acta 1858 (7 Pt B) (2016) 1619–1634.
[154] N. Ben-Tal, A. Ben-Shaul, A. Nicholls, B. Honig, Free-energy determinants of alpha-helix insertion into lipid bilayers, Biophys. J. 70 (4) (1996) 1803–1812.
[155] F.L. Brown, Continuum simulations of biomembrane dynamics and the importance of hydrodynamic effects, Q. Rev. Biophys. 44 (4) (2011) 391–432.
[156] Y.C. Zhou, B. Lu, A.A. Gorfe, Continuum electromechanical modeling of protein-membrane interactions, Phys. Rev. E Stat. Nonlinear Soft Matter Phys. 82 (4 Pt 1) (2010) 041923.
[157] W.M. Botello-Smith, X. Liu, Q. Cai, Z. Li, H. Zhao, R. Luo, Numerical Poisson-Boltzmann model for continuum membrane systems, Chem. Phys. Lett. 555 (2013) 274–281.
[158] N.R. Latorraca, K.M. Callenberg, J.P. Boyle, M. Grabe, Continuum approaches to understanding ion and peptide interactions with the membrane, J. Membr. Biol. 247 (5) (2014) 395–408.
[159] (a) P.B. Canham, The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell, J. Theor. Biol. 26 (1) (1970) 61–81;
(b) W. Helfrich, Elastic properties of lipid bilayers: theory and possible experiments, Z. Naturforsch. Teil C: Biochem. Biophys. Biol. Virol. 28 (11) (1973) 693–703.
[160] S. Tristram-Nagle, R. Zhang, R.M. Suter, C.R. Worthington, W.J. Sun, J.F. Nagle, Measurement of chain tilt angle in fully hydrated bilayers of gel phase lecithins, Biophys. J. 64 (4) (1993) 1097–1109.
[161] N. Dan, P. Pincus, S.A. Safran, Membrane-induced interactions between inclusions, Langmuir 9 (11) (1993) 2768–2771.
[162] K. Mayawala, D.G. Vlachos, J.S. Edwards, Spatial modeling of dimerization reaction dynamics in the plasma membrane: Monte Carlo vs. continuum differential equations, Biophys. Chem. 121 (3) (2006) 194–208.
[163] S. Markosyan, P.M. De Biase, L. Czapla, O. Samoylova, G. Singh, J. Cuervo, D.P. Tieleman, S.Y. Noskov, Effect of confinement on DNA, solvent and counterion dynamics in a model biological nanopore, Nano 6 (15) (2014) 9006–9016.
[164] S.Y. Noskov, W. Im, B. Roux, Ion permeation through the alpha-hemolysin channel: theoretical studies based on Brownian dynamics and Poisson-Nernst-Plank electrodiffusion theory, Biophys. J. 87 (4) (2004) 2299–2309.
[165] L. Rems, M. Tarek, M. Casciola, D. Miklavcic, Properties of lipid electropores II: comparison of continuum-level modeling of pore conductance to molecular dynamics simulations, Bioelectrochemistry (2016).
[166] J. Yoo, M.B. Jackson, Q. Cui, A comparison of coarse-grained and continuum models for membrane bending in lipid bilayer fusion pores, Biophys. J. 104 (4) (2013) 841–852.
[167] F. Pazos, M. Helmer-Citterich, G. Ausiello, A. Valencia, Correlated mutations contain information about protein-protein interaction, J. Mol. Biol. (1997) 271.
[168] L.C. Xue, D. Dobbs, V. Honavar, HomPPI: a class of sequence homology based protein-protein interface prediction methods, BMC Bioinform. 12 (1) (2011) 1–24.
[169] S.J. de Vries, A.M.J.J. Bonvin, CPORT: a consensus Interface predictor and its performance in prediction-driven docking with HADDOCK, PLoS One 6 (3) (2011) e17695.
[170] H. Neuvirth, R. Raz, G. Schreiber, ProMate: a structure based prediction program to identify the location of protein-protein binding sites, J. Mol. Biol. 338 (2004).
[171] S. Ahmad, K. Mizuguchi, Partner-aware prediction of interacting residues in protein-protein complexes from sequence data, PLoS One 6 (12) (2011) e29104.
