Published in final edited form as:

Expert Opin Drug Discov. 2008 March ; 3(3): 343–355. doi:10.1517/17460441.3.3.343.

Advances in the Development and Application of Computational Methodologies for Structural Modeling of G-Protein Coupled Receptors

Juan Carlos Mobarec, Eng¹ and Marta Filizola, PhD¹

Juan Carlos Mobarec: ; Marta Filizola: marta.filizola@mssm.edu

¹Department of Structural and Chemical Biology, Mount Sinai School of Medicine, Icahn Medical Institute Building, 1425 Madison Avenue, Box 1677, New York, NY 10029-6574, Tel: 212-241-8634; Fax: 212-849-2456

Abstract

Background—Despite the large amount of experimental data accumulated in the past decade on G-protein coupled receptor (GPCR) structure and function, understanding of the molecular mechanisms underlying GPCR signaling is still far from being complete, thus impairing the design of effective and selective pharmaceuticals.

Objective—Understanding of GPCR function has been challenged even further by more recent experimental evidence that several of these receptors are organized in the cell membrane as homo- or hetero-oligomers, and that they may exhibit unique pharmacological properties. Given the complexity of these new signaling systems, researcher's efforts are turning increasingly to molecular modeling, bioinformatics and computational simulations for mechanistic insights of GPCR functional plasticity.

Methods—We review here current advances in the development and application of computational approaches to improve prediction of GPCR structure and dynamics, thus enhancing current understanding of GPCR signaling.

Results/Conclusions—Models resulting from use of these computational approaches further supported by experiments are expected to help elucidate the complex allosterism that propagates through GPCR complexes, ultimately aiming at successful structure-based rational drug design.

Keywords

bioinformatics; computer simulations; dynamics; GPCRs; membrane proteins; modeling; oligomerization; databases

1. Introduction

G-protein coupled receptors (GPCRs) are among the largest and most important families of integral membrane proteins in fundamental and applied research. Although nearly one half of currently marketed drugs bind to GPCRs [1], only 10% of these receptors have endogenous ligands [2]. GPCRs transmit molecular information from the extracellular (outer) side to the intracellular (inner) side of the cell, thus mediating many intracellular responses. They are characterized by 7 transmembrane (TM) helices connected by extracellular and intracellular

loops, and by an extracellular N-terminus and an intracellular C-terminus (Figure 1). Based on sequence similarity within the seven TMs, GPCRs can be grouped into six families: the rhodopsin family A, the secretin family B, the glutamate receptor family C, the fungal pheromone family D, the cAMP receptor family E, and the frizzled/smoothed receptor family F.

GPCRs can recognize structurally diverse ligands ranging from photons to ions, amino acids, small organic molecules, lipids, peptides, or proteins. The location of the ligand binding pocket is known for many of these receptors [3]. Specifically, small organic molecules are known to bind within the TM bundle, while peptides and proteins interact with the N-terminus and/or the extracellular loop regions. Despite these differences in ligands and ligand-binding sites, the majority of GPCRs can interact and activate one or more of the 16 known α -subunits of heterotrimeric G-proteins. Besides, compelling recent evidence suggests that some GPCRs can also activate G protein-independent signalling pathways [4,5].

Due to limitations in the purification and crystallization of GPCRs, until recently the only known three-dimensional (3D) structures of GPCRs were those corresponding to several inactive forms [6–10] and photoreaction intermediates [11–14] of inactive visual rhodopsin bound to the ground-state chromophore 11-cis-retinal. The very recent crystal structures of human β 2-adrenergic receptor with or without the partial inverse agonist carazolol [15,16] provided the first evidence for similarities and differences between the specialized GPCR rhodopsin for the detection of light and GPCRs for diffusible hormones and neurotransmitters such as β 2-adrenergic receptor. The impact of these diverse GPCR structures in drug discovery and their limitations have recently been discussed in the literature [17]. Even in the presence of the new crystal structures, however, the specific changes that lead to the complete activation and signaling of GPCRs remain mostly unknown.

The traditional mechanism of GPCR signalling refers to a process by which the monomeric receptor undergoes significant conformational changes upon ligand binding. Recent evidence suggests that this process involves one or more conformational intermediates of the receptor [18]. Although much biochemical and biophysical data are consistent with GPCR's ability in both detergent solution and nanoscale planar phospholipid bilayer systems (nanodiscs) to bind and activate G-proteins in a monomeric form [18–23], many recent studies support the hypothesis that G-protein coupling in cell membranes involves the formation of GPCR dimers or oligomers. These complex systems have been suggested to exhibit specific signaling cascades, specific pharmacological properties and specific internalization and recycling properties [24–34], thus complicating our understanding of the signaling mechanism of GPCRs. This new perspective has turned increasingly to molecular modeling, bioinformatics and computational simulations to help predict and explain the dynamics of ligand-GPCR, GPCR-GPCR, and GPCR-G protein (or other proteins of the signaling cascade) interactions.

We review here computational approaches that have recently been applied to GPCRs with the goal of increasing current understanding of receptor structure and function. Specifically, these approaches have provided advances in structural modeling of GPCRs with emphasis on: a) construction of both active and inactive forms of GPCRs, b) identification of structurally and functionally important residues, c) prediction of intermolecular interactions, d) elucidation of dynamic behavior, and e) data mining. Models resulting from use of these computational approaches further refined on the basis of more detailed structural information obtained from experiments are expected to help elucidate the molecular mechanisms at the basis of GPCR function, ultimately leading to the design of more effective and selective pharmaceuticals. While describing the contribution of these computational approaches to address open questions in GPCR structure and function, we also highlight pitfalls and the remaining challenges.

2. Prediction of GPCR structures

Based on experimental evidence that some of the conformational changes occurring in rhodopsin upon activation are similar to those observed in other GPCRs, rhodopsin crystal structures (inactive forms) have often been used as templates in the homology modeling of the transmembrane region of several GPCR subtypes, including family C GPCRs [35]. While use of this template has often been proven valid on the basis of available experimental data [36, 37], the vast amino acid sequence variability within the GPCR superfamily has prompted the development of alternative *de novo* (knowledge-based approaches) and *ab initio* (first-principles approaches) computational methods to build complete inactive and/or activated molecular models of GPCRs without using specific homologous template structures. Results of these methods are summarized in the following sections.

2.1 Comprehensive Modeling

Obtaining accurate 3D models of GPCRs by *de novo* or *ab initio* methods is a multifaceted undertaking. We were among the first investigators [38–40] to develop *de novo* methods to build GPCR models when no rhodopsin crystal structure was available. Briefly, we built bundles of ideal helices using hydrophobicity properties of the receptors and geometric parameters deriving from low resolution cryoelectron microscopy density maps [41,42]. Using similar principles, the Goddard group later on developed methods (MembStruk and HierDock) to predict structures of GPCRs, as well as ligand binding sites and relative binding affinities [43–45]. The MembStruk-predicted structure of bovine rhodopsin deviated from the crystal by 2.84 Å root mean-square deviation (RMSD) in the transmembrane region main-chain atoms. A conformation with a RMSD of 3.7 Å from the native-state structure was predicted using an approach based on evolutionary conservation, hydrophobicity, and intermediate-resolution structures [46]. A similar RMSD value was obtained by prediction of a rhodopsin structure using a method only based on sequence conservation patterns [28]. Slightly better results were obtained by Skolnick's threading assembly refinement (TASSER) method, which was recently used to generate structure predictions for 907 putative GPCRs in the human genome [47], and allowed to build a rhodopsin model with a global C α RMSD from native structure of 4.6 Å, and a RMSD in the transmembrane helix region of 2.1 Å. Much less accurate was the predictive ability of Baker's Rosetta-Membrane *ab-initio* method whose best rhodopsin model had a RMSD of 9.2 Å to the native structure over all residues and a RMSD of 3.8 Å over only 91 core residues [48].

Although the GPCR conformational space is much smaller compared to that sampled by water-soluble proteins, it is still very complicated due to: a) unexpected structural diversity (as revealed by the latest β 2 adrenergic receptor structure [15,16] compared to rhodopsin crystal structure), b) non-conserved non-canonical elements (e.g., proline kinks and other deviations from standard α -helical conformations), c) loops of variable lengths, d) presence of cavities that can accommodate water molecules and/or different ligands, and e) possible influence of interacting GPCR subunits or other proteins of the signaling cascade. Thus, although *de novo* and *ab initio* methods may suggest reasonable TM arrangements of GPCRs, the accuracy of their predictions is limited by the experimental information available. As a result, homology modeling is considered a more reliable technique whenever applicable [49].

2.2 Loop Modeling

Loop regions in GPCRs exhibit low sequence identity and variable length due to insertions and deletions. It was recently suggested that the accuracy of 3D models of membrane proteins decreases rapidly below 30% sequence identity [50]. Thus, homology modeling using the loops of known GPCR structures is unfeasible in most cases. *De novo/ab initio* computational approaches using either coarse-grained backbone dihedral sampling [51] or Monte Carlo (MC)

simulations in a temperature annealing protocol combined with a scaled collective variables (SCV) technique [52,53] have recently been shown to accurately predict loop regions of GPCRs with decreasing performance at increasing loop lengths. For instance, “global” RMSD values (RMSDs calculated for all heavy atoms of the loop backbones overlapped over the corresponding helical stems of the 1GZM or 1U19 rhodopsin crystal structures) of the Nikiforovich’s low-energy conformers [51] of the individual intracellular (IC) and extracellular (EC) loops with respect to 1GZM (or 1U19; RMSD values are in parentheses) were as follows: 1.9 (2.0) Å for IC1 (7 nonstem residues; amino acids 65–71); 2.5 (1.8) Å for IC2 (10 nonstem residues; 140–149); 5.0 (5.0) Å for IC3 (20 nonstem residues; 226–245); 2.1 (2.0) Å for EC1 (8 nonstem residues; 100–107); 4.7 (4.9) Å for EC2 loop (27 nonstem residues; 173–199); and 1.0 (1.3) Å for EC3 (8 nonstem residues; 278–285). Slightly better results were obtained by Mehler and coworkers [52,53] for the shortest loops of rhodopsin (IC1, EC1, and EC3). Specifically, these studies allowed to identify low-energy conformers with RMSD values (Ca atoms only) of 0.5, 0.5, and 1.1 Å for IC1, EC1, and EC3 loops, respectively. Also reasonably good results were obtained for the long IC3 loop of bovine rhodopsin using multicanonical molecular dynamics [54], which yielded RMSD values as small as 3.6 Å (“local” RMSD value obtained by overlapping all residues in loop fragment 227–244 to the x-ray structure of the 1HZX entry in the PDB where the missing fragment 236–240 was artificially restored).

Although *de novo/ab initio* predictions for GPCR loops shorter than 10 residues have been quite reliable when using reasonable physical-chemical force fields incorporating descriptions of a water rather than vacuum environment, prediction of longer and interacting loops is still a very challenging task. The reasons for the failure of current *de novo/ab initio* approaches in predicting long loops stem from the difficulty in carrying out sufficiently complete searches of their conformational spaces, as well as insufficiently accurate force fields. Several groups, including ours [55], are currently trying to find new more effective ways to improve structural characterization of long loops of GPCRs in both their inactive and active forms.

2.3 Activated Models

Ligand-induced activation of GPCRs results in multiple allosteric conformational changes that propagate throughout the receptor structure, ultimately triggering different signaling cascades [18]. In contrast to inactive states of GPCRs, no crystal structures of active states are available to date. Given the absence of experimental structural information, several investigators, including ourselves, have applied computational strategies to predict activated models of GPCRs [51,56–59].

Specifically, we used activation-specific structural data available for rhodopsin-like GPCRs (e.g., site-directed spin-labeling, cysteine accessibility, infrared (IR) spectroscopy, site-directed crosslinking, and engineered zinc binding studies) as structural constraints and a simulated annealing molecular dynamics technique to build a number of different active state models of rhodopsin starting from its known inactive structure. These models suggest that TM3, TM5 and TM6 play an important role in activation. A similar approach using distance restraints and molecular dynamics simulations had been used by Gouldson and coworkers [59] early on to build models of the active forms of rhodopsin and the β 2-adrenergic receptor. The main changes in their active model of rhodopsin compared to the inactive one involved TM4, TM5, TM6 and TM7. Isin and coworkers [57] recently proposed a rhodopsin activated model using an elastic network normal mode analysis in conjunction with experimental data. This model suggested the presence of a global hinge site near the retinal-binding pocket, which might represent the originating site for signal transduction to both cytoplasmic and extracellular ends. Using simplified energy calculations, Nikiforovich and coworkers [51] searched for sterically and energetically reasonable rearrangements of the TM regions of rhodopsin upon

retinal isomerization. Comparison of the resulting different models with site-directed spin-label experimental data allowed to identify the most plausible conformation of the TM bundle, which was completed with loops, N- and C-terminal fragments. A comparative analysis of molecular dynamics studies of several normal and/or mutated GPCR subtypes in their free and/or ligand-bound forms was also used by Fanelli and coworkers to infer about structural features that are unique to specific conformational states [60]. Specifically, weakening of the salt-bridge formed by the highly conserved arginine of the E/DRY motif, and the increase in solvent accessibility of the cytosolic interface between helices 3 and 6 were found to be largely responsible for the difference between inactive and active states of GPCRs.

Although the rhodopsin activated models generated to date appear to satisfy most of the experimental data known for GPCRs, novel predictions deriving from their analyses still await experimental validation. Moreover, it remains to be determined whether or not all GPCRs share the same activated forms.

3. Structurally and functionally important residues

Amino acid conservation and/or co-variation of sequence positions have been used as an indication of strong evolutionary pressure at specific GPCR sites, thus suggesting these positions as potential targets for structural and functional studies [34,61,62]. Most of the identified positions had known functional roles (e.g., residues of the E/DRY and NPXXY motifs), thus demonstrating the accuracy of these methods. In addition to evolutionary-based methods, an analysis of flexibility and rigidity of rhodopsin amino-acids using the FIRST (Floppy Inclusion and Rigid Substructure Topography) method recently allowed to identify residues that may be important for stability and folding of rhodopsin, as well as other GPCRs [63]. By simulating rhodopsin thermal unfolding by systematically breaking hydrogen bonds [64], this method suggested that the C110-C187 disulfide bridge, the residues at interface of TM and extracellular (EC) domains, and the rhodopsin residues in the retinal binding pocket are part of the stability core of this prototype GPCR. Specifically, these residues are: F103 (EC2), C110(3.25 according to Ballesteros & Weinstein's generic numbering [65]), E113 (3.28), G114(3.29), R177(IC2), P180(IC2), C185(IC2), S186(IC2), and C187(IC2) [63]. *In vitro* mutational studies confirmed that 90% of the rhodopsin residues predicted by FIRST caused complete or partial protein unfolding upon mutation. Whether these results can be generalized to other GPCRs is still an open question.

4. Intermolecular interactions

It is largely accepted that GPCRs are part of a complex allosteric machinery regulated by protein-protein interactions with the same or different copies of GPCR subtypes as well as other proteins of the signaling cascade. Thus, computational methods designed to help identify protein-protein interaction interfaces have become more and more popular in their application to GPCRs. These methods can be divided in "sequence-based methods" (based on sequence and genomic information), and "structure-based/docking methods" (based on a variety of selection criteria and exhaustive searches of the interaction space). Many of these methods with considerations of their accuracy and reliability have been reviewed recently [34]. We review here the results of more recent developments and applications of sequence- and structure-based methods to predict protein-protein interactions of GPCRs with G-proteins, as well as GPCRs with other GPCRs. These results set the basis for current structural modeling of GPCR complexes.

4.1 Sequence-based methods

Several labs, including ours, have used evolutionary-based bioinformatic methods to predict the dimerization/oligomerization interfaces of GPCRs (see [34,66] for recent reviews), as well

as GPCR residues at the interface with G-proteins [67–69]. A compilation of the most frequently predicted residues by these methods at the interface between GPCRs point to specific regions of TMs 4–6 [34]. In contrast, the most frequently predicted residues at the interface between GPCRs and G-proteins are within the intracellular loops. A recent analysis of multiple sequence alignments of GPCRs of known coupling selectivity [67] has also allowed to identify residues that may be responsible for this selectivity. Of note, these studies showed that: 1) most residues involved in coupling selectivity are located at the intracellular loops; 2) the occurrence of positively/negatively charged amino acids of the characteristic residues varies depending on the G-protein coupling selectivity, and 3) some characteristic residues are located far from the GPCR/G-protein binding, near the extracellular terminus of transmembrane helices. Since these methods rely on the statistical significance of the dataset of sequences used to produce multiple sequence alignments, their application may be limited in the case of some GPCRs whose sequence is unknown for many organisms.

4.2 Structure-based/docking methods

Structure-based/docking methods predict the spatial interaction between two input molecules by matching local complementary features on the protein surfaces. Geometric complementarity, which had first been implemented in the GRAMM computer algorithm to model the interactions of transmembrane helices of GPCRs, thus serving as a guide to build 3D models of receptor bundles [70], has also been used to build plausible configurations of GPCR oligomers [71]. Among the 100 dimer models generated by the GRAMM algorithm for 5HT4 receptor oligomers, 40% of the most favorable complexes exhibited TM2 and TM4 at the interface, and the other 40% were TM4–6/TM4–6 complexes. Electrostatic and shape complementarities between the crystal structures of dark rhodopsin and heterotrimeric transducin (Gt) were used to sample the roto-translational space of GPCRs with respect to G-proteins, and to obtain so-called “precoupled models”. These studies showed that dark adapted rhodopsin can bind transducin in spite of variations in the rhodopsin C-terminal length and orientation [56]. More recent docking studies by the same group between heterotrimeric transducin and monomeric, dimeric and tetrameric rhodopsin have proposed that monomeric dark rhodopsin holds the molecular determinants for transducin recognition, favoring a complex with a 1:1 stoichiometry [72].

Molecular models of the complex between active rhodopsin and the heterotrimeric G protein complex have also appeared in recent literature [73–75]. These models do not share many commonalities with “precoupled models”. For instance, a recent model by Nikiforovich et al. differs from the “precoupled model” of Fanelli & Dell’Orco by a -45° rotation of Gt $\alpha\beta\gamma$ around the axis normal to the membrane surface. Although all these complex models include structural features established experimentally, their complete validation has yet to come.

5. Dynamic Behavior

Computational approaches using both standard molecular dynamics simulations and coarse-grained models recently allowed to infer about the dynamic behavior of GPCRs in both monomeric and oligomeric arrangements. The results of these studies are summarized below.

5.1 Molecular dynamics simulations

Results of several simulations of rhodopsin monomer in realistic explicit membrane environments with and without cholesterol are available in the literature [76–83]. Members of the Blue Gene group at IBM have recently contributed long MD simulations of rhodopsin in a bilayer composed of 1-stearoyl-2-docosahexaenoyl-phosphatidylcholine (SDPC), 1-stearoyl-2-docosahexaenoyl-phosphatidylethanolamine (SDPE), and cholesterol. The results of these simulations helped explain experimental observations such as the stabilizing effect of

cholesterol on the inactive state of rhodopsin. The same group [83] also carried out 26 simulations of rhodopsin in a realistic membrane setting for the duration of 100 ns each, and demonstrated that even 100 ns of simulation were not sufficient to sample the fluctuations of most loops of rhodopsin with statistical certainty. Kong and Karplus [79] recently determined a network involving the highly conserved motifs D(E)RY and NPxxY that couples retinal to the cytoplasmic interface based on the interaction-correlation matrix calculated by MD simulations of rhodopsin in an explicit membrane environment. They also suggested that the major signal-transduction pathway of rhodopsin involves side chains of TM6 and TM7 based on the results of MD simulations with restraints on each transmembrane helix.

In light of recent experimental evidence suggesting that GPCRs are organized in the cell membrane as functional homo- or hetero-oligomers, we carried out nanosecond time-scale MD simulations of a rhodopsin dimer and compared the results to the monomer simulated in the same type of bilayer membrane model composed of an equilibrated unit cell of hydrated palmitoyl-oleoyl-phosphatidyl choline (POPC) [84]. A combined essential dynamics analysis of the MD results of the rhodopsin monomer and dimer revealed an intrinsic asymmetry of the system, which was accentuated at the dimerization interface.

Although consonant with several experimental observations, the results of molecular dynamics simulations are limited by the system size and sub-microsecond time scales accessible with current computer hardware and algorithms. The following section summarizes the results of alternative classes of methods that have recently been applied to study the motion of GPCR complexes.

5.2 Coarse-grained models

We recently studied motion of monomeric and oligomeric arrangements of rhodopsin [58] using an approximate - yet powerful - normal mode analysis (NMA) technique termed elastic network model (ENM) [85]. Specifically, we studied the effect of oligomerization on the dynamics and inter-residue correlations of rhodopsin monomer by comparing the lowest frequency normal modes of rhodopsin monomer with the lowest frequency normal modes of the protein dimeric and tetrameric arrangements derived from inferences from atomic force microscopy [86] and cross-linking data [87]. Results from these simulations indicated a profound effect of oligomerization on the dynamics of the rhodopsin monomer, which was mainly manifested at interfacial regions. These studies also suggested that rigid body rotation of interacting TM4s along their own helical axes is an unlikely mechanism for the experimentally demonstrated conformational rearrangement of the GPCR dimeric interface occurring upon activation [87]. In contrast, protomer displacement (i.e. clockwise rotation of each protomer involved in intra-dimeric interaction along the membrane axis as viewed from the cytoplasm) and protomer exchange (i.e. clockwise rotation of each protomer involved in intra-dimeric interaction followed by their sliding within the array as viewed from the cytoplasm) resulted to be equally feasible dynamic motions, but exhibiting differential motional changes.

To investigate the effect of the phospholipid bilayer membrane on rhodopsin spontaneous self-assembly, another coarse-grained molecular dynamics model termed CGMD recently allowed for computationally efficient calculations of the structure and dynamics of rhodopsin assemblies for larger time and length scales than accessible to atomistic models [88]. Specifically, systems with up to 16 rhodopsin molecules at a protein-to-lipid ratio of 1:100 in four different phospholipid environments were simulated during 8 microseconds. A local membrane deformation was proposed to play a key role in protein-protein association, and contact zones involving helices TM1/TM2/H8, TM4/TM5, and TM6/TM7 appeared by the end of the 6–8-microsecond simulations.

Albeit intriguing, inferences from all the studies described above still await experimental validation.

6. Dedicated databases and servers

As a proof of the biological and physiological importance of GPCRs, several dedicated websites (either databases or web server tools) have become available in the public domain over the years. Table 1 lists a number of dedicated free websites that manage useful information for structural modeling of GPCRs. This list includes information about the website names, contents, URLs, and date of last update/release whenever available. The type of information that these websites manage can be divided into seven main categories (depicted as circles in Figure 2): Sequences, Mutations, Ligands, Structures/Models, Signaling Pathways, Interacting Proteins, and Specialized Repositories. As shown in Figure 2, there is very little overlap between the different circles, as an indication of poor data integration among current websites on GPCRs. Given the importance of GPCRs in many biological processes, databases that integrate information from different repositories of these receptors, their interacting partners, tissue distribution, and mutations are more and more desirable. For instance, integrated documentation resources such as InterPro [89] and PathGuide [90] for protein families, domains, sites, or biological pathways are very valuable since they help advance biological knowledge by deriving protein signatures from the combination of several databases that use different methodologies and biological information on well-characterized proteins, including GPCRs.

A more dedicated, yet integrated database is the recently available GPCR-Ligand Database GLIDA [91], which correlates information between GPCRs and their ligands. Other recent integration-promoting databases related to GPCRs are RINGdb [92] and the GPCR-NaVa (GPCR Natural Variants) database [93]. RINGdb is a relational database of GPCRs and the regulators of G-protein signaling (RGS), a family of proteins that share a conserved signature domain (RGS domain) that binds directly to activated G subunits to modulate G-protein signaling. The GPCR-NaVa database integrates information on natural variants that occur in humans from databases and scientific papers.

Classical GPCR databases such as GPCRDB [94] and tinyGRAP [95] have been designed with a strong emphasis on monomers. These databases include sequence information, alignments of monomers, visualization of monomer units, and mutation information for receptor monomers. The recent explosion of information about GPCR oligomerization, however, has recently inspired the development of a web-based information system that will store computational and experimental information about GPCR oligomers. This GPCR-Oligomerization Knowledge Base (GPCR-OKB) is currently under construction in our lab, and will be linked to the reputable GPCRDB [94] for information about GPCR monomers. We recently completed a formal description (ontology) of the information concepts required to describe GPCR-OKB and of the relationships between these concepts [32]. This ontology was designed in collaboration with experimental experts in the field of GPCR oligomerization, who also provide a quality control check for GPCR-OKB.

In addition to databases, several dedicated web-based tools have been designed to improve structural knowledge of GPCRs. The most recent ones either allow to screen for new members of the GPCR superfamily, or to predict coupling selectivity. For instance, PRED-GPCR [96] is a classification server that uses signature profile hidden Markov models (HMMs) to screen sequences against well characterized GPCR families. New signatures are added as they become available, and the classification system is re-trained every year. The G-protein and Receptor Interaction Feature Finding Instrument (GRIFFIN) [97] is a server that predicts GPCR and G-protein coupling selectivity based on a support vector machine (SVM) and HMMs with high

sensitivity and specificity (>85% on average). GPCRclass [98] is a specialized server for the identification and classification of members from the amine sub-family receptors. The classification method used by GPCRclass is based on amino acid and dipeptide composition, and uses SVMs as classifiers.

7. Conclusion

Supporting evidence that GPCRs are organized in the cell membrane as oligomers is accumulating at a fast pace. This new perspective further complicates current understanding of the molecular mechanisms underlying GPCR function, thus making researchers turn increasingly to computational approaches to generate new valuable testable hypotheses. Recent advances in the development and application of computational approaches for structural modeling of GPCRs have focused on: 1) comprehensive modeling of active and inactive forms of these receptors, 2) identification of structurally and functionally important residues, including residues at interfaces, 3) system dynamics, and 4) data mining. Despite the reasonably accurate models generated by these approaches, vastly validated on the basis of available experimental data, novel predictions are still awaiting confirmation. Thus, more detailed structural information obtained from experiments is needed to complete validation of structural models of these receptors deriving from computation and to provide new insights towards the rational design of more effective and selective pharmaceuticals.

8. Expert opinion

The combination of experimental explorations of GPCRs with computational modeling and simulation studies is continuing to play an increasingly central role in modern drug discovery. Figure 3 shows the information flow between experimental data and computational methods in research on GPCR structure and dynamics with directional or bi-directional arrows indicating the relationships between the different modules. Despite the wealth of data generated by experimental and computational approaches, understanding of GPCR structure, dynamics, and function is still very limited. On the one hand, the discovery that GPCRs may function as homo- or hetero-oligomers has complicated the view of GPCR signaling. In fact, unique signaling units with unexpected combinations of pharmacological properties come now into play suggesting entirely new ways of developing successful drugs. On the other hand, the limited high resolution structural information available has hampered knowledge and understanding of GPCR features. Indeed, the presence of only two high resolution GPCR structures, i.e. rhodopsin and $\beta 2$ adrenergic receptor, can offer no more than a narrow view of the structural space of these receptors.

Notwithstanding the reduced conformational space of GPCRs in comparison to water-soluble proteins, their architectures may not be as simple as initially thought. In fact, the new crystal structure of $\beta 2$ adrenergic receptor has revealed unexpected structural diversities in the ligand binding site and other regions of this receptor compared to the rhodopsin crystal structures. These structural differences highlight the challenges in using a specific GPCR structure as a template model for the large GPCR family. However, while *de novo* and *ab initio* computational approaches could be used as alternative approaches to build complete inactive and activated molecular models of GPCRs without using specific homologous template structures, homology modeling, when applicable, may reach higher levels of performance and reliability as largely demonstrated for other membrane proteins. The reason is that current *de novo* and *ab initio* programs mostly use energy-based scoring functions originally developed for water-soluble proteins, and are therefore of limited practical relevance.

As more detailed experimental structural information will expand our knowledge of GPCRs, better methods will be developed that will allow more accurate *de novo* and *ab initio* structure

predictions of these receptors. In particular, information is needed to address the following questions: 1) Do homologous GPCRs share exactly the same structural and functional elements? 2) Do different GPCRs exhibit the same dimerization/oligomerization interfaces? 3) What conformational changes do GPCR oligomers undergo upon activation? 4) What are the elements that determine the specificity of interaction of GPCR oligomers with G-proteins or other proteins of the signaling cascade? 5) How do different ligands trigger different signaling cascades? 6) What are the molecular determinants responsible for GPCR allostery and modulation? 7) What is the effect of the environment (e.g. lipid, other proteins, etc.) on GPCR dynamics and function? We anticipate that greater integrated efforts by the experimental and computational communities focusing on the properties of GPCR assemblies rather than monomers are needed to answer these questions, and to advance knowledge on GPCR structure and dynamics. This considerable undertaking is justified by the prospect of achieving a better understanding of the allosteric mechanisms that mediate GPCR function, which is essential to design effective and selective pharmaceuticals.

Bibliography

1. Drews J. Drug discovery: a historical perspective. *Science* 2000;287:1960–1964. [PubMed: 10720314]
2. Vassilatis DK, Hohmann JG, Zeng H, Li F, Ranchalis JE, Mortrud MT, et al. The G protein-coupled receptor repertoires of human and mouse. *Proc Natl Acad Sci U S A* 2003;100:4903–4908. [PubMed: 12679517]
3. Ji TH, Grossmann M, Ji I. G protein-coupled receptors. I. Diversity of receptor-ligand interactions. *J Biol Chem* 1998;273:17299–17302. [PubMed: 9651309]
4. Azzi M, Charest PG, Angers S, Rousseau G, Kohout T, Bouvier M, et al. Beta-arrestin-mediated activation of MAPK by inverse agonists reveals distinct active conformations for G protein-coupled receptors. *Proc Natl Acad Sci U S A* 2003;100:11406–11411. [PubMed: 13679574]
5. Lefkowitz RJ, Pierce KL, Luttrell LM. Dancing with different partners: protein kinase a phosphorylation of seven membrane-spanning receptors regulates their G protein-coupling specificity. *Mol Pharmacol* 2002;62:971–974. [PubMed: 12391258]
6. Teller DC, Okada T, Behnke CA, Palczewski K, Stenkamp RE. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry* 2001;40:7761–7772. [PubMed: 11425302]
7. Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc Natl Acad Sci U S A* 2002;99:5982–5987. [PubMed: 11972040]
8. Li J, Edwards PC, Burghammer M, Villa C, Schertler GF. Structure of bovine rhodopsin in a trigonal crystal form. *J Mol Biol* 2004;343:1409–1438. [PubMed: 15491621]
9. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J Mol Biol* 2004;342:571–583. [PubMed: 15327956]
10. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 2000;289:739–745.745 [PubMed: 10926528]
••The first crystal structure of a GPCR.
11. Salom D, Lodowski DT, Stenkamp RE, Le Trong I, Golczak M, Jastrzebska B, et al. Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc Natl Acad Sci U S A* 2006;103:16123–16128. [PubMed: 17060607]
12. Nakamichi H, Buss V, Okada T. Photoisomerization mechanism of rhodopsin and 9-cis-rhodopsin revealed by x-ray crystallography. *Biophys J* 2007;92:L106–L108. [PubMed: 17449675]
13. Nakamichi H, Okada T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc Natl Acad Sci U S A* 2006;103:12729–12734. [PubMed: 16908857]
14. Sekharan S, Sugihara M, Weingart O, Okada T, Buss V. Protein assistance in the photoisomerization of rhodopsin and 9-cis-rhodopsin--insights from experiment and theory. *J Am Chem Soc* 2007;129:1052–1054. [PubMed: 17263385]

15. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* 2007;318:1258–1265.1265 [PubMed: 17962520] ••First evidence of differences and similarities between the specialized visual rhodopsin and GPCRs for diffusible hormones and neurotransmitters.
16. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, et al. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* 2007;450:383–387. [PubMed: 17952055]
17. Kobilka B, Schertler GFX. New G-protein-coupled receptor crystal structures: insights and limitations. *Trends in Pharmacological Sciences* 2008;29:79–83. [PubMed: 18194818]
18. Whorton MR, Bokoch MP, Rasmussen SG, Huang B, Zare RN, Kobilka B, et al. A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proc Natl Acad Sci U S A* 2007;104:7682–7687. [PubMed: 17452637]
19. Chabre M, le Maire M. Monomeric G-protein-coupled receptor as a functional unit. *Biochemistry* 2005;44:9395–9403. [PubMed: 15996094]
20. Bayburt TH, Leitz AJ, Xie G, Oprian DD, Sligar SG. Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *J Biol Chem* 2007;282:14875–14881. [PubMed: 17395586]
21. Ernst OP, Gramse V, Kolbe M, Hofmann KP, Heck M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. *Proc Natl Acad Sci U S A* 2007;104:10859–10864. [PubMed: 17578920]
22. Leitz AJ, Bayburt TH, Barnakov AN, Springer BA, Sligar SG. Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology. *Biotechniques* 2006;40:601–602. [PubMed: 16708760]04, 06, passim.
23. Jastrzebska B, Fotiadis D, Jang GF, Stenkamp RE, Engel A, Palczewski K. Functional and structural characterization of rhodopsin oligomers. *J Biol Chem* 2006;281:11917–11922. [PubMed: 16495215]
24. Pin JP, Neubig R, Bouvier M, Devi L, Filizola M, Javitch JA, et al. International Union of Basic and Clinical Pharmacology. LXVII. Recommendations for the recognition and nomenclature of G protein-coupled receptor heteromultimers. *Pharmacol Rev* 2007;59:5–13.13 [PubMed: 17329545] •The first rules to discern the type of information available on GPCR oligomers.
25. Milligan G. A day in the life of a G protein-coupled receptor: the contribution to function of G protein-coupled receptor dimerization. *Br J Pharmacol.* 2007
26. Kent T, McAlpine C, Sabetnia S, Presland J. G-protein-coupled receptor heterodimerization: assay technologies to clinical significance. *Curr Opin Drug Discov Devel* 2007;10:580–589.
27. Javitch JA. The ants go marching two by two: oligomeric structure of G-protein-coupled receptors. *Mol Pharmacol* 2004;66:1077–1782. [PubMed: 15319448]
28. Park Y, Helms V. Assembly of transmembrane helices of simple polytopic membrane proteins from sequence conservation patterns. *Proteins* 2006;64:895–905. [PubMed: 16807902]
29. Prinster SC, Hague C, Hall RA. Heterodimerization of g protein-coupled receptors: specificity and functional significance. *Pharmacol Rev* 2005;57:289–298. [PubMed: 16109836]
30. Terrillon S, Bouvier M. Roles of G-protein-coupled receptor dimerization. *EMBO Rep* 2004;5:30–34. [PubMed: 14710183]
31. Pflieger KD, Eidne KA. Monitoring the formation of dynamic G-protein-coupled receptor-protein complexes in living cells. *Biochem J* 2005;385:625–637. [PubMed: 15504107]
32. Skrabanek L, Murcia M, Bouvier M, Devi L, George SR, Lohse MJ, et al. Requirements and ontology for a G protein-coupled receptor oligomerization knowledge base. *BMC Bioinformatics* 2007;8:177. [PubMed: 17537266]
33. Bulenger S, Marullo S, Bouvier M. Emerging role of homo- and heterodimerization in G-protein-coupled receptor biosynthesis and maturation. *Trends Pharmacol Sci* 2005;26:131–137. [PubMed: 15749158]
34. Filizola M, Weinstein H. The study of G-protein coupled receptor oligomerization with computational modeling and bioinformatics. *Febs J* 2005;272:2926–2938. [PubMed: 15955053]
35. Binet V, Duthey B, Lecaillon J, Vol C, Quoyer J, Labesse G, et al. Common structural requirements for heptahelical domain function in class A and class C G protein-coupled receptors. *J Biol Chem* 2007;282:12154–12163. [PubMed: 17310064]

36. Ballesteros JA, Shi L, Javitch JA. Structural mimicry in G protein-coupled receptors: implications of the high-resolution structure of rhodopsin for structure-function analysis of rhodopsin-like receptors. *Mol Pharmacol* 2001;60:1–19. [PubMed: 11408595]
37. Visiers I, Ballesteros JA, Weinstein H. Three-dimensional representations of G protein-coupled receptor structures and mechanisms. *Methods Enzymol* 2002;343:329–371. [PubMed: 11665578]
38. Filizola M, Perez JJ, Carteni-Farina M. BUNDLE: a program for building the transmembrane domains of G-protein-coupled receptors. *J Comput Aided Mol Des* 1998;12:111–118. [PubMed: 9690171]
39. Filizola M, Carteni-Farina M, Perez JJ. Modeling of the 3D structure of rhodopsin using a de novo approach to build G-protein coupled receptors. *Journal of Physical Chemistry* 1999;103:2520–2527.
40. Alkorta I, Du P. Sequence divergence analysis for the prediction of seven-helix membrane protein structures: II A 3-D model of human rhodopsin. *Protein Eng* 1994;7:1231–1238. [PubMed: 7855138]
41. Schertler GF, Hargrave PA. Projection structure of frog rhodopsin in two crystal forms. *Proc Natl Acad Sci U S A* 1995;92:11578–11582. [PubMed: 8524807]
42. Schertler GF, Villa C, Henderson R. Projection structure of rhodopsin. *Nature* 1993;362:770–772. [PubMed: 8469290]
43. Trabanino RJ, Hall SE, Vaidehi N, Floriano WB, Kam VW, Goddard WA 3rd. First principles predictions of the structure and function of g-protein-coupled receptors: validation for bovine rhodopsin. *Biophys J* 2004;86:1904–1921. [PubMed: 15041637]
44. Floriano WB, Vaidehi N, Goddard WA 3rd, Singer MS, Shepherd GM. Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor. *Proc Natl Acad Sci U S A* 2000;97:10712–10716. [PubMed: 11005853]
45. Vaidehi N, Floriano WB, Trabanino R, Hall SE, Freddolino P, Choi EJ, et al. Prediction of structure and function of G protein-coupled receptors. *Proc Natl Acad Sci U S A* 2002;99:12622–12627. [PubMed: 12351677]
46. Fleishman SJ, Harrington S, Friesner RA, Honig B, Ben-Tal N. An automatic method for predicting transmembrane protein structures using cryo-EM and evolutionary data. *Biophys J* 2004;87:3448–3459. [PubMed: 15339802]
47. Zhang Y, Devries ME, Skolnick J. Structure modeling of all identified G protein-coupled receptors in the human genome. *PLoS Comput Biol* 2006;2:e13. [PubMed: 16485037]
48. Yarov-Yarovoy V, Schonbrun J, Baker D. Multipass membrane protein structure prediction using Rosetta. *Proteins* 2006;62:1010–1025. [PubMed: 16372357]
49. Punta M, Forrest LR, Bigelow H, Kernysky A, Liu J, Rost B. Membrane protein prediction methods. *Methods* 2007;41:460–474. [PubMed: 17367718]
50. Forrest LR, Tang CL, Honig B. On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys J* 2006;91:508–517. [PubMed: 16648166]
51. Nikiforovich GV, Marshall GR. Modeling flexible loops in the dark-adapted and activated states of rhodopsin, a prototypical G-protein-coupled receptor. *Biophys J* 2005;89:3780–3789. [PubMed: 16199504]
52. Mehler EL, Hassan SA, Kortagere S, Weinstein H. Ab initio computational modeling of loops in G-protein-coupled receptors: lessons from the crystal structure of rhodopsin. *Proteins* 2006;64:673–690. [PubMed: 16729264]
53. Mehler EL, Periole X, Hassan SA, Weinstein H. Key issues in the computational simulation of GPCR function: representation of loop domains. *J Comput Aided Mol Des* 2002;16:841–853. [PubMed: 12825797]
54. Watanabe YS, Fukunishi Y, Nakamura H. Modelling of third cytoplasmic loop of bovine rhodopsin by multicanonical molecular dynamics. *J Mol Graph Model* 2004;23:59–68. [PubMed: 15331054]
55. Bortolato A, Bandyopadhyay D, Mehler EL, Filizola M. Improving Prediction of G-Protein Coupled Receptor Loops. *Biophysical Journal*, POS 1974. 2008
56. Fanelli F, Dell'Orco D. Rhodopsin activation follows precoupling with transducin: inferences from computational analysis. *Biochemistry* 2005;44:14695–14700. [PubMed: 16274216]
57. Isin B, Rader AJ, Dhiman HK, Klein-Seetharaman J, Bahar I. Predisposition of the dark state of rhodopsin to functional changes in structure. *Proteins* 2006;65:970–983. [PubMed: 17009319]

58. Niv MY, Skrabanek L, Filizola M, Weinstein H. Modeling activated states of GPCRs: the rhodopsin template. *J Comput Aided Mol Des* 2006;20:437–448. [PubMed: 17103019]
59. Gouldson PR, Kidