Genome-wide Membrane Protein Structure Prediction

Stefano Piccoli[1,+], Eda Suku[1,+], Marianna Garonzi[1,+] and Alejandro Giorgetti[1,2,3,]*

[1]*Applied Bioinformatics Group, Dept. of Biotechnology, University of Verona, strada Le grazie 15, 37134, Verona, Italy;* [2]*German Research School for Simulation Sciences, Juelich, Germany;* [3]*Center for Biomedical Computing (CBMC), University of Verona, strada Le grazie 8, 37134, Verona, Italy*

Abstract: Transmembrane proteins allow cells to extensively communicate with the external world in a very accurate and specific way. They form principal nodes in several signaling pathways and attract large interest in therapeutic intervention, as the majority pharmaceutical compounds target membrane proteins. Thus, according to the current genome annotation methods, a detailed structural/functional characterization at the protein level of each of the elements codified in the genome is also required. The extreme difficulty in obtaining high-resolution three-dimensional structures, calls for computational approaches. Here we review to which extent the efforts made in the last few years, combining the structural characterization of membrane proteins with protein bioinformatics techniques, could help describing membrane proteins at a genome-wide scale. In particular we analyze the use of comparative modeling techniques as a way of overcoming the lack of high-resolution three-dimensional structures in the human membrane proteome.

INTRODUCTION

A lipidic membrane sorrounds cells with the aim of isolating them from the external world. This barrier has also an important role in cellular communication because it is the target of all the extracellular stimuli acting on the cell. So the membrane presents a major target for environmental stimuli acting upon a living cell. Several proteins, i.e. integral membrane proteins, are thus specialized in detecting extra cellular signals and translating the information to the cell, allowing a response. Membrane proteins are involved in several signaling pathways.

The original human genome sequencing project estimated 20% of the total gene count of 31,778 genes to code for membrane proteins [1]. This enormous number indicates the importance of membrane proteins for organism survival. Unfortunately, due to the difficulties in expression at the experimental level, just few of them have been deeply characterized. For example, if we consider the total amount of membrane proteins for which the three-dimensional structure is known, a feature absolutely needed for a complete functional characterization, we can observe that the structure of just 397 unique membrane proteins has been solved by X-ray crystallography (http://blanco.biomol.uci.edu/mpstruc/listAll/list). A total of 40 of them are human membrane proteins (38 transmembrane alpha protein and 2 monotopic membrane protein).

While 26 of these structures were solved by X-ray crystallography just 14 were solved by Nuclear magnetic resonance (NMR) techniques. These numbers represent a small portion of the entire human membrane proteome. It is evident from the structures that have been deposited so far, that only all-alpha and beta-barrel structural organizations are present in nature. Indeed most of membrane proteins in the Protein Data Bank (PDB: http://www.pdb.org), i.e. 67%, consist of bundles of transmembrane helices with different tilting with respect to the membrane plane and to each other.

Recently, membrane protein crystallography techniques have reached higher levels of applicability, allowing the resolution of proteins of the upmost importance for the scientific community. Anyway, although the revolution in 'resolution' techniques, we have to admit that there is still an enormous gap between annotated protein sequences and structures. The need of membrane protein characterization, thus calls for alternative innovative approaches. One of these consisted in using an extensive combination of computational biology techniques with molecular biology validating experiments.

Here we review how these could help describing membrane proteins at a genome-wide scale. In particular, we will try to give a numerical idea about the possibility to use homology modeling techniques in order to obtain structural models of all the human membrane proteome.

PROTEIN STRUCTURAL BIOINFORMATICS

Proteins are linear polymers of aminoacidic residues, modeled by events of random change and natural selection along a period of millions of years, fulfilling a delicate sequence-structure-function relationship. Thus, evolutionarily related

*Address correspondence to this author at the Applied Bioinformatics Group, Dept. of Biotechnology, University of Verona, strada Le grazie 15, 37134, Verona, Italy; Tel/Fax: +39 0458027905;
E-mail: alejandro.giorgetti@univr.it

[+] *Contributed equally to the work.*

proteins (homologs) that descend from a common ancestor accumulated small changes still conserving the same general fold. This forms the basis of the technique known as comparative or homology modeling [2, 3]. Chothia and Lesk, in a seminal article [4] have aligned the sequences and structures of all the soluble proteins with known structure (up to 1986), finding a correlation between the evolutionary distance and structural divergence between evolutionary related proteins. Summarizing, comparative modeling is a method that allows the prediction of protein structures using as a template a member of the family for which the three-dimensional structure (3D-structure) is known [2, 3]. Very recently, we have checked for the existence of correlation between evolutionary relationship and structure similarity on membrane proteins, following the same protocol as Chothia and Lesk [4, 5]. By using the LGA server (http://protein-model.org/) we have aligned, at the sequence and structural level, the core region of all the membrane proteins with known three-dimensional structure (565 pairs) and produced the graph of (Fig. **1**) (Here we have improved the graph by including the recently solved protein structures). The structural divergence between two evolutionary correlated proteins is measured as their Root Mean Square Deviation (RMSD).

For comparison purposes the core of the protein was considered as composed by all the amino acids present in secondary structure elements and those regions not diverging for more than 3 Å, as in [4]. The p value was $2,2E^{-73}$, indicating the statistical significance of the graph. We have then calculated the least square fit of the data, as performed in [4], obtaining the black continuous curve of (Fig. **1**). The calculated $R^2 = 0.83$ value demonstrates the high correlation between structure and sequence similarity. If we consider that Chothia and Lesk have calculated that plot more than 28 years ago considering just 32 pairs of soluble proteins, we have to admit that the result for 565 pairs of membrane proteins is astonishing. Although the noise in the low identity region (high fraction of mutated residues) seems to be higher, in the last few years (see Application Cases), several

membrane proteins, present in this 'twilight region', that is, sharing sequence identities lower than 20% with their templates, were successfully modeled.

The question at this point can be: Would it be possible to obtain accurate structural models of the entire human membrane proteome?

TRANSMEMBRANE DOMAIN MODELING STATISTICS

A recent project, *Survey of the human transmembrane proteome* [6], allowed the assessment of the feasibility for predicting the structure by comparative modeling techniques of the human membrane protein domains. The tools and the database developed in the ambit of this project are associated with ModPipe [7], an automated modelling pipeline. In this project, all the human transmembrane all-α domains were identified, clustered and classified into three different classes, i.e. enzymes, transporters and receptors, totalizing 3417 non redundant domains with at least two α-helices. These domains were then funneled through two different automatic pipelines, i.e. ModPipe and BLASTCLUST [8], with the aim of identifying putative templates. While with ModPipe templates were identified for 984 domains (threshold level of 25% sequence identity - SI), BLASTCLUST found 1201 ones (threshold level of 25% SI and 70% of coverage). We have repeated the procedure for searching homologues of known structure of the human transmembrane domains using an additional methodology based on the program HMMER [9] and considering e-value thresholds instead of SI. Each of the 3838 domain was used for the generation of Hidden Markov profiles, (https://modbase. comp-bio.ucsf.edu/projects/membrane/downloads: Domain Sequences THM). The HMMER program was then used to align the profiles against all the proteins archived in the Protein Data Bank (PDB: http://www.pdb.org). We found at least one template for 1881 domains with the e-value lower than the threshold (<0.01). The analysis of the coverage of the 1881 sequences led to the following results: 215 domains

Fig. (1). RMSD versus % Sequence Identity of the core of membrane proteins. The core of all the membrane proteins found in the Membrane Proteins with Known Structure Database (http://blanco.biomol.uci.edu/mpstruc/listAll/list) was aligned at the structural and sequence level. The relationship between structural similarity (D) and the fraction of mutated residues (H) calculated by performing a least square fit of the data is: $D = 0,71E^{0,93H}$ (black continuous curve).

with 100% coverage, 979 with a coverage greater than 70%, 276 with a coverage between 50% and 70% and 420 domains with a coverage between 50% and 20%. These numbers, obtained by different techniques, show that around 1/3 of the human membrane proteome could be modeled using comparative modeling techniques with different levels of accuracy.

MEMBRANE PROTEIN MODELING: APPLICATION CASES

In order to assess the use of homology modeling techniques in membrane protein characterization we analyzed the literature of the last 5 years. From NCBI-Pubmed we found 408 research articles and 33 reviews. We have analyzed the contributions considering the three main categories of membrane proteins, i.e. transport proteins, receptors and membrane enzymes [1]. For each category we selected representative articles in which homology modeling gives a decisive contribution.

Transporters

Homology modeling technique was used for example, i) to study the conformational changes between the holo and apo physiological states of the ATP-binding cassette (ABC) superfamily of proteins [10, 11], ii) to study the water and glicerol permeability and response to drug inhibitors of aquaporins [12, 13]. The latter was an example in which the combination of computational techniques with experiments, i.e. stopped-flow analysis and inhibitions assays, allowed the description of complex systems. In order to study Cys-loop ligand-gated ion channels like g-amino butyric acid type A receptors (GABAARs) and glycine receptors (GlyRs), modeling data were used to set up mutagenesis experiments aimed at characterizing glycosilation sites found to be altered in disease states of neuronal activity [14-16]. Voltage-gated proton channels (HV1s) homology models combined with electrophysiology experiments, like patch clamp, were used to characterize the open conformation and accessibilities of the channel [17]. Cation channels (Na+, K+, Ca2+) modeling sessions were instead used to characterize the different activation states of the human ether-a-go-go related gene 1 (hERG1) K+ ion channel [18] and of the CaV1.2 calcium channels [19]. The activation mechanisms in cyclic nucleotide channels have been also characterized by the use of a combination of site-directed mutagenesis experiments and homology modeling techniques [20-23]. Using molecular dynamic and metadynamics (MTD), a method that allows exploring multidimensional free energy surfaces (FESs), an alternative Na+ binding site of Sodium-Galactose Transporter (SGLT) symporter protein was predicted [24]. Homology Modeling combined with virtual docking simulations was used to study the interaction between Acid-sensing ASIC channels and toxins [25]. Kranjc [26] applied homology modeling in order to construct a model of arginine rich domain of the calcium-activated anion channel bestrophin and evaluate how specific mutations affect its capacity to bind calcium ions. The works presented in this section represent the power of combined experimental/computational approaches. Indeed, from what regards ion channels, they gave a fundamental contribution to the understanding of the molecular determinants of the gating mechanisms and helped

to characterize the role of residues involved in ion binding at atomic level. In the case of carrier and solute transporters the computational/experimental characterization of the transport mechanisms allowed the gain insights into the large conformational changes that occur upon transportation of solutes.

For recent and extensive reviews on this category see [27-29].

Receptors

Homology modeling was also extensively used in the last years to unravel the functioning and structural features of given transmembrane receptors. For example TLR8, a member of the Toll-like receptors (TLRs) family, were studied by a combination of comparative modeling, molecular dynamics simulations and virtual docking and computational mutagenesis studies. These studies were carried out in order to investigate the interactions of the receptors and its cognate antiviral compound R848 that activates the full TLR8 pathway [30]. Scavenger receptors, expressed by endothelial cells (SREC-I), are membrane protein involved in the endocytosis of lipoproteins: trough homology modeling and experimental procedures the effect of glycosilation on its physiological activities was studied [31].

G protein-coupled receptors (GPCRs) form the largest membrane-bound receptor family expressed by mammalians (encompassing ca. 4% of the protein-coding human genome, 25) and are of paramount importance for pharmaceutical intervention (ca. 40% of currently marketed drugs target GPCRs) [32]. Recently, there was an explosion in the crystallography of GPCRs [33-37] determining a breakthrough in the field of signaling processes. It was of the outmost importance for the scientific community, indeed, the 2012 Nobel Prize in Chemistry was awarded to Brian Kobilka (Stanford University) and Robert Lefkowitz (Duke University) for their structural work on the GPCRs. Anyway, while about 5500 GPCR sequences are public, there are actually only 19 crystallized structures (as reported in the http://blanco.biomol.uci.edu/mpstruc/listAll/list) providing exciting opportunities for structure-based drug design methods that can now use increasingly reliable homology models of GPCR targets [38, 39]. Successful computational models of GPCRs have been used for virtual screening, enriching the rate of ligand hits relative to a random collection of compounds, with hit rates ranging from 3 to 21% [40, 41], comparable to virtual screening success rates with X-ray structures [38]. Furthermore, research aimed at elucidating the underlying principles determining the molecular responsiveness range of GPCRs that mediate senses, such as taste [42, 43] and odor [44] receptors, depends on the ability to build reliable models of the interaction sites. A crucial step in understanding specificity and promiscuity in molecular recognition and structure-based design is to identify the residues that are important for ligand binding. Researchers have successfully applied homology-based of non-rhodopsin GPCRs structure modeling approaches to ligand-binding elucidation [41, 45, 46]. Interestingly, the bigger amount of works based on homology modeling (we found 88 articles by last five years) is dedicated to the rhodopsin-like GPCRs (class A), [47-49] and Taste2 (T2R) GPCRs [43, 50-53]. Several approaches to modeling GPCRs have been described [41], including *ab*

[41], including *ab initio* [42, 54] and template based [40, 49, 55]. The relevance of the articles presented here resides in the fact that the combination of homology modeling and/or molecular docking combined with a rapid growing number of protein structures deposited in the PDB database, together with the availability of functional assays, allowed to represent the ligand-binding interactions at an unprecedented level of detail. This kind of approaches would allow in a near future the intelligent design of pharmaceutical compounds, using a structure-based strategy. For an extensive review on GPCRs see [56].

Recently Malo *et al.* [57] proposed a new method to generate all-atom models of the membrane-spanning part of TM proteins that repacks secondary structure elements of a homology model guided by a ligand and a limited set of experimental and evolutionary restraints.

Membrane Enzymes

Homology modeling techniques were also applied to the characterization of membrane enzymes. Using molecular modeling validated by site-directed mutagenesis experiments, the molecular interaction between F1-ATPase and the glycine-extended form of gastrin G-gly was characterized [58]. The active site residues of the 25-hydroxycholesterol-7α-hydroxylase and oxysterol-7α-hydroxylase CYP7B1, CYP7 family members were characterized by comparative modeling techniques combined with virtual docking experiments and validated using site-directed mutational analysis [59]. These works were the first attempts in which a systematic experimental/computational combined efforts were applied to this family of enzymes, and allowed the identification of fundamental residues involved in ligand/substrate binding.

METHODOLOGIES FOR GENOME-WIDE MODELING APPROACHES

The possibility of obtaining high quality models through the use of comparative modeling techniques will allow the development of automatic tools able to perform genome-wide membrane protein modeling. Indeed, in the last few years several efforts in this direction have been carried out for proteins in general, using two methods: ModPipe [7] and the Modelling and Assessment of ISoforms Through Automated Server (MAISTAS) [60], a tool particularly designed for modelling all the annotated isoform(s) of a protein sequence. For the specific case of membrane proteins, some of us have recently developed a server dedicated to the automated modelling of GPCRs, i.e. the GPCRs Online Modelling and Docking webserver (GOMoDo) (http://molsim.sci.univr.it/gomodo). This web tool performs automatic homology modelling and ligand docking of GPCR receptors. Moreover, in the next future it will be extended to offer the user the possibility of modelling membrane proteins in an automatized way, contributing to the automatic modelling of the entire human membrane proteome.

In the case of the 2/3 of proteins that could not be covered by homology modelling, efforts have been also undertaken by implementing techniques such as fold recognition, threading in particular. Indeed, in the ambit of the Encode project, recently Jones and collaborators have used the Gen-

THREADER [61, 62] method to automatically model the structure of a set of protein variants that could not be modelled by homology modelling techniques [63]. This approach could be extended to the human membrane proteome, until the Structural genomics projects will be able to offer at least a template for each of the families of membrane proteins. Another server is the Bologna Annotation Resource Plus (BAR+) [64] that is based on a large-scale genome cross comparison, taking advantage of a huge pairwise sequence comparison that includes 988 complete proteomes. One of the principal features, is the structure prediction at genome-wide level, using fold recognition methods.

CONCLUSION

Membrane proteins are of the outmost importance for the survival of any living being. A deep insight into the molecular mechanisms underlying their function is thus needed for a complete characterization of the way our cells exchange information with the environment. A detailed annotation of the human membrane proteome implies necessarily, the structural characterization of membrane proteins. As stated in the previous sections, structural biologists still find difficult to overcome the implicit limitations of crystallographic techniques on membrane proteins. These difficulties call for techniques that combine *in vitro* and *in silico* approaches.

Here we have reviewed novel biological applications in which combinations of experimental/computational multi-tasking approaches were shown to be of fundamental importance for obtaining a deep characterization of complex systems that include membrane proteins. In particular, homology modeling represents a powerful instrument because it is applicable to many different biological systems. Indeed, the astonishing improvements in crystallographic techniques combined with comparative modeling techniques will allow the characterization of almost the complete human membrane proteome in the near future.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflicts of interest.

REFERENCES

[1] Almén, M.S.; Nordström, K.J.; Fredriksson, R.; Schiöth, H.B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.,* **2009**, 7, 50.

[2] Tramontano, A.; Cozzetto, D.; Giorgetti, A.; Raimondo, D. *Methods in Molecular Biology, vol. 413: Protein Structure Prediction*, 2nd ed.; Humana Press Inc.: Totowa, New York, **2008**.

[3] Cozzetto, D.; Giorgetti, A.; Raimondo, D.; Tramontano, A. The evaluation of protein structure prediction results. *Mol. Biotechnol.,* **2008**, 39, 1-8.

[4] Chothia, C.; Lesk, A.M. The relation between the divergence of sequence and structure in proteins. *EMBO J.,* **1986**, 5, 823-826.

[5] Giorgetti, A.; Piccoli, S. Knowledge Based Membrane Protein Structure Prediction. From X-Ray Crystallography to Bioinformatics and Back to Molecular Biology. In: *Current Trends in X-Ray Crystallography*, 1st ed.; Chandrasekaran, A.; Eds.; InTech: Rijeka, Croatia, **2011**; pp. 349-364.

[6] Pieper, U.; Webb, B.M.; Barkan, D.T.; Schneidman-Duhovny, D.;

Schlessinger, A.; Braberg, H.; Yang, Z.; Meng, E.C.; Pettersen, E.F.; Huang, C.C.; Datta, R.S.; Sampathkumar, P.; Madhusudhan, M.S.; Sjölander, K.; Ferrin, T.E.; Burley, S.K.; Sali, A. ModBase, a database of annotated comparative protein structure models, and associated resources. *Nucleic Acids Res.*, **2011**, 39, 465-474.

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

[8]   Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis S.; Madden, TL. NCBI BLAST: a better web interface. *Nucleic Acids Res.* **2008**, 36, 5-9.

[9]   Finn, R.D.; Clements, J.; Eddy, S.R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research,* **2011**, 39, 29-37.

[10]   Mornon, J.P.; Lehn, P.; Callebaut, I. Molecular models of the open and closed states of the whole human CFTR protein. *Cell Mol. Life Sci.*, **2009**, 66, 3469-3486.

[11]   Shi, Z.; Tiwari, A.K.; Shukla, S.; Robey, R.W.; Singh, S.; Kim, I.W.; Bates, S.E.; Peng, X.; Abraham, I.; Ambudkar, S.V.; Talele, T.T.; Fu, L.W.; Chen, ZS. Sildenafil reverses ABCB1- and ABCG2- mediated chemotherapeutic drug resistance. *Cancer Res.*, **2011**, 71, 3029-3041.

[12]   Yakata, K.; Tani, K.; Fujiyoshi, Y. Water permeability and characterization of aquaporin-11. *J. Struct. Biol.*, **2011**, 174, 315-320.

[13]   Martins, A.P.; Marrone, A.; Ciancetta, A.; Galán Cobo, A.; Echevarría, M.; Moura, TF; Re, N.; Casini, A; Soveral G. Targeting aquaporin function: potent inhibition of aquaglyceroporin-3 by a gold-based compound. *PLoS One,* **2012**: Available from: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0037435.

[14]   Haeger, S.; Kuzmin, D.; Detro-Dassen, S.; Lang, N.; Kilb, M.; Tsetlin, V.; Betz, H.; Laube, B.; Schmalzing, G. An intramembrane aromatic network determines pentameric assembly of Cys-loop receptors. *Nat. Struct. Mol. Biol.*, **2010**, 17, 90-98.

[15]   Castro, P.A.; Figueroa, M.; Yevenes, G.E.; San Martin, L.S.; Aguayo, LG. The basicproperty of Lys385 is important for potentiation of the human a1 glycine receptor by ethanol. *J. Pharmacol. Exp. Ther.*, **2012**, 340, 339-349.

[16]   Gurba, K.N.; Hernandez, C.C.; Hu, N.; Macdonald, R.L. GABRB3 mutation; G32R, associated with childhood absence epilepsy alters α1β3γ2L γ-aminobutyric acid type A (GABAA) receptor expression and channel gating. *J. Biol. Chem.*, **2012**, 287, 12083-97.

[17]   Kulleperuma, K.; Smith, S.M.; Morgan, D.; Musset, B.; Holyoake, J.; Chakrabarti, N.; Cherny, V.V.; DeCoursey, T.E.; Pomès, R. Construction and validation of a homology model of the human voltage-gated proton channel hHV1. *J. Gen. Physiol.*, **2013**, 141, 445-465.

[18]   Durdagi, S.; Deshpande, S.; Duff, H.J.; Noskov, S.Y. Modeling of open; closed; and open-inactivated states of the hERG1 channel: structural mechanisms of the state-dependent drug binding. *J. Chem. Inf. Model.*, **2012**, 52, 2760-2774.

[19]   Depil, K.; Beyl, S.; Stary-Weinzinger, A.; Hohaus, A.; Timin, E.; Hering, S. Timothy mutation disrupts the link between activation and inactivation in Ca(V)1.2 protein. *J. Biol. Chem.,* **2011**, 286, 31557-31564.

[20]   Giorgetti, A.; Nair, A.V.; Codega, P.; Torre, V.; Carloni P. Structural basis of gating of CNG channels. *FEBS Lett.*, **2005**, 579, 1968-1972.

[21]   Mazzolini, M.; Marchesi, A.; Giorgetti, A.; Torre, V. Gating in CNGA1 channels. *Pflugers Arch., EJP.,* **2010**, 459, 547-555.

[22]   Giorgetti, A.; Carloni, P. Molecular modeling of ion channels: structural predictions. *Curr Opin Chem Biol.*, **2003**, 7, 150-156.

[23]   Giorgetti, A.; Carloni, P.; Mistrik, P.; Torre, V. A homology model of the pore region of HCN channels. *Biophys J.*, **2005**, 89, 932-944.

[24]   Bisha, I.; Laio, A.; Magistrato, A.; Giorgetti, A.; Sgrignani, J. A Candidate Ion-Retaining State in the Inward-Facing Conformation of Sodium/Galactose Symporter: Clues from Atomistic Simulations *J. Chem. Theory Comput.,* **2013**, 9,1240-1246.

[25]   Pietra, F. Docking and MD simulations of the interaction of the tarantula peptide psalmotoxin-1 with ASIC1a channels using a homology model. *J Chem. Inf. Model.*, **2009**, 49, 972-7.

[26]   Kranjc, A.; Grillo, F.W.; Rievaj, J.; Boccaccio, A.; Pietrucci, F.;

Menini, A; Carloni, P.; Anselmi, C. Regulation of bestrophins by Ca2+: a theoretical and experimental study. *PLoS On*e, **2009**: Available from: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0004672

[27]   Ravna, A.W.; Sylte, I. Homology modeling of transporter proteins (carriers and ion channels). *Methods Mol. Biol.*, **2012**, 857, 281-299.

[28]   Palmeira, A.; Sousa, E.; Vasconcelos, M.H.; Pinto, M.; Fernandes, M.X. Structure and ligand-based design of P-glycoprotein inhibitors: a historical perspective. *Curr. Pharm. Des.*, **2012**, 18, 4197-4214.

[29]   Terada, T.; Inui, K. Recent advances in structural biology of peptide transporters. *Curr. Top Membr.*, **2012**, 70, 257-274.

[30]   Govindaraj, R.G.; Manavalan, B.; Basith, S.; Choi, S. Comparative analysis ofspecies-specific ligand recognition in Toll-like receptor 8 signaling: a hypothesis. *PLoS One,* **2011**: Available from: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0025118

[31]   Sano, M.; Korekane, H.; Ohtsubo, K.; Yamaguchi, Y.; Kato, M.; Shibukawa, Y.; Tajiri, M.; Adachi, H.; Wada, Y.; Asahi, M.; Taniguchi, N. N-glycans of SREC-I (scavenger receptor expressed b endothelial cells): essential role for ligand binding, trafficking and stability. *Glycobiology,* **2012**, 22, 714-724.

[32]   Schöneberg, T.; Schulz, A.; Biebermann, H.; Hermsdorf, T.; Römpler, H.; Sangkuhl, K. Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacol. Ther.,* **2004**, 104, 173-206.

[33]   Heilker, R.; Wolff, M.; Tautermann, C.S.; Bieler, M. G-protein-coupled receptor-focused drug discovery using a target class platform approach. *Drug. Discov. Today,* **2009**, 14, 231-240.

[34]   Rosenbaum, D.M.; Rasmussen, S.G.F.; Kobilka, B.K. The structure and function of G-protein-coupled receptors. *Nature,* **2009**, 459, 356-363.

[35]   Topiol, S.; Sabio, M. X-ray structure breakthroughs in the GPCR transmembrane region. *Biochem. Pharmacol.,* **2009**, 78, 11-20.

[36]   Sprang, S.R. Cell signalling: binding the receptor at both ends. *Nature*, **2011**,469, 172-173.

[37]   Xu, F.: Wu, H.; Katritch, V.; Han, G.W.; Jacobson, K.A.; Gao, Z.G.; Cherezov, V.; Stevens, R.C. Structure of an agonist-bound human A2A adenosine receptor. *Science,* **2011**, 332, 322-327.

[38]   Mobarec, J.C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: which structural template to use? *J. Med. Chem.,* **2009**, 52, 5207-5216.

[39]   Yarnitzky, T.; Levit, A.; Niv, M.Y. Homology modeling of G-protein-coupled receptors with X-ray structures on the rise. *Curr Opin. Drug. Discov. Devel.,* **2010**, 13, 317-325.

[40]   Senderowitz, H.; Marantz, Y. G Protein-coupled receptors: target-based *in silico* screening. *Curr. Pharm. Des.,* **2009**, 15, 4049-4068.

[41]   de Graaf, C.; Rognan, D. Customizing G Protein-coupled receptor models for structure-based virtual screening. *Curr. Pharm. Des.,* **2009**, 15, 4026-4048.

[42]   Cui, M.; Jiang, P.; Maillet, E.; Max, M.; Margolskee, R.F.; Osman, R. The heterodimeric sweet taste receptor has multiple potential ligand binding sites. *Curr. Pharm. Des.,* **2006**, 12, 4591-4600.

[43]   Brockhoff, A.; Behrens, M.; Niv, M.Y.; Meyerhof, W. Structural requirements of bitter taste receptor activation. *Proc. Natl. Acad. Sci. U S A.,* **2010**, 107, 11110-11115.

[44]   Reisert, J.; Restrepo, D. Molecular tuning of odorant receptors and its implication for odor signal processing. *Chem. Senses.,* **2009**, 34, 535-545.

[45]   Petrel, C.; Kessler, A.; Dauban, P.; Dodd, R.H.; Rognan, D.; Ruat, M. Positive and negative allosteric modulators of the Ca^{2+}-sensing receptor interact within overlapping but not identical binding sites in the transmembrane domain. *J. Biol. Chem.,* **2004**, 279, 18990-18997.

[46]   Bhattacharya, S.; Subramanian, G.; Hall, S.; Lin, J.; Laoui, A.; Vaidehi, N. Allosteric antagonist binding sites in class B GPCRs: corticotropin receptor 1. *J. Comput. Aided Mol. Des.,* **2010**, 24, 659-674.

[47]   Niv, M.Y.; Skrabanek, L.; Filizola, M.; Weinstein, H. Modeling activated states of GPCRs: the rhodopsin template. *J. Comput. Aided Mol. Des.,* **2006**, 20, 437-448.

[48]   Niv, M.Y.; Filizola, M. Influence of oligomerization on the dynamics of G-protein-coupled receptors as assessed by normal mode analysis. *Proteins,* **2008**, 71, 575-586.

[49]   Ivanov, A.A.; Barak, D.; Jacobson, K.A. Evaluation of homology modeling of G-protein-coupled receptors in light of the A (2A) adenosine receptor crystallographic structure. *J. Med. Chem.,* **2009**,

52, 3284-3292.

[50]   Slack, J.P.; Brockhoff, A.; Batram, C.; Menzel, S.; Sonnabend, C.; Born, S.; Galindo, M.M.; Kohl, S.; Thalmann, S.; Ostopovici-Halip, L.; Simons, C.T.; Ungureanu, I.; Duineveld, K.; Bologa, C.G.; Behrens, M.; Furrer, S.; Oprea, T.I.; Meyerhof, W. Modulation of bitter taste perception by a small molecule hTAS2R antagonist. *Curr. Biol.,* **2010,** 20, 1104-1109.

[51]   Biarnés, X.; Marchiori, A.; Giorgetti, A.; Lanzara, C.; Gasparini, P.; Carloni, P.; Born, S.; Brockhoff, A.; Behrens, M.; Meyerhof, W. Insights into the binding of Phenyltiocarbamide (PTC) agonist to its target human TAS2R38 bitter receptor. *PLoS ONE,* **2010**: Available from: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0012394

[52]   Leguèbe, M.; Nguyen, C.; Capece, L.; Hoang, Z.; Giorgetti, A.; Carloni, P. Hybrid molecular mechanics/coarse-grained simulations for structural prediction of G-protein coupled receptor/ligand complexes. *PLoS One,* **2012**: Available from: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0047332

[53]   Marchiori, A.; Capece, L.; Giorgetti, A.; Gasparini, P.; Behrens, M.; Carloni, P.; Meyerhof, W. Coarse-Grained/Molecular Mechanics of the TAS2R38 Bitter Taste Receptor: Experimentally-Validated Detailed Structural Prediction of Agonist Binding. *PLoS One.*, **2013**: Available from: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0064675

[54]   Vaidehi, N.; Pease, J.E.; Horuk, R. Modeling small molecule-compound binding to G-protein-coupled receptors. *Methods Enzymol.,* **2009**, 460, 263-288.

[55]   Simms, J.; Hall, N.E.; Lam, P.H.; Miller, L.J.; Christopoulos, A.; Abagyan, R.; Sexton, P.M. Homology modeling of GPCRs. *Methods Mol. Biol.,* **2009**, 552, 97-113.

[56]   Levit, A.; Barak, D.; Behrens, M.; Meyerhof, W.; Niv, M.Y. Homology Model-Assisted Elucidation of Binding Sites in GPCRs In: *Membrane Protein Structure and Dynamics: Methods and Protocols (Methods in Molecular Biology)*, 1st ed.; Vaidehi, N.; Klein-Seetharaman, J.; Eds.; Springer Science Business Media: London, **2012**, pp. 179-205.

[57]   Malo, M.; Persson, R.; Svensson, P.; Luthman, K.; Brive, L. Development of 7TM receptor-ligand complex models using ligand-biased, semi-empirical helix-bundle repacking in torsion space: application to the agonist interaction of the humandopamine D2 receptor. *J. Comput. Aided. Mol. Des*., **2013**, 27, 277-291.

[58]   Kowalski-Chauvel, A.; Najib, S.; Tikhonova, I.G.; Huc, L.; Lopez, F.; Martinez, L.O.; Cohen-Jonathan-Moyal, E.; Ferrand, A.; Seva, C. Identification of the F1-ATPase at the cell surface of colonic epithelial cells: role in mediating cell proliferation. *J. Biol. Chem*., **2012**, 287, 41458-41468.

[59]   Siam, A.; Brancale, A.; Simons, C. Comparative modeling of 25-hydroxycholesterol-7α-hydroxylase (CYP7B1): ligand binding and analysis of hereditary spastic paraplegia type 5 CYP7B1 mutations. *J. Mol. Model*., **2012**, 18, 441-453.

[60]   Floris, M.; Raimondo, D.; Leoni, G.; Orsini, M.; Marcatili, P.; Tramontano, A. MAISTAS: a tool for automatic structural evaluation of alternative splicing products. *Bioinformatics*. **2011**, 27, 1625-1629.

[61]   McGuffin, L.J.; Street S.A.; Bryson, K.; Sørensen, S.A.; Jones DT. The Genomic Threading Database: a comprehensive resource for structural annotations of the genomes from key organisms. *Nucleic Acids Res.*, **2004**, 32, 196-199.

[62]   McGuffin, L.J.; Smith, R.T.; Bryson, K.; Sørensen, S.A.; Jones DT. High throughput profile-profile based fold recognition for the entire human proteome. *BMC Bioinformatics*, **2006**, 7, 288.

[63]   Tress, M.L.; Martelli, P.L.; Frankish, A.; Reeves, G.A.; Wesselink, J.J.; Yeats, C.; Olason, P.I.; Albrecht, M.; Hegyi, H.; Giorgetti, A.; Raimondo, D.; Lagarde, J.; Laskowski, R.A.; López, G.; Sadowski, M.I.; Watson, J.D.; Fariselli, P.; Rossi, I.; Nagy, A.; Kai, W.; Størling, Z.; Orsini, M.; Assenov, Y.; Blankenburg, H.; Huthmacher, C.; Ramírez, F.; Schlicker, A.; Denoeud, F.; Jones, P.; Kerrien, S.; Orchard, S.; Antonarakis, S.E.; Reymond, A.; Birney, E.; Brunak, S.; Casadio, R.; Guigo, R.; Harrow, J.; Hermjakob, H.; Jones, D.T.; Lengauer, T.; Orengo, CA.; Patthy, L.; Thornton, JM.; Tramontano, A.; Valencia A. The implications of alternative splicing in the ENCODE protein complement. *Proc Natl Acad Sci U S A.*, **2007**, 104, 5495-5500.

[64]   Piovesan, D.; Martelli, PL.; Fariselli, P.; Zauli, A.; Rossi, I.; Casadio R. BAR-PLUS: the Bologna Annotation Resource Plus for functional and structural annotation of protein sequences. *Nucleic Acids Res.*, **2011**, 39, 197-202.