[172] F. ul Amir Afsar Minhas, B.J. Geiss, A. Ben-Hur, PAIRpred: partner-specific prediction of interacting residues from sequence and structure, Proteins 82 (7) (2014) 1142–1155.
[173] L.C. Xue, D. Dobbs, A.M.J.J. Bonvin, V. Honavar, Computational prediction of protein interfaces: a review of data driven methods, FEBS Lett. 589 (23) (2015) 3516–3526.
[174] Y. Murakami, K. Mizuguchi, Applying the naive Bayes classifier with kernel density estimation to the prediction of protein-protein interaction sites, Bioinformatics (Oxford, England) 26 (2010).
[175] A. Porollo, J. Meller, Prediction-based fingerprints of protein-protein interactions, Proteins 66 (2007).
[176] (a) C. Feinauer, H. Szurmant, M. Weigt, A. Pagnani, Inter-protein sequence co-evolution predicts known physical interactions in bacterial ribosomes and the Trp operon, PLoS One 11 (2) (2016), e0149166. ;
(b) D.T. Jones, T. Singh, T. Kosciolek, S. Tetchner, MetaPSICOV: combining coevolution methods for accurate prediction of contacts and long range hydrogen bonding in proteins, Bioinformatics (Oxford, England) 31 (7) (2015) 999–1006.
[177] (a) B. Rost, C. Sander, Combining evolutionary information and neural networks to predict protein secondary structure, Proteins: Struct. Funct. Bioinform. 19 (1) (1994) 55–72;
(b) J.A. Capra, R.A. Laskowski, J.M. Thornton, M. Singh, T.A. Funkhouser, Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure, PLoS Comput. Biol. 5 (12) (2009) e1000585;
(c) J.A. Capra, M. Singh, Predicting functionally important residues from sequence conservation, Bioinformatics (Oxford, England) 23 (15) (2007) 1875–1882;
(d) J. Ng, R. Li, K. Morgan, J. Simon, Evolutionary conservation and predicted structure of the Drosophila extra sex combs repressor protein, Mol. Cell. Biol. 17 (11) (1997) 6663–6672.
[178] (a) H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules, Nucleic Acids Res. 44 (Web Server issue) (2016) W344–50;

(b) C.M. Buslje, J. Santos, J.M. Delfino, M. Nielsen, Correction for phylogeny, small number of observations and data redundancy improves the identification of coevolving amino acid pairs using mutual information, Bioinformatics 25 (9) (2009) 1125–1131;
(c) A. Ciancetta, D. Sabbadin, S. Federico, G. Spalluto, S. Moro, Advances in computational techniques to study GPCR-ligand recognition, Trends Pharmacol. Sci. 36 (12) (2015) 878–890;
(d) D. De Juan, F. Pazos, A. Valencia, Emerging methods in protein co-evolution, Nat. Rev. Genet. 14 (2013);
(e) A. Dehzangi, K. Paliwal, J. Lyons, A. Sharma, A. Sattar, Proposing a highly accurate protein structural class predictor using segmentation-based features, BMC Genomics 15 (Suppl. 1) (2014);
(f) J.P. Dekker, A. Fodor, R.W. Aldrich, G. Yellen, A perturbation-based method for calculating explicit likelihood of evolutionary co-variance in multiple sequence alignments, Bioinformatics 20 (2004);
(g) M. Figliuzzi, H. Jacquier, A. Schug, O. Tenaillon, M. Weigt, Coevolutionary landscape inference and the context-dependence of mutations in Beta-lactamase TEM-1, Mol. Biol. Evol. 33 (2016);
(h) A.A. Fodor, R.W. Aldrich, On evolutionary conservation of thermodynamic coupling in proteins, J. Biol. Chem. 279 (2004);
(i) K.Y. Yip, P. Patel, P.M. Kim, D.M. Engelman, D. McDermott, M. Gerstein, An integrated system for studying residue coevolution in proteins, Bioinformatics 24 (2008);
(j) J. Yu, M. Vavrusa, J. Andreani, J. Rey, P. Tuffery, R. Guerois, InterEvDock: a docking server to predict the structure of protein-protein interactions using evolutionary information, Nucleic Acids Res. 44 (W1) (2016) W542–9.
