BRIEFINGS IN BIOINFORMATICS. VOL 15. NO 2. 155–168
Advance Access published on 23 March 2013
doi:10.1093/bib/bbt015

Bioinformatics approaches for functional annotation of membrane proteins

M. Michael Gromiha and Yu-Yen Ou

Submitted: 15th January 2013; Received (in revised form): 19th February 2013

Abstract

Membrane proteins perform diverse functions in living organisms such as transporters, receptors and channels. The functions of membrane proteins have been investigated with several computational approaches, such as developing databases, analyzing the structure–function relationship and establishing algorithms to discriminate different type of membrane proteins. However, compilation of bioinformatics resources for the functions of membrane proteins is not well documented compared with their structural aspects. In this comprehensive review, we elaborately focus on three aspects of membrane protein functions: (i) databases for different types of membrane proteins based on their functions including transporters, receptors and ion channels, annotated functional data for genomes, as well as functionally important amino acid residues in membrane proteins obtained from experimental data, (ii) analysis of membrane protein functions based on their structures, motifs, amino acid properties and other features and (iii) algorithms for discriminating different types of membrane proteins and annotating them in genomic sequences. In addition, we provide a list of online resources for the databases and web servers for functional annotation of membrane proteins.

Keywords: *membrane protein; channel; structure; function; transporter; receptor; motif; database*

INTRODUCTION

Membrane proteins play several important roles in living organisms such as transporters, receptors, channels and so on. They are embedded into lipid bilayers and have amino acid sequences, which will fold with a hydrophobic region in contact with the alkane chains of the lipids and polar surface in contact with the aqueous phases on both sides of the membrane and the polar head groups of the lipids. Membrane proteins have become attractive targets for pharmaceutical industries, and about 20–30% of protein sequences in genomes are identified as membrane proteins. On structural point of view, they are of two types, (i) transmembrane helical proteins, which span the cytoplasmic membrane with α-helices [1], and (ii) transmembrane strand proteins, which traverse the outer membranes of gram negative bacteria with β-strands [2]. The discrimination of transmembrane helical and strand proteins, prediction of their membrane spanning segments, thermodynamics and modeling of membrane proteins have been reviewed in detail [1–17].

On functional aspect, membrane proteins are involved in the transport of ions and molecules across the membrane, bind to small molecules at the extracellular space and recognize the immune system, cell-cell signaling and energy transducers. They are generally classified into transporters, ion channels and receptors. Transporters span the cell membrane with specific membrane topology,

Corresponding author. M. Michael Gromiha, Department of Biotechnology, Indian Institute of Technology Madras, Chennai 600036, Tamilnadu, India. Tel.: +91-44-2257-4138; Fax: +91-44-2257-4102; E-mail: gromiha@iitm.ac.in

M. Michael Gromiha is the Head of Protein Bioinformatics Lab at the Department of Biotechnology, Indian Institute of Technology Madras, Chennai 600036, Tamilnadu, India. His main research interest is on deciphering sequence-structure–function relationship in globular and membrane proteins using biophysics, bioinformatics and computational biology techniques. He is the author of a book on Protein Bioinformatics: From Sequence to Function published by Elsevier/Academic Press.

Yu-Yen Ou is an Assistant professor in the Department of Computer Science and Engineering, Yuan-Ze University, Taiwan. His fields of professional interests are bioinformatics, machine learning and data mining.

energy coupling mechanism and substrate specificities [18,19]. They form an intricate system of pumps and channels through which they deliver essential nutrients, eject waste products and assist the cell to sense environmental conditions [20].

Ion channels are membrane proteins, which facilitate the diffusion of ions across biological membranes. They provide a high-conducting hydrophilic pathway across hydrophobic interior of the membrane. Channels function into two different ways, such as gating and ion selectivity. The conformational change of a protein to keep the channel closed or open is termed as gating and it allows the transport of a molecule at the open state [21]. Channels can also discriminate the size and charge of the traversing molecule, which is termed as ion selectivity.

Membrane receptors are specialized protein molecules in the membranes of cells, to which external molecules (hormones, neurotransmitters, drugs) attach, triggering changes in the function of the cell. Essentially they receive external chemical signals and pass to the interior of the cell. Numerous receptor types such as nuclear receptors, G-protein coupled receptors, olfactory receptors (ORs) etc. are found within a cell, and each type is linked to a specific biochemical pathway. Recently it has been reported that the ORs in the antennas of malaria mosquito detect odorants produced by humans [22,23].

The classification of membrane proteins based on their function is an important problem toward the advancement of structural and functional genomics. In this review, we focus on three different aspects: (i) bioinformatics databases for the classification of membrane proteins based on their functions, annotated channels, receptors and transporters, motifs in membrane proteins as well as experimentally determined functionally important amino acid residues in membrane proteins, (ii) structure–function relationship in membrane proteins using amino acid properties, motifs and other features and (iii) discrimination algorithms and web servers for identifying channels, pores, transporters, transporting targets and so on. The G-protein coupled receptors have been well studied with the development of databases [24,25] and algorithms [26–29], and this topic has been extensively reviewed in the literature [30–33]. Hence, in this review, we focus on other classes of membrane proteins.

DATABASES FOR FUNCTIONAL ASPECTS OF MEMBRANE PROTEINS

The information about the functions of membrane proteins are increasing rapidly and it prompted the necessity of developing and maintaining databases to support scientific community. Consequently, different databases have been constructed on various directions, such as the overall collection of experimentally known functional information, specific data on transporters, channels and receptors, annotated functions in genomes, motifs in membrane proteins and so on. Table 1 lists a set of databases available for functional aspects of membrane proteins.

Transporters

Saier *et al.* [34] followed the transporter classification (TC) system approved by the International Union of Biochemistry and Molecular Biology and developed the Transporter Classification Database (TCDB). It is a comprehensive database for transporters and a brief description about it as described in Gromiha *et al.* [35] is given below. TCDB contains sequence, structure, function and evolutionary information about transport systems from a variety of living organisms. It is a curated repository for factual information compiled largely from over 10 000 published references. It uses a functional/phylogenetic system of classification, and currently encompasses about 5600 representative protein sequences and the proteins are classified into more than 600 families (26 June 2012). Transport systems are classified on the basis of five components such as (i) transporter class (i.e., channel, carrier, primary active transporter or group translocator), (ii) transporter subclass, which in the case of primary active transporters refers to the energy source used to drive transport, (iii) transporter family, (iv) subfamily and (v) substrate or range of substrates transported. The browsing options available in TCDB are illustrated in Figure 1 [35]. It also has the facility to retrieve data with several search options. TCDB is freely accessible at http://www.tcdb.org.