[179] (a) R. Jackups Jr., J. Liang, Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction, J. Mol. Biol. 354 (4) (2005) 979–993;
(b) H. Naveed, Y. Xu, R. Jackups Jr., J. Liang, Predicting three-dimensional structures of transmembrane domains of beta-barrel membrane proteins, J. Am. Chem. Soc. 134 (3) (2012) 1775–1781;
(c) H. Naveed, R. Jackups Jr., J. Liang, Predicting weakly stable regions, oligomerization state, and protein-protein interfaces in transmembrane domains of outer membrane proteins, Proc. Natl. Acad. Sci. U. S. A. 106 (31) (2009) 12735–12740.
[180] V. Nanda, D. Hsieh, A. Davis, Prediction and design of outer membrane protein-protein interactions, Methods Mol. Biol. 1063 (2013) 183–196.
[181] S. Hayat, P. Walter, Y. Park, V. Helms, Prediction of the exposure status of transmembrane beta barrel residues from protein sequence, J. Bioinforma. Comput. Biol. 9 (1) (2011) 43–65.
[182] N. Zaki, S. Bouktif, S. Lazarova-Molnar, A combination of compositional index and genetic algorithm for predicting transmembrane helical segments, PLoS One 6 (7) (2011) 1–8.
[183] Y. Park, S. Hayat, V. Helms, Prediction of the burial status of transmembrane residues of helical membrane proteins, BMC Bioinform. 8 (2007) 302.
[184] M. Hayat, A. Khan, WRF-TMH: predicting transmembrane helix by fusing composition index and physicochemical properties of amino acids, Amino Acids 44 (5) (2013) 1317–1328.
[185] Z. Yuan, F. Zhang, M.J. Davis, M. Boden, R.D. Teasdale, Predicting the solvent accessibility of transmembrane residues from protein sequence, J. Proteome Res. 5 (5) (2006) 1063–1070.
[186] M. Hayat, M. Tahir, PSOFuzzySVM-TMH: identification of transmembrane helix segments using ensemble feature space by incorporated fuzzy support vector machine, Mol. BioSyst. (2015).
[187] B. Cao, A. Porollo, R. Adamczak, M. Jarrell, J. Meller, Enhanced recognition of protein transmembrane domains with prediction-based structural profiles, Bioinformatics (Oxford, England) 22 (3) (2006) 303–309.
[188] A. Fuchs, A. Kirschner, D. Frishman, Prediction of helix-helix contacts and interacting helices in polytopic membrane proteins using neural networks, Proteins: Struct. Funct. Bioinform. 74 (4) (2009) 857–871.
[189] C.L. Worth, T.L. Blundell, On the evolutionary conservation of hydrogen bonds made by buried polar amino acids: the hidden joists, braces and trusses of protein architecture, BMC Evol. Biol. 10 (2) (2010) 161.
[190] H. Batoulis, T.H. Schmidt, P. Weber, J.G. Schloetel, C. Kandt, T. Lang, Concentration dependent ion-protein interaction patterns underlying protein oligomerization behaviours, Sci Rep 6 (2016) 24131.
[191] H. Zhou, Y. Zhou, Predicting the topology of transmembrane helical proteins using mean burial propensity and a hidden-Markov-model-based method, Protein science : a publication of the Protein Society 12 (7) (2003) 1547–1555.
[192] T. Beuming, H. Weinstein, A knowledge-based scale for the analysis and prediction of buried and exposed faces of transmembrane domain proteins, Bioinformatics (Oxford, England) 20 (12) (2004) 1822–1835.
[193] S. Ahmad, Y.H. Singh, Y. Paudel, T. Mori, Y. Sugita, K. Mizuguchi, Integrated prediction of one-dimensional structural features and their relationships with conformational flexibility in helical membrane proteins, BMC Bioinform. 11 (1) (2010) 533.
[194] A. Drozdetskiy, C. Cole, J. Procter, G.J. Barton, JPred4: a protein secondary structure prediction server, Nucleic Acids Res. (2015).
[195] B. Reva, Y. Antipin, C. Sander, Predicting the functional impact of protein mutations: application to cancer genomics, Nucleic Acids Res. (2011).
[196] T.A. Hopf, C.P.I. Schärfe, J.P.G.L.M. Rodrigues, A.G. Green, O. Kohlbacher, C. Sander, A.M.J.J. Bonvin, D.S. Marks, Sequence co-evolution gives 3D contacts and structures of protein complexes, elife 3 (2014), e03430.
[197] I.S. Moreira, P.A. Fernandes, M.J. Ramos, Protein–protein docking dealing with the unknown, J. Comput. Chem. 31 (2) (2010) 317–342.
[198] I. Halperin, B. Ma, H. Wolfson, R. Nussinov, Principles of docking: an overview of search algorithms and a guide to scoring functions, Proteins 47 (4) (2002) 409–443.
[199] (a) A.L. Lomize, H.I. Mosberg, Thermodynamic model of secondary structure for alpha-helical peptides and proteins, Biopolymers 42 (2) (1997) 239–269;
(b) A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, Anisotropic solvent model of the lipid bilayer. 2. Energetics of insertion of small molecules, peptides, and proteins in membranes, J. Chem. Inf. Model. 51 (4) (2011) 930–946.
[200] A.A. Polyansky, A.O. Chugunov, P.E. Volynsky, N.A. Krylov, D.E. Nolde, R.G. Efremov, PREDDIMER: a web server for prediction of transmembrane helical dimers, Bioinformatics (Oxford, England) 30 (6) (2014) 889–890.
[201] B.K. Mueller, S. Subramaniam, A. Senes, A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical Calpha-H hydrogen bonds, Proc. Natl. Acad. Sci. U. S. A. 111 (10) (2014) E888–95.
[202] R.F. Walters, W.F. DeGrado, Helix-packing motifs in membrane proteins, Proc. Natl. Acad. Sci. U. S. A. 103 (37) (2006) 13658–13663.
[203] T. Braun, J. Koehler Leman, O.F. Lange, Combining evolutionary information and an iterative sampling strategy for accurate protein structure prediction, PLoS Comput. Biol. 11 (12) (2015), e1004661.
[204] R. Fredriksson, M.C. Lagerstrom, L.G. Lundin, H.B. Schioth, The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints, Mol. Pharmacol. 63 (6) (2003) 1256–1272.
[205] X. Ding, X. Zhao, A. Watts, G-protein-coupled receptor structure, ligand binding and activation as studied by solid-state NMR spectroscopy, Biochem. J. 450 (3) (2013) 443–457.
[206] J.H. Park, T. Morizumi, Y. Li, J.E. Hong, E.F. Pai, K.P. Hofmann, H.W. Choe, O.P. Ernst, Opsin, a structural model for olfactory receptors? Angew. Chem. 52 (42) (2013) 11021–11024.
[207] B.K. Kobilka, G protein coupled receptor structure and activation, Biochim. Biophys. Acta 1768 (4) (2007) 794–807.
[208] B.C. Heng, D. Aubel, M. Fussenegger, An overview of the diverse roles of G-protein coupled receptors (GPCRs) in the pathophysiology of various human diseases, Biotechnol. Adv. 31 (8) (2013) 1676–1694.
[209] P. Maurice, J.L. Guillaume, A. Benleulmi-Chaachoua, A.M. Daulat, M. Kamal, R. Jockers, GPCR-interacting proteins, major players of GPCR function, Adv. Pharmacol. (San Diego, Calif.) 62 (2011) 349–380.
[210] D. Sengupta, A. Chattopadhyay, Molecular dynamics simulations of GPCR–cholesterol interaction: an emerging paradigm, Biochim. Biophys. Acta Biomembr. 1848 (9) (2015) 1775–1782.
[211] M. Scarselli, J.G. Donaldson, Constitutive internalization of G protein-coupled receptors and G, J. Biol. Chem. 284 (6) (2009) 3577–3585.
[212] S. Mehmood, V. Corradi, H.G. Choudhury, R. Hussain, P. Becker, D. Axford, S. Zirah, S. Rebuffat, D.P. Tieleman, C.V. Robinson, K. Beis, Structural and functional basis for lipid synergy on the activity of the antibacterial peptide ABC transporter McjD, J. Biol. Chem. (2016).
[213] J. Jasti, H. Furukawa, E.B. Gonzales, E. Gouaux, Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH, Nature 449 (7160) (2007) 316–323.
[214] O.M. Becker, S. Shacham, Y. Marantz, S. Noiman, Modeling the 3D structure of GPCRs: advances and application to drug discovery, Curr. Opin. Drug Discov. Dev. 6 (3) (2003) 353–361.
[215] N.G. Tsvetanova, R. Irannejad, M. von Zastrow, G protein-coupled receptor (GPCR) signaling via heterotrimeric G proteins from endosomes, J. Biol. Chem. 290 (11) (2015) 6689–6696.
[216] C.A. Johnston, D.P. Siderovski, Receptor-mediated activation of heterotrimeric G-proteins: current structural insights, Mol. Pharmacol. 72 (2) (2007) 219–230.
[217] T.B. Patel, Single transmembrane spanning heterotrimeric G protein-coupled receptors and their signaling cascades, Pharmacol. Rev. 56 (3) (2004) 371–385.
[218] J.A. Ballesteros, H. Weinstein, [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors, Methods Neurosci. 25 (1995) 366–428.
[219] (a) U. Gether, S. Lin, P. Ghanouni, J.A. Ballesteros, H. Weinstein, B.K. Kobilka, Agonists induce conformational changes in transmembrane domains III and VI of the β2 adrenoceptor, EMBO J. 16 (22) (1997) 6737–6747;
(b) J.D. Urban, W.P. Clarke, M. Von Zastrow, D.E. Nichols, B. Kobilka, H. Weinstein, J.A. Javitch, B.L. Roth, A. Christopoulos, P.M. Sexton, Functional selectivity and classical concepts of quantitative pharmacology, J. Pharmacol. Exp. Ther. 320 (1) (2007) 1–13;
(c) D. Zhang, H. Weinstein, Signal transduction by a 5-HT2 receptor: a mechanistic hypothesis from molecular dynamics simulations of the three-dimensional model of the receptor complexed to ligands, J. Med. Chem. 36 (7) (1993) 934–938.
[220] S. Hertig, N.R. Latorraca, R.O. Dror, Revealing atomic-level mechanisms of protein allostery with molecular dynamics simulations, PLOS Comput. Biol. 12 (6) (2016) e1004746.
[221] (a) F. Fanelli, M. Seeber, A. Felline, D. Casciari, F. Raimondi, Quaternary structure predictions and structural communication features of GPCR dimers, Prog. Mol. Biol. Transl. Sci. 117 (2013) 105–142;
(b) K.C. Jonas, F. Fanelli, I.T. Huhtaniemi, A.C. Hanyaloglu, Single molecule analysis of functionally asymmetric G protein-coupled receptor (GPCR) oligomers reveals diverse spatial and structural assemblies, J. Biol. Chem. 290 (7) (2015) 3875–3892.
[222] B. Farran, An update on the physiological and therapeutic relevance of GPCR oligomers, Pharmacol. Res. 117 (2017) 303–327.
[223] J.J. Linderman, Modeling of G-protein-coupled receptor signaling pathways, J. Biol. Chem. 284 (9) (2009) 5427–5431.

[224] B.D. Coleman, A. Marivin, K. Parag-Sharma, V. DiGiacomo, S. Kim, J.S. Pepper, J. Casler, L.T. Nguyen, M.R. Koelle, M. Garcia-Marcos, Evolutionary conservation of a GPCR-independent mechanism of trimeric G protein activation, Mol. Biol. Evol. (2015).
[225] M.Y.S. Kalani, N. Vaidehi, S.E. Hall, R.J. Trabanino, P.L. Freddolino, M.A. Kalani, W.B. Floriano, V.W.T. Kam, W.A. Goddard, The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists, Proc. Natl. Acad. Sci. U. S. A. 101 (11) (2004) 3815–3820.
[226] I.S. Moreira, L. Shi, Z. Freyberg, S.S. Ericksen, H. Weinstein, J.A. Javitch, Structural basis of dopamine receptor activation, in: K.A. Neve (Ed.), The Dopamine Receptors, Humana Press, Totowa, NJ 2010, pp. 47–73.
[227] E.B.V. Arnam, K.N. McCleary, F. Liu, R. Abrol, H.A. Lester, W.A. Goddard, Dougherty III, Investigation of dopamine receptor structure and function by structure prediction and unnatural amino acid mutagenesis, Biophys. J. 102 (3) (2010) 247a.
[228] O. Sensoy, I.S. Moreira, G. Morra, Understanding the differential selectivity of Arrestins toward the phosphorylation state of the receptor, ACS Chem. Neurosci. 7 (9) (2016) 1212–1224.
[229] I.S. Moreira, Structural features of the G-protein/GPCR interactions, Biochim. Biophys. Acta Gen. Subj. 1840 (1) (2014) 16–33.
[230] Y. Han, I.S. Moreira, E. Urizar, H. Weinstein, J.A. Javitch, Allosteric communication between protomers of dopamine class A GPCR dimers modulates activation, Nat. Chem. Biol. 5 (9) (2009) 688–695.
[231] O. Sensoy, H. Weinstein, A mechanistic role of helix 8 in GPCRs: computational modeling of the dopamine D2 receptor interaction with the GIPC1–PDZ-domain, Biochim. Biophys. Acta Biomembr. 1848 (4) (2015) 976–983.
[232] J.M. Perez-Aguilar, J. Shan, M.V. LeVine, G. Khelashvili, H. Weinstein, A functional selectivity mechanism at the serotonin-2A GPCR involves ligand-dependent conformations of intracellular loop 2, J. Am. Chem. Soc. 136 (45) (2014) 16044–16054.
[233] J.A. Dalton, X. Gomez-Santacana, A. Llebaria, J. Giraldo, Computational analysis of negative and positive allosteric modulator binding and function in metabotropic glutamate receptor 5 (in)activation, J. Chem. Inf. Model. 54 (5) (2014) 1476–1487.
[234] (a) M.H. Saier Jr., C.V. Tran, R.D. Barabote, TCDB: the transporter classification database for membrane transport protein analyses and information, Nucleic Acids Res. 34 (Database issue) (2006) D181–6;
(b) M.H. Saier Jr., V.S. Reddy, D.G. Tamang, A. Vastermark, The transporter classification database, Nucleic Acids Res. 42 (Database issue) (2014) D251–8;
(c) M.H. Saier Jr., V.S. Reddy, B.V. Tsu, M.S. Ahmed, C. Li, G. Moreno-Hagelsieb, The transporter classification database (TCDB): recent advances, Nucleic Acids Res. 44 (D1) (2016) D372–9;
(d) M.H. Saier Jr., M.R. Yen, K. Noto, D.G. Tamang, C. Elkan, The transporter classification database: recent advances, Nucleic Acids Res. 37 (Database issue) (2009) D274–8.
[235] S.A. Chen, Y.Y. Ou, T.Y. Lee, M.M. Gromiha, Prediction of transporter targets using efficient RBF networks with PSSM profiles and biochemical properties, Bioinformatics (Oxford, England) 27 (15) (2011) 2062–2067.
[236] A. Pohl, P.F. Devaux, A. Herrmann, Function of prokaryotic and eukaryotic ABC proteins in lipid transport, Biochim. Biophys. Acta 1733 (1) (2005) 29–52.
[237] M. Dean, Y. Hamon, G. Chimini, The human ATP-binding cassette (ABC) transporter superfamily, J. Lipid Res. 42 (7) (2001) 1007–1017.
[238] (a) M. Dean, T. Annilo, Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates, Annu. Rev. Genomics Hum. Genet. 6 (2005) 123–142;
(b) R.S. Molday, M. Zhong, F. Quazi, The role of the photoreceptor ABC transporter ABCA4 in lipid transport and Stargardt macular degeneration, Biochim. Biophys. Acta 1791 (7) (2009) 573–583.
[239] E.J. Tarling, T.Q. de Aguiar Vallim, P.A. Edwards, Role of ABC transporters in lipid transport and human disease, Trends Endocrinol. Metab. 24 (7) (2013) 342–350.
[240] (a) B. Marquez, F. Van Bambeke, ABC multidrug transporters: target for modulation of drug pharmacokinetics and drug-drug interactions, Curr. Drug Targets 12 (5) (2011) 600–620;
(b) R. van de Ven, R. Oerlemans, J.W. van der Heijden, G.L. Scheffer, T.D. de Gruijl, G. Jansen, R.J. Scheper, ABC drug transporters and immunity: novel therapeutic targets in autoimmunity and cancer, J. Leukoc. Biol. 86 (5) (2009) 1075–1087;
(c) S. Li, W. Zhang, X. Yin, S. Xing, H.Q. Xie, Z. Cao, B. Zhao, Mouse ATP-binding cassette (ABC) transporters conferring multi-drug resistance, Anti Cancer Agents Med. Chem. 15 (4) (2015) 423–432;
(d) Z. Chen, T. Shi, L. Zhang, P. Zhu, M. Deng, C. Huang, T. Hu, L. Jiang, J. Li, Mammalian drug efflux transporters of the ATP binding cassette (ABC) family in multidrug resistance: a review of the past decade, Cancer Lett. 370 (1) (2016) 153–164.
[241] M. Pinto, D. Digles, G.F. Ecker, Computational models for predicting the interaction with ABC transporters, Drug Discov. Today Technol. 12 (2014) e69–77.
[242] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (1) (2000) 235–242.
[243] M.M. Gromiha, P. Anoosha, D. Velmurugan, K. Fukui, Mutational studies to understand the structure-function relationship in multidrug efflux transporters: applications for distinguishing mutants with high specificity, Int. J. Biol. Macromol. 75 (2015) 218–224.
[244] N. Akyuz, E.R. Georgieva, Z. Zhou, S. Stolzenberg, M.A. Cuendet, G. Khelashvili, R.B. Altman, D.S. Terry, J.H. Freed, H. Weinstein, O. Boudker, S.C. Blanchard, Transport domain unlocking sets the uptake rate of an aspartate transporter, Nature 518 (7537) (2015) 68–73.
[245] O. Boudker, R.M. Ryan, D. Yernool, K. Shimamoto, E. Gouaux, Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter, Nature 445 (7126) (2007) 387–393.
[246] Y. Ji, V.L.G. Postis, Y. Wang, M. Bartlam, A. Goldman, Transport mechanism of a glutamate transporter homologue Glt(Ph), Biochem. Soc. Trans. 44 (3) (2016) 898–904.
[247] J. Setiadi, G. Heinzelmann, S. Kuyucak, Computational studies of glutamate transporters, Biomol. Ther. 5 (4) (2015) 3067–3086.
[248] C. Grewer, A. Gameiro, T. Rauen, SLC1 glutamate transporters, Arch. Eur. J. Physiol. 466 (1) (2014) 3–24.
[249] G. Heinzelmann, S. Kuyucak, Molecular dynamics simulations of the mammalian glutamate transporter EAAT3, PLoS One 9 (3) (2014), e92089.
[250] (a) B. Giros, M. Jaber, S.R. Jones, R.M. Wightman, M.G. Caron, Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter, Nature 379 (6566) (1996) 606–612;
(b) D. Bengel, D.L. Murphy, A.M. Andrews, C.H. Wichems, D. Feltner, A. Heils, R. Mossner, H. Westphal, K.P. Lesch, Altered brain serotonin homeostasis and locomotor insensitivity to 3, 4-methylenedioxymethamphetamine ("ecstasy") in serotonin transporter-deficient mice, Mol. Pharmacol. 53 (4) (1998) 649–655;
(c) F. Xu, R.R. Gainetdinov, W.C. Wetsel, S.R. Jones, L.M. Bohn, G.W. Miller, Y.M. Wang, M.G. Caron, Mice lacking the norepinephrine transporter are supersensitive to psychostimulants, Nat. Neurosci. 3 (5) (2000) 465–471.
[251] G.E. Torres, R.R. Gainetdinov, M.G. Caron, Plasma membrane monoamine transporters: structure, regulation and function, Nat. Rev. Neurosci. 4 (1) (2003) 13–25.
[252] H. Koldso, J. Grouleff, B. Schiott, Insights to ligand binding to the monoamine transporters-from homology modeling to LeuBAT and dDAT, Front. Pharmacol. 6 (2015) 208.
[253] Y. Shi, Common folds and transport mechanisms of secondary active transporters, Annu. Rev. Biophys. 42 (2013) 51–72.
[254] D.A. Doyle, Molecular insights into ion channel function (review), Mol. Membr. Biol. 21 (4) (2004) 221–225.
[255] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, Ion channels and the electrical properties of membranes, in: Science, G (Ed.), Molecular Biology of the Cell, 4th editionGarland Science, New York, 2002.
[256] D.C. Gadsby, Ion channels versus ion pumps: the principal difference, in principle, Nat. Rev. Mol. Cell Biol. 10 (5) (2009) 344–352.
[257] S. Feske, H. Wulff, E.Y. Skolnik, Ion channels in innate and adaptive immunity, Annu. Rev. Immunol. 33 (2015) 291–353.
[258] A. Waszkielewicz, A. Gunia, N. Szkaradek, K. Słoczyńska, S. Krupińska, H. Marona, Ion channels as drug targets in central nervous system disorders, Curr. Med. Chem. 20 (10) (2013) 1241–1285.
[259] H. Hirata, Review of molecular structure and function of ion channels, Nihon Rinsho Jpn. J. Clin. Med. 51 (4) (1993) 1065–1082.
[260] S. Saha, J. Zack, B. Singh, G.P. Raghava, VGIchan: prediction and classification of voltage-gated ion channels, Genomics Proteomics Bioinformatics 4 (4) (2006) 253–258.
[261] W.J. Gallin, P.A. Boutet, VKCDB: voltage-gated K+ channel database updated and upgraded, Nucleic Acids Res. 39 (Database issue) (2011) D362–6.
[262] N. Le Novere, J.P. Changeux, LGICdb: the ligand-gated ion channel database, Nucleic Acids Res. 29 (1) (2001) 294–295.
[263] R. Ranjan, G. Khazen, L. Gambazzi, S. Ramaswamy, S.L. Hill, F. Schürmann, H. Markram, Channelpedia: an integrative and interactive database for ion channels, Front. Neuroinform. 5 (2011).
[264] M. Grabe, H.C. Lai, M. Jain, Y.N. Jan, L.Y. Jan, Structure prediction for the down state of a potassium channel voltage sensor, Nature 445 (7127) (2007) 550–553.
[265] H.N. Khan, M.H. Khan, H. Rashid, Structure Prediction and Analysis of Mouse Amiloride-Sensitive Cation Channel 2, Neuronal Using Bioinformatics Tools, Emerging Technologies, 2009. ICET 2009. International Conference on, 19–20 Oct. 2009, 2009 168–172.
[266] Y. Gofman, C. Scharfe, D.S. Marks, T. Haliloglu, N. Ben-Tal, Structure, dynamics and implied gating mechanism of a human cyclic nucleotide-gated channel, PLoS Comput. Biol. 10 (12) (2014), e1003976.
[267] G. Fernández-Ballester, A. Fernández-Carvajal, M. González-Ros Jé, A. Ferrer-Montiel, Ionic channels as targets for drug design: a review on computational methods, Pharmaceutics 3 (4) (2011) 932–953.
[268] M. Weingarth, A. Prokofyev, E.A. van der Cruijsen, D. Nand, A.M. Bonvin, O. Pongs, M. Baldus, Structural determinants of specific lipid binding to potassium channels, J. Am. Chem. Soc. 135 (10) (2013) 3983–3988.
[269] A. Rosenhouse-Dantsker, S. Noskov, S. Durdagi, D.E. Logothetis, I. Levitan, Identification of novel cholesterol-binding regions in Kir2 channels, J. Biol. Chem. 288 (43) (2013) 31154–31164.