Ren *et al.* [19] classified the transporters into different types and families based on their mode of transport, bioenergetics, molecular phylogeny and substrate specificities and developed a database for annotated transporters in genomes. The annotation has been made with the available experimental information as well as bioinformatics approaches. The database contains the facility to search data based on

Table 1: List of databases for membrane protein functions

Name	URL
TCDB, transport classification database	http://www.tcdb.org
TransportDB, annotated transport systems in genomes	http://www.membranetransport.org/
Medicago truncatula transporter database	http://bioinformatics.cau.edu.cn/MtTransporter/
ABCdb, ABC transporter repertories	http://www-abcdb.biotoul.fr
ABCMdb, mutations in ABC transporters	http://abcmutations.hegelab.org/
GATMD: γ-aminobutyric acid transporter mutagenesis database	http://physiology.sci.csupomona.edu/GATMD/
YTPdb, yeast transport protein database	http://homes.esat.kuleuven.be/ytpdb/
TSdb, transporter substrate database	http://tsdb.cbi.pku.edu.cn
Olfactory receptor database	http://senselab.med.yale.edu/senselab/ordb
HORDE, human olfactory data explorer	http://genome.weizmann.ac.il/horde/
Channelpedia	http://www.ionchannels.org/
Ligand-gated ion channels database	http://www.ebi.ac.uk/compneur-srv/LGICdb
VKCDB, Voltage-gated K(+) channel database	http://vkcdb.biology.ualberta.ca
MIPModDB, resource for the superfamily of major intrinsic proteins	http://bioinfo.iitk.ac.in/MIPModDB
MeMotif, motifs in α-helical membrane proteins	http://projects.biotec.tu-dresden.de/memotif
TMFunction, functionally important residues in membrane proteins	http://tmbeta-genome.cbrc.jp/TMFunction/
dbPTM, posttranslational modifications of proteins	http://dbPTM.mbc.nctu.edu.tw/

text and BLAST search options. The database is also cross-linked with other related resources. It is available at http://www.membranetransport.org/. On this line, databases have also been developed for specific classes of transporters, and Miao *et al*. [36] reported a comprehensive database for *Medicago truncatula* transporters. It has the classifications from 162 families and seven types based on their transporting mode and energy coupling mechanisms. The database is available at http://bioinformatics.cau.edu.cn/MtTransporter/.

Fichant *et al*. [37] developed a database of ABC (ATP binding cassette) transporters in genomes based on their transporting compounds. A query page has been developed to retrieve ABC transporter repertories and analyze them based on functional and evolutionary perspectives. The database is available at http://www-abcdb.biotoul.fr. Gyimesi *et al*. [38] set up a database of ABC protein mutants from the information available in the literature using an automatic data mining approach. A web interface has been set up to compare mutations in different ABC transporters using sequence alignments and three-dimensional (3D) structural models. It is available at http://abcmutations.hegelab.org/. Anderson *et al*. [39] developed a web-accessible relational database of manually annotated biochemical, functional and pharmacological data reported on GAT1—the most intensely studied Gamma-aminobutyric acid (GABA) transporters. This GABA Transporter Mutagenesis Database (GATMD) is available at http://physiology.sci.csupomona.edu/GATMD/. Brohee *et al*. [40] developed a yeast transport protein database (YTPdb), for the classification and annotation of yeast transporters, which are based on the functional criteria such as subcellular location or their substrate compounds. YTPdb allows queries at various levels, from highly specific (e.g., ammonium as a substrate or vacuole as a location) to broader (e.g., cation as a substrate or inner membranes as location), and the database is available at http://homes.esat.kuleuven.be/ytpdb/. Zhao *et al*. [41] developed a transporter substrate database (TSdb) as a central repository of substrate information of transporters as well as their annotations. All the substrates in the database are mapped to Kyoto Encyclopedia of Genes and Genomes (KEGG) ligand compound database [42], which will be useful to map all the substrate to the KEGG pathway. TSdb is available at http://tsdb.cbi.pku.edu.cn/.

Receptors

Crasto *et al*. [43] constructed the OR database as a central repository for OR and OR-like gene and proteins sequences. They have developed an algorithm to automatically download sequences from GenBank and Uniprot into OR database. The data obtained from different resources have been structured with database architecture. HORDE is another database for ORs, which mainly deals with human ORs. It includes the information on genomic variations, classification, orthologs and other features. The database is available at http://genome.weizmann.ac.il/horde/.

Figure 1: Browsing data from Transport Classification Database. (**a**) hierarchical representation of transporters, such as class, subclass and family; (**b**) proteins belonging to a specific family, sugar porin family; (**c**) details about a typical protein, LamB, maltoporin. Figure was taken from Gromiha *et al*. [35].

Ion channels

Ion channels are transmembrane pores® proteins that allow the passage of ions into and out of a cell. There are hundreds of different ion channels and they are distinguished based on their ion selectivity, gating mechanism and sequence similarity [35]. Ion channels can be voltage-gated, ligand-gated, pH-gated or mechanically gated. These gating criteria along with a combination of sequence similarity and ion selectivity further subdivide ion channels into several subtypes, voltage-gated potassium channels, voltage-gated sodium channels, voltage-gated calcium channels, chloride channels, ligand-gated channels and so on. Ranjan *et al.* [44] designed a database for ion channels to store the information on ion channels and models. It integrates and highlights recent publications with relevant information. It has about 200 annotated ion channels, and the database Channepedia is available at http://www.ionchannels.org/.

Donizelli *et al.* [45] developed a nonredundant manually curated resource for ligand-gated ion channels, such as nicotinic, ATP, GABA and glutamate ionotropic receptors. It has both nucleic acid and protein sequences and facilities to search with keywords and sequence similarity. It is also equipped with customized multiple sequence alignments and manipulation of protein structures. The database is available at http://www.ebi.ac.uk/compneur-srv/LGICdb/. Gallin and Boutet [46] set up a voltage-gated K^+ channel database with a comprehensive set of sequence data, which can be directly used for further analysis. It has more than 2000 entries, and the channels are categorized into subfamilies by phylogenetic analysis and hidden Markov models. The database is available at http://vkcdb.biology.ualberta.ca. Gupta *et al.* [47] developed a database of major intrinsic proteins, which contains channel proteins with the following information: source, gene structure, sequence features, substitutions in the conserved Asn-Pro-Ala (NPA) motifs, structural model, the residues forming the selectivity filter and channel radius profile. MIPModDB database is available at http://bioinfo.iitk.ac.in/MIPModDB.

Functional motifs in membrane proteins

Specific motifs are overrepresented in membrane proteins and they play important roles for their structure and function [17]. It has been reported that GxxxG motif mediates helix packing in membrane proteins [48]. The glycine zippers, (G,A,S)xxx GxxxG and GxxxGxxx(G,S,T), with multiple GxxxG motifs are shown to have important functional roles in many membrane proteins [49]. The mutations in these motifs are shown to have deleterious effects on the function of membrane proteins [17,49,50]. Further, the motif Po.G..Hy.Hy (Po, polar residue; G, glycine; Hy, large hydrophobic residue) is identified as a sorting signal for the insertion and integration of mitochondrial β-barrel membrane proteins [51–53]. Owing to the importance of specific motifs, Marsico *et al.* [54] developed a database with more than 2000 experimentally known and computationally predicted motifs in transmembrane helical proteins. It is available at http://projects.biotec.tu-dresden.de/memotif.

Functionally important residues in membrane proteins

Gromiha *et al.* [55] developed a database TMFunction, which is a collection of experimentally observed functional residues in membrane proteins reported in the literature. The major aspects of this database have been reviewed in Gromiha *et al.* [35] and are discussed below. It contains the information about experimental functional data (IC50, measure of the effectiveness of a compound in inhibiting biological function; V_{max}, maximal velocity of transport; relative activity of mutants with respect to wild-type protein, binding affinity, dissociation constant) along with sequence, structure, mutation and literature information. A web interface has been set up with different search and display options so that users have the feasibility to get the relevant data. TMFunction is freely available at http://tmbeta-genome.cbrc.jp/TMFunction/. The usage of TMFunction is illustrated with the following example. The data obtained for the protein 'olfactory receptor', function 'receptor' and single mutants is shown in Figure 2A. The terms entry, protein, Uniprot ID, mutation, parameter, data, function, experiments and Pubmed ID have been selected for displaying the results (Figure 2B). Figure 2C shows the final results obtained with the search conditions and display options.

Posttranslational modifications of membrane proteins

Lu *et al.* [56] collected experimental and curated data on posttranslational modifications (PTM) of proteins and set up a database, dbPTM, for research

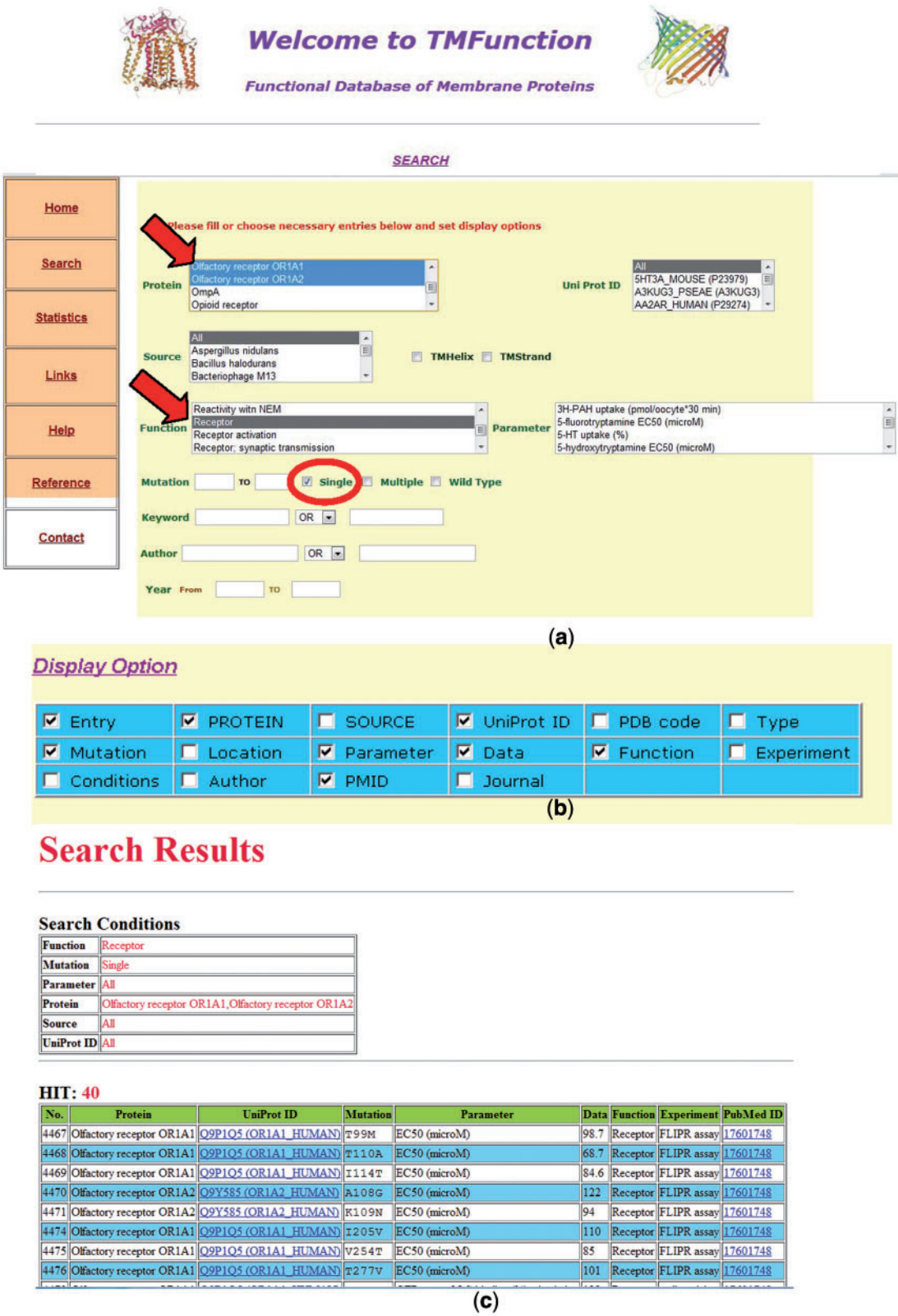

Figure 2: An example of searching conditions, display options and results of TMFunction: (**a**) main menu for the search options in TMFunction. The items protein name (olfactory receptor), function (receptor) and single mutants are selected for search as indicated by arrows/circles; (**b**) display options in TMFunction. We have selected entry, protein, Uniprot ID, mutation, parameter, data, function and PMID to show in the output; (**c**) part of the results obtained from TMFunction.

community. It also has the information on substrate specificity of PTM sites and functional association of PTMs between substrates and their interacting proteins. Specifically, structural topologies of membrane proteins have been integrated with PTM sites for understanding their structure–function relationships. In addition, graphical representations have been provided for indicating the PTM substrate sites and structural topology of membrane proteins.

STRUCTURE–FUNCTION RELATIONSHIP IN MEMBRANE PROTEINS

The experimental data related with membrane protein functions available in different databases have been used to understand the relationship between structure and function of membrane proteins. Gromiha *et al*. [57] related a set of physical, chemical, energetic and conformational properties of amino acid residues [58] with the change of odor response (or compound's potency or half maximal effective concentration, EC50) due to amino acid substitutions in ORs. In this method, the mutation induced changes in property values $\Delta P(i)$ was computed using the following equation:

$$\Delta P(i) = P_{mut}(i) - P_{wild}(i)$$

where, $P_{mut}(i)$ and $P_{wild}(i)$ are, respectively, the property value of the i^{th} mutant and wild-type residues, and i varies from 1 to N, total number of mutants. The computed difference in property values ΔP was related with experimental EC50 or odorant response using single correlation coefficient. The analysis revealed that both the characteristics of odorant molecule (ligand) and amino acid properties are important for odor response and EC50 [57]. Further, the information about the neighboring residues along the sequence and surrounding residues in 3D structures are also found to be important for understanding the functions of ORs on mutations. The sequence and structural effects have been included using the following expressions:

$$P\text{seq}(i) = \left[\sum_{j=i-k}^{j=i+k} P_j(i)\right] - P_{mut}(i)$$

where, $\Sigma P_j(i)$ is the total property value of the segment of $(2k+1)$ residues ranging from $i-k$ to $i+k$ about the i^{th} residue of wild type.

$$P_{str}(i) = P_{sur}(i) - P_{mut}(i);\ P_{sur}(i) = \sum_j n_{ij}.P_j$$

where, n_{ij} is the total number of type j residues surrounding the i^{th} residue of the protein within a distance of 8 Å, and P_j is the property value of residue type j.

Crasto [59] focused on ORs and reviewed the methods developed for modeling ORs, binding modes with odorant molecules, molecular dynamics simulations and the locations in ORs for ligand binding. Frelet and Klein [60] addressed the structure–function relationship in ABC transporters on various aspects such as the occurrence of two ATP binding and hydrolysis domains together with two ABC signatures, the nature of individual nucleotide-binding domains (independent or interacting), location of substrate binding sites and the functional link between ATP hydrolysis and transport process.

The analysis of amino acid composition in three major classes of transporters such as channels/pores, electrochemical and active transporters reveals the importance of specific amino acid residues with functions of transporters. Channels/pores catalyze facilitated diffusion (by an energy-independent process) by passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism. Electrochemical transporters are the ones that that use a carrier-mediated process. Active transporters use a primary source of energy to drive active transport of a solute against a concentration gradient.

Gromiha and Yabuki [61] revealed the importance of specific amino acid residues for different classes of membrane proteins, which were reviewed in Gromiha *et al*. [35]. Asn is dominant in channels/pores among all the transporters and, interestingly, it plays an important role to the stability and function of β-barrel membrane proteins [21,62]. Glu is another amino acid that shows high difference between channels and electrochemical transporters, and the residues Glu166 and Glu148 are important for the channel function in CIC chloride channel proteins [21]. The residues Phe and Leu are dominant in electrochemical transporters. In addition, the composition of Ala, Ile, Val and Trp are higher in electrochemical transporters compared with channels/pores and active transporters. This observation is supported with the fact that in glycerol-3-phosphate transporter, the space between helices 1 and 7 is filled by nine aromatic side chains and the occurrence of bulky aromatic residues helps to close the pore completely [63].

Marsico *et al*. [64] introduced a structural fragment clustering technique for comparing sequential motifs

with 3D structural fragments. They obtained 213 nonredundant motifs, and 58 of them were assigned to function using the information available in the literature. Seventy percent of the motifs are found in cofactor, ligand and ion binding sites, 30% at protein interaction interfaces and 12% bind specific lipids such as glycerol or cardiolipins.

Imai *et al*. [52,53] analyzed the existence of a specific motif Po.G..Hy.Hy (Po, polar residue; G, glycine; Hy, large hydrophobic residue) in mitochondrial β-barrel membrane proteins, which is reported to be a sorting signal for insertion into membrane and integration in lipid phase [51]. The analysis has been done with several sets of protein sequences from different categories, such as mitochondrial β-barrel membrane proteins with 90 and 40% sequence identities of known proteins, inner membrane proteins, soluble β-barrel proteins, matrix proteins etc., and we noticed that the signal is specific to mitochondrial β-barrel membrane proteins. Further, we have analyzed the existence of amino acid residues in different locations of the motifs in 70 mitochondrial β-barrel membrane proteins, and the results are presented in Table 2. We observed that the amino acid residues satisfy the pattern of the motif in all locations; Gly is conserved in all the sequences; hydrophobic residues are conserved with 89, 100 and 100% at three different positions in the considered set of proteins. These occurrences are remarkably higher than random chances mentioned in Table 2.

FUNCTIONAL DISCRIMINATION OF MEMBRANE PROTEINS

Several methods have been proposed to discriminate membrane proteins based on their functions such as classification of transporters, different types of ion channels, transporting targets, conformation of substrates entering into the channels and so on. These aspects will be discussed in this section and some of them have been described in Gromiha *et al*. [35]. The web servers, which discriminate membrane proteins based on their functions are listed in Table 3.

Discrimination of transporters and classification of transporter types

Functional annotation of genomes requires information, which can be obtained just from amino acid sequence. For the past two decades, several methods have been reported to identify membrane proteins based on their structure with high accuracy and are reviewed in detail. Recently, different algorithms have been proposed to identify membrane proteins based on their functions such as transporters, channels and receptors as well as their subclassifications. The discrimination methods follow a standard procedure as described below: (i) development of nonredundant datasets from well-established databases; the dataset is divided into two sets, one for training/cross-validation and another for independent test, (ii) identification of features and are generally, amino acid composition, amino acid occurrence, residue pair preference, amino acid properties, positions specific scoring matrices etc. (iii) constructing algorithms for discrimination, both statistical methods and machine learning techniques, (iv) assessing the performance using different measures, sensitivity, specificity, accuracy, precision, recall, F-measure and receiver operator characteristics estimated by area under the curve (AUC) and (v) validating with different methods, n-fold cross-validation, jack-knife test etc. The details about the

Table 2: The frequency of amino acid groups of β-signal motif sequences from 70 MBOMP homologs

Amino acid category	Motif position								
	Background	Po	X	G	hy	X	Hy	X	Hy
Large hydrophobic, Hy (L, I, V, F, M, W, Y)	0.32	0.03	**0.96**	0	**0.89**	0	**I**	0.16	**I**
Small hydrophobic, hy (A, C)	0.08	0.07	0.04	0	0.1	0.11	0	0.11	0
Glycine, G	0.07	0.16	0	**I**	0	0.24	0	0	0
Polar & +ve charge (K, R, H, S, T, N, Q)	0.39	**0.74**	0	0	0.01	0.5	0	0.61	0
−ve charge (D, E)	0.09	0	0.00	0.00	0.00	0.14	0.00	0.11	0.00
Proline	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Background: overall frequency of each amino acid group in the entire set of sequences. Bold letters indicate the conserved positions. Data were taken from [53].

Table 3: List of web servers for functional discrimination of membrane proteins

Name	URL
Discrimination of transporters	http://tmbeta-genome.cbrc.jp/disc-function/
TransportTP, transporter prediction	http://bioinfo3.noble.org/transporter
Transporter-RBF, classification of transporters	http://rbf.bioinfo.tw/~sachen/TCpredict/Transporter-RBF.php
Transporter Targets, prediction of transporter targets	http://rbf.bioinfo.tw/~sachen/ttrbf.html
SLITHER, conformation of transporting substrates	http://bioinfo.mc.ntu.edu.tw/slither/
P-glycoprotein substrate	http://pgp.althotas.com
VGIChan, classification of ion channels	http://www.imtech.res.in/raghava/vgichan/
Predicting ion channels and their types	http://cobi.uestc.edu.cn/people/hlin/tools/IonchanPred/
VKCpred, voltage-gated K+ channel subfamilies	http://cobi.uestc.edu.cn/people/hlin/tools/VKCPred/
MolAxis, identification of channels	http://bioinfo3d.cs.tau.ac.il/MolAxis
MEDELLER, modeling membrane proteins	http://medeller.info
PoreLogo, visualization of pore-lining residues	http://www.ebi.ac.uk/thornton-srv/software/PoreLogo

development of features, assessment and validation procedures are discussed in the literature [4].

Gromiha and Yabuki [61] reported that the amino acid occurrence could discriminate a set of membrane non-transporters and transporters with the 5-fold cross-validation accuracy of about 80% using k-nearest neighbor methods. A web server has been developed for discriminating transporters from other proteins, and it is available at http://tmbeta-genome.cbrc.jp/disc-function/. The usage of Position specific scoring matrix (PSSM) profiles and amino acid properties showed an increase of 5–10% in discrimination accuracy [65]. Li *et al*. [66] used traditional homology-based methods and machine learning techniques to detect the transporters in a set of completed genomes and showed a recall and precision of 82% in yeast proteome. The prediction server is available at http://bioinfo3.noble.org/transporter.

The next step is to classify the transporters into different classes, families and superfamilies. Efforts have been taken on this direction and significant progress has been reported. Gromiha and Yabuki [61] used different machine learning methods for discriminating channels/pores, electrochemical and active transporters, which showed an accuracy of 64% using amino acid occurrence. Ou *et al*. [65] used PSSM profiles and amino acid properties for discriminating these three classes of transporters and obtained an average accuracy of 78% in a test set of 118 proteins.

Schaadt and Helms [67] compared the similarity of transporters deposited in TCDB and annotated transporters in Arabidopsis thaliana using amino acid composition and classified the proteins into three families. They reported that the composition of transmembrane and non-transmembrane regions could classify four different families with an accuracy of 80%. Ou *et al*. [65] considered six major families in TCDB and developed a method based on PSSM and amino acid properties to classify them, which showed an average accuracy of 69%, with an improvement of 8% over amino acid composition. Li *et al*. [68] used nearest neighbor algorithm to distinguish 484 transporter families and reported a 5-fold cross-validated accuracy of 72.3%. Recently, Chen *et al*. [69] used the information of comparative BLAST scores to discriminate seven different superfamilies of transporters deposited in TCDB.

Prediction of transporting targets

Transporters catalyze the active transport of molecules across biological membranes and it is essential to annotate transporters specific to different substrates for the advancement of functional genomics. It is a difficult task and experimental annotation is scarce for our understanding. Hence, computational methods have been developed for identifying the transporters based on their targets. Schaadt *et al*. [70] used amino acid composition, characteristics of amino acid residues and conservation to detect transporters based on different substrates, amino acids, oligopeptides, phosphates and hexoses and showed an accuracy of 75–90%.

Chen *et al*. [71] considered four major classes of transporting targets such as (i) electron, (ii) protein/mRNA, (iii) ion and (iv) others and analyzed the characteristic features of transporters associated with these targets using amino acid properties. They have used various features, amino acid composition, residue pair preference, amino acid properties and PSSM profiles and developed an algorithm based on radial

basis function (RBF) networks to discriminate transporters with different transporting targets. This method showed an AUC of 0.90, 0.86, 0.77 and 0.86, respectively, for electron, protein/mRNA, ion and other transporters using PSSM and biochemical properties. A protocol has been developed to identify transporters based on their transporting targets, and the details are shown in Figure 3. In this procedure, PSSM features and other properties specific to each target are used to classify a transporter into four types of transporter targets. This procedure yields four results (for each transport type), and the transporter target is assigned according to the greatest preference (Figure 3).

Further, the work on transporting targets has been viewed through different perspectives, such as the conformation of a substrate to enter into the membrane and the interaction of a probable substrate with the protein. Lee *et al*. [72] developed a program SLITHER, which generates contiguous conformations of a molecule along a curved tunnel inside a protein and the binding free energy profile along the predicted channel pathway. The method can also be used to predict the ability of a substrate to crawl through an inner channel or half channel of proteins across surmountable energy barriers. The service is available at http://bioinfo.mc.ntu.edu.tw/slither/. Bikadi *et al*. [73] developed a support vector machine method for predicting the substrates to P-glycoprotein as well as their interactions. They showed that molecular docking to P-glycoprotein structures could predict the geometry of the protein–substrate complexes and reported an accuracy of 80% on an independent dataset of 32 compounds. The prediction and docking methods are available at http://pgp.althotas.com.

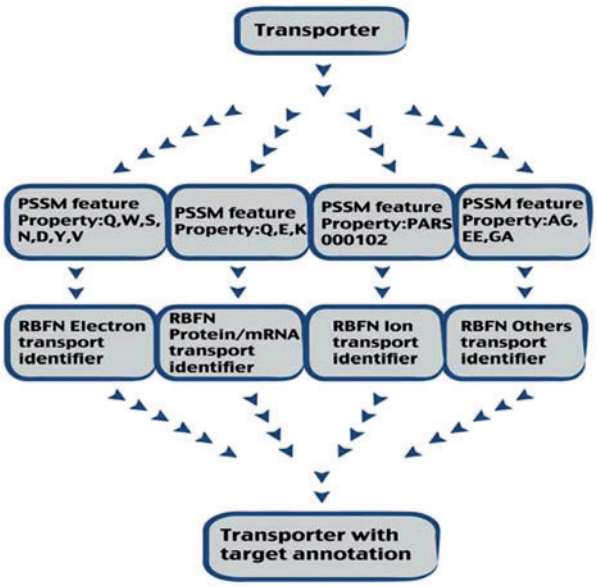

Figure 3: The architecture for annotating transporter targets with three steps [71]: (i) PSSM profiles for specific features, (ii) RBF networks for each target and (iii) final classification.

Identification of ion channels and discrimination of different classes

Ion channels are integral membrane proteins that enable the passage of inorganic ions across cell membranes. They are key components for physiological functions, and the role of ion channels in diseases has been described in a special issue in Progress in Biophysics and Molecular Biology [74]. Yaffe *et al*. [75] developed a web server for identifying channels in membrane proteins. It takes a protein structure in PDB format and lists all the main channels to membrane proteins. In addition, for each channel, the gating residues and the narrowest radius are also given along with a full list of the lining residues and the channel surface in a 3D graphical representation. The server is available at http://bioinfo3d.cs.tau.ac.il/MolAxis.

Ion channels are classified into different types such as voltage-gated potassium, calcium, sodium and ligand-gated channels, and they perform different functions. Hence, several methods have been proposed to discriminate the ion channels and classify them into different groups, which were reviewed in Gromiha *et al*. [35]. For example, Saha *et al*. [76] developed a method based on support vector machines to discriminate ion channels and non-ion channels and reported an accuracy of 89% to discriminate them. Further, the ion channels have been classified into potassium, sodium, calcium and chloride channels with one against others and obtained an average accuracy of 97.8%. A web server, VGIchan, has been developed for predicting and classifying voltage-gated ion channels, and it is available at www.imtech.res.in/raghava/vgichan/.

Lin and Ding [77] used a feature selection technique, analysis of variance and support vector machines to detect ion channels and classify them. They showed an accuracy of 86.6% for discriminating ion channels and non-ion channels. Further, voltage- and ligand-gated channels are distinguished with an accuracy of 92.6%, and four types of channels (potassium, sodium, calcium and anion) are classified

with an accuracy of 87.8%. A web server has been developed for prediction purposes, and it is available at http://cobi.uestc.edu.cn/people/hlin/tools/IonchanPred/. Recently, Chen and Lin [78] classified the voltage-gated K+ channels into seven subfamilies using amino acid composition and residue pair preference. The prediction models are available at http://cobi.uestc.edu.cn/people/hlin/tools/VKCPred/. The ion channels were also classified into ion channel target and ion channel non-target proteins using different properties of amino acid residues such as composition, hydrophobicity, polarity, polarizability and normalized van der Waals volume and RBF networks [79].

On a different approach, Li and Gallin [80] proposed a model for predicting the half-activation voltage of a voltage-gated potassium channel based on its amino acid sequence. They showed that the method could predict within the mean absolute error of 7.0 mV. Willett *et al.* [81] used binary kernel discrimination technique for predicting ion channel activity.

Channels and pores

In TCDB, channels and pores are grouped in the same class to include the transport systems that catalyze facilitated diffusion (by an energy-independent process) by passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism. However, channels have α-helical conformation, whereas pores have β-strands in their membrane spanning segments. As the amino acid sequences in the membrane spanning α-helical and β-strand segments are different, the former one is dominated with a stretch of hydrophobic residues and the latter one has both hydrophobic and polar/charged residues, classification of channels and pores could be done with high accuracy. As expected, amino acid composition alone could discriminate channels and pores with an accuracy of 92.4% [35,61].

MODELING AND VISUALIZATION

Transport proteins are targeted by several prescribed drugs and they can be potential targets for potential drug development. Hence, several methods have been proposed for modeling transporter proteins and mainly using homology modeling approach. Kelm *et al.* [82] presented a membrane protein–specific homology-based coordinate generation method, MEDELLER, and optimized it to build highly reliable core models. They showed that the method performed better than the methods developed mainly for globular proteins. Further, 3D structures of transporters have been modeled to understand the channel function and substrate specificity [83–90]. Recently, Ravna and Sylte [91] reviewed the computational approaches for constructing homology models of carriers and ion channels.

Oliva *et al.* [92] developed an automatic tool, PoreLogo for analysing, visualizing and comparing the amino acid composition of transmembrane channels and its conservation across the corresponding protein family. It is available at http://www.ebi.ac.uk/thornton-srv/software/PoreLogo/.

CONCLUSIONS

This review is focused on the bioinformatics approaches for understanding the functions of membrane proteins. The first part focused on the development of databases for different aspects of functions such as transporters, ion channels, receptors, motifs and functionally important residues in membrane proteins. The contents and features of these databases have been outlined. The second part deals with the structure–function relationship in membrane proteins on various perspectives such as relationship between amino acid properties and functional data of OR mutants, amino acid composition and importance of specific residues in transporters and sorting signal for mitochondrial β-barrel membrane proteins. Last part is devoted to prediction methods and web servers. The discrimination methods for identifying membrane transporters and predicting transporters into three classes, six families and transporting targets have been described. The methods reported for identifying channels and predicting the subclasses of ion channels have been discussed. Further, the approaches for modeling membrane proteins of functional importance and visualizing the amino acid composition of channel proteins have been mentioned. In addition, we have also provided a list of databases and online tools for functional annotation of proteins. In essence, this comprehensive review provides the insights on understanding the functions of membrane proteins on different perspectives.

Key points

- Bioinformatics approaches have made significant contribution to understand the functions of membrane proteins.
- A wealth of data on membrane protein functions have been deposited for receptors, transporters, channels and functionally important amino acid residues.
- Structure–function relationship for membrane proteins have been elucidated.
- Algorithms and tools for annotating different types of membrane proteins based on their function in genomic sequences have been developed.

FUNDING

Indian Institute of Technology Madras research grant (BIO/10-11/540/NFSC/MICH to M.M.G.) (in part), National Science Council (NSC) of Taiwan (NSC 101-2221-E-155-068 to Y-Y.O.) (in part).

References

1. White SH, Wimley WC. Membrane protein folding and stability: physical principles. *Ann Rev Biophys Biomol Str* 1999;**28**:319–65.
2. Schulz GE. Transmembrane beta-barrel proteins. *Adv Protein Chem* 2003;**63**:47–70.
3. Gromiha MM, Suwa M. Current developments on beta-barrel membrane proteins: sequence and structure analysis, discrimination and prediction. *Curr Prot Pept Science* 2007;**8**:580–99.
4. Gromiha M. *Protein Bioinformatics: From Sequence to Function*. New Delhi: Elsevier/Academic Press, 2010.
5. Chen CP, Rost B. State-of-the-art in membrane protein prediction. *Appl Bioinformatics* 2002;**1**:21–35.
6. Tusnády GE, Simon I. Topology prediction of helical transmembrane proteins: how far have we reached? *Curr Protein Pept Sci* 2010;**11**(7):550–61.
7. Fleishman SJ, Ben-Tal N. Progress in structure prediction of alpha-helical membrane proteins. *Curr Opin Struct Biol* 2006; **16**(4):496–504.
8. Hurwitz N, Pellegrini-Calace M, Jones DT. Towards genome-scale structure prediction for transmembrane proteins. *Philos Trans R Soc Lond B Biol Sci* 2006;**361**(1467): 465–75.
9. Senes A. Computational design of membrane proteins. *Curr Opin Struct Biol* 2011;**21**(4):460–6.
10. Fiedler S, Broecker J, Keller S. Protein folding in membranes. *Cell Mol Life Sci* 2010;**67**(11):1779–98.
11. Fleishman SJ, Unger VM, Ben-Tal N. Transmembrane protein structures without X-rays. *Trends Biochem Sci* 2006; **31**(2):106–13.
12. Punta M, Forrest LR, Bigelow H, *et al*. Membrane protein prediction methods. *Methods*. 2007;**41**(4):460–74.
13. Hedin LE, Illergård K, Elofsson A. An introduction to membrane proteins. *J Proteome Res* 2011;**10**(8):3324–31.
14. von Heijne G. Membrane-protein topology. *Nat Rev Mol Cell Biol* 2006;**7**(12):909–18.
15. Galdiero S, Galdiero M, Pedone C. beta-Barrel membrane bacterial proteins: structure, function, assembly and interaction with lipids. *Curr Protein Pept Sci* 2007;**8**(1):63–82.
16. Tamm LK, Hong H, Liang B. Folding and assembly of beta-barrel membrane proteins. *Biochim Biophys Acta* 2004; **1666**(1–2):250–63.
17. Nam HJ, Jeon J, Kim S. Bioinformatic approaches for the structure and function of membrane proteins. *BMB Rep* 2009;**42**(11):697–704.
18. Saier MH. A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* 2000;**64**(2):354–411.
19. Ren QH, Chen KX, Paulsen IT. TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Res* 2007;**35**:D274–9.
20. Saier MH, Tran CV, Barabote RD. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res* 2006;**34**:D181–6.
21. Dutzler R, Campbell EB, MacKinnon R. Gating the selectivity filter in ClC chloride channels. *Science* 2003; **300**:108–12.
22. Leal WS. Behavioural neurobiology: the treacherous scent of a human. *Nature* 2010;**464**(7285):37–8.
23. Carey AF, Wang G, Su CY, *et al*. Odorant reception in the malaria mosquito Anopheles gambiae. *Nature* 2010; **464**(7285):66–71.
24. Horn F, Weare J, Beukers MW, *et al*. GPCRDB: an information system for G protein-coupled receptors. *Nucleic Acids Res* 1998;**26**(1):275–9.
25. Satagopam VP, Theodoropoulou MC, Stampolakis CK, *et al*. GPCRs, G-proteins, effectors and their interactions: human-gpDB, a database employing visualization tools and data integration techniques. Database *(Oxford)* 2010;**2010**: baq019.
26. Papasaikas PK, Bagos PG, Litou ZI, *et al*. PRED-GPCR: GPCR recognition and family classification server. *Nucleic Acids Res* 2004;**32**(Web Server issue):W380–2.
27. Yabuki Y, Muramatsu T, Hirokawa T, *et al*. GRIFFIN: a system for predicting GPCR-G-protein coupling selectivity using a support vector machine and a hidden Markov model. *Nucleic Acids Res* 2005;**33**(Web Server issue): W148–53.
28. Bhasin M, Raghava GP. GPCRpred: an SVM-based method for prediction of families and subfamilies of G-protein coupled receptors. *Nucleic Acids Res* 2004; **32**(Web Server issue):W383–9.
29. Xiao X, Wang P, Chou KC. GPCR-2L: predicting G protein-coupled receptors and their types by hybridizing two different modes of pseudo amino acid compositions. *Mol. BioSyst* 2011;**7**:911–9.
30. Nemoto W, Toh H. Membrane interactive alpha-helices in GPCRs as a novel drug target. *Curr Protein Pept Sci* 2006; **7**(6):561–75.
31. Yarnitzky T, Levit A, Niv MY. Homology modeling of G-protein-coupled receptors with X-ray structures on the rise. Curr Opin Drug Discov Devel. 2010;**13**(3): 317–25.
32. Liu T, Tang GW, Capriotti E. Comparative modeling: the state of the art and protein drug target structure prediction. *Comb Chem High Throughput Screen* 2011;**14**(6): 532–47.

33. Selent J, Kaczor AA. Oligomerization of G protein-coupled receptors: computational methods. *Curr Med Chem* 2011; **18**(30):4588–605.
34. Saier MH, Yen MR, Noto K, *et al*. The Transporter Classification Database: recent advances. *Nucleic Acids Res* 2009;**37**:D274–8.
35. Gromiha MM, Ou YY, Chen SA. Structural and functional discrimination of membrane proteins. In: Gupta S. P. (ed). *Ion Channels and Their Inhibitors*. Germany: Springer, 2011; 1–32.
36. Miao Z, Li D, Zhang Z, *et al*. Medicago truncatula transporter database: a comprehensive database resource for M. *truncatula transporters*. BMC Genomics 2012;**13**:60.
37. Fichant G, Basse MJ, Quentin Y. ABCdb: an online resource for ABC transporter repertories from sequenced archaeal and bacterial genomes. *FEMS Microbiol Lett* 2006; **256**(2):333–9.
38. Gyimesi G, Borsodi D, Sarankó H, *et al*. ABCMdb: A database for comparative analysis of protein mutations in ABC transporters, and potential framework for a general application. *Hum Mutat* 2012;**33**:1547–56.
39. Anderson CM, Kidd PD, Eskandari S. GATMD: γ-Aminobutyric Acid Transporter Mutagenesis Database. *Database* 2010;**2010**:baq028.
40. Brohée S, Barriot R, Moreau Y, *et al*. YTPdb: a wiki database of yeast membrane transporters. *Biochim Biophys Acta* 2010;**1798**(10):1908–12.
41. Zhao M, Chen Y, Qu D, *et al*. TSdb: A database of transporter substrates linking metabolic pathways and transporter systems on a genome scale via their shared substrates. *Sci. China Life Sci.* 2011;**54**:60–4.
42. Kanehisa M, Goto S, Hattori M, *et al*. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 2006;**34**:D354–7.
43. Crasto C, Marenco L, Miller P, *et al*. Olfactory receptor database: a metadata-driven automated population from sources of gene and protein sequences. *Nucleic Acids Res* 2002;**30**(1):354–60.
44. Ranjan R, Khazen G, Gambazzi L, *et al*. Channelpedia: an integrative and interactive database for ion channels. *Front Neuroinform* 2011;**5**:36.
45. Donizelli M, Djite MA, Le Novère N. LGICdb: a manually curated sequence database after the genomes. *Nucleic Acids Res* 2006;**34**(Database issue):D267–9.
46. Gallin WJ, Boutet PA. VKCDB: voltage-gated K+ channel database updated and upgraded. *Nucleic Acids Res* 2011; **39**(Database issue):D362–6.
47. Gupta AB, Verma RK, Agarwal V, *et al*. MIPModDB: a central resource for the superfamily of major intrinsic proteins. *Nucleic Acids Res* 2012;**40**(Database issue): D362–9.
48. MacKenzie KR, Prestegard JH, Engelman DM. A transmembrane helix dimer: structure and implications. *Science* 1997;**276**:131–3.
49. Kim S, Jeon TJ, Oberai A, *et al*. Transmembrane glycine zippers: physiological and pathological roles in membrane proteins. *Proc Natl Acad Sci USA* 2005;**102**:14278–83.
50. Plotkowski ML, Kim S, Phillips ML, *et al*. Transmembrane domain of myelin protein zero can form dimers: possible implications for myelin construction. *Biochemistry* 2007;**46**: 12164–73.
51. Kutik S, Stojanovski D, Becker L, *et al*. Dissecting membrane insertion of mitochondrial beta-barrel proteins. *Cell* 2008;**132**(6):1011–24.
52. Imai K, Gromiha MM, Horton P. Mitochondrial beta-barrel proteins, an exclusive club? *Cell* 2008;**135**(7): 1158–9.
53. Imai K, Fujita N, Gromiha MM, *et al*. Eukaryote-wide sequence analysis of mitochondrial β-barrel outer membrane proteins. *BMC Genomics* 2011;**12**:79.
54. Marsico A, Scheubert K, Tuukkanen A, *et al*. MeMotif: a database of linear motifs in alpha-helical transmembrane proteins. *Nucleic Acids Res* 2010;**38**(Database issue):D181–9.
55. Gromiha MM, Yabuki Y, Suresh MX, *et al*. TMFunction: database for functional residues in membrane proteins. *Nucleic Acids Res* 2009;**37**:D201–4.
56. Lu CT, Huang KY, Su MG, *et al*. dbPTM 3.0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. *Nucleic Acids Res* 2013;**41**(D1):D295–305.
57. Gromiha MM, Harini K, Sowdhamini R, *et al*. Relationship between amino acid properties and functional parameters in olfactory receptors and discrimination of mutants with enhanced specificity. *BMC Bioinformatics* 2012; **13(Suppl 7)**:S1.
58. Gromiha MM, Oobatake M, Sarai A. Important amino acid properties for enhanced thermostability from mesophilic to thermophilic proteins. *Biophys Chem* 1999;**82**(1):51–67.
59. Crasto CJ. Computational biology of olfactory receptors. *Curr Bioinform* 2009;**4**(1):8–15.
60. Frelet A, Klein M. Insight in eukaryotic ABC transporter function by mutation analysis. *FEBS Lett* 2006;**580**(4): 1064–84.
61. Gromiha MM, Yabuki Y. Functional discrimination of membrane proteins using machine learning techniques. *BMC Bioinformatics* 2008;**9**:135.
62. Gromiha MM, Suwa M. A simple statistical method for discriminating outer membrane proteins with better accuracy. *Bioinformatics* 2005;**21**:961–8.
63. Huang YF, Lemieux MJ, Song JM, *et al*. Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. *Science* 2003;**301**:616–20.
64. Marsico A, Henschel A, Winter C, *et al*. Structural fragment clustering reveals novel structural and functional motifs in alpha-helical transmembrane proteins. *BMC Bioinformatics* 2010;**11**:204.
65. Ou YY, Chen SA, Gromiha MM. Classification of transporters using efficient radial basis function networks with position-specific scoring matrices and biochemical properties. *Proteins Struct Funct Bioinfo* 2010;**78**:1789–97.
66. Li HQ, Benedito VA, Udvardi MK, *et al*. TransportTP: a two-phase classification approach for membrane transporter prediction and characterization. *BMC Bioinformatics* 2009;**10**:418.
67. Schaadt NS, Helms V. Functional classification of membrane transporters and channels based on filtered TM/non-TM amino acid composition. *Biopolymers* 2012;**97**(7): 558–67. doi:10.1002/bip.22043.
68. Li HQ, Dai XB, Zhao XC. A nearest neighbor approach for automated transporter prediction and categorization from protein sequences. *Bioinformatics* 2008;**24**: 1129–36.

69. Chen JS, Reddy V, Chen JH, *et al*. Phylogenetic characterization of transport protein superfamilies: superiority of SuperfamilyTree programs over those based on multiple alignments. *J Mol Microbiol Biotechnol* 2011;**21**(3–4):83–96.
70. Schaadt NS, Christoph J, Helms V. Classifying substrate specificities of membrane transporters from Arabidopsis thaliana. *J Chem Inf Model* 2010;**50**(10):1899–905.
71. Chen SA, Ou YY, Lee TY, *et al*. Prediction of transporter targets using efficient RBF networks with PSSM profiles and biochemical properties. *Bioinformatics* 2011;**27**(15): 2062–7.
72. Lee PH, Kuo KL, Chu PY, *et al*. SLITHER: a web server for generating contiguous conformations of substrate molecules entering into deep active sites of proteins or migrating through channels in membrane transporters. *Nucleic Acids Res* 2009;**37**(Web Server issue):W559–64.
73. Bikadi Z, Hazai I, Malik D, *et al*. Predicting P-glycoprotein-mediated drug transport based on support vector machine and three-dimensional crystal structure of P-glycoprotein. *PLoS One* 2011;**6**(10):e25815.
74. Han J, Nilius B, Earm YE, *et al*. Ion channels: growing evidence for the role in disease. *Prog Biophys Mol Biol* 2010; **103**:1.
75. Yaffe E, Fishelovitch D, Wolfson HJ, *et al*. MolAxis: a server for identification of channels in macromolecules. *Nucleic Acids Res* 2008;**36**(Web Server issue):W210–5.
76. Saha S, Zack J, Singh B, *et al*. VGIchan: prediction and classification of voltage-gated ion channels. *Genomics Proteomics Bioinformatics* 2006;**4**:253–8.
77. Lin H, Ding H. Prediction of ion channels and their types. *J Theor Biol* 2011;**269**(1):64–9.
78. Chen W, Lin H. Identification of voltage-gated potassium channel subfamilies from sequence information using support vector machine. *Comput Biol Med* 2012;**42**(4):504–7.
79. Huang C, Zhang RJ, Chen ZQ, *et al*. Predict potential drug targets from the ion channel proteins based on SVM. *J Theor Biol* 2010;**262**:750–6.
80. Li B, Gallin WJ. Computational identification of residues that modulate voltage sensitivity of voltage-gated potassium channels. *BMC Struct Biol* 2005;**5**:16.
81. Willett P, Wilton D, Hartzoulakis B, *et al*. Prediction of ion channel activity using binary kernel discrimination. *J Chem Inf Model* 2007;**47**(5):1961–6.
82. Kelm S, Shi J, Deane CM. MEDELLER: homology-based coordinate generation for membrane proteins. *Bioinformatics* 2010;**26**(22):2833–40.
83. Durell SR, Hao Y, Guy HR. Structural models of the transmembrane region of voltage-gated and other K+ channels in open, closed, and inactivated conformations. *J Struct Biol* 1998;**121**(2):263–84.
84. Capener CE, Kim HJ, Arinaminpathy Y, *et al*. Ion channels: structural bioinformatics and modelling. *Hum Mol Genet* 2002;**11**(20):2425–33.
85. Bass RB, Locher KP, Borths E, *et al*. The structures of BtuCD and MscS and their implications for transporter and channel function. *FEBS Lett* 2003;**555**(1):111–5.
86. Tai K, Fowler P, Mokrab Y, *et al*. Molecular modeling and simulation studies of ion channel structures, dynamics and mechanisms. *Methods Cell Biol* 2008;**90**:233–65.
87. Ravna AW, Sylte I, Sager G. Binding site of ABC transporter homology models confirmed by ABCB1 crystal structure. *Theor Biol Med Model* 2009;**6**:20.
88. Honorat M, Falson P, Terreux R, *et al*. Multidrug resistance ABC transporter structure predictions by homology modeling approaches. *Curr Drug Metab* 2011; **12**(3):268–77.
89. Michino M, Chen JH, Stevens RC, *et al*. FoldGPCR: Structure prediction protocol for the transmembrane domain of G protein-coupled receptors from class A. *Proteins Struct Funct Bioinf* 2010;**78**:2189–201.
90. de Brevern AG. 3D structural models of transmembrane proteins. *Methods Mol Biol* 2010;**654**:387–401.
91. Ravna AW, Sylte I. Homology modeling of transporter proteins (carriers and ion channels). *Methods Mol Biol* 2012; **857**:281–99.
92. Oliva R, Thornton JM, Pellegrini-Calace M. PoreLogo: a new tool to analyse, visualize and compare channels in transmembrane proteins. *Bioinformatics* 2009;**25**(23): 3183–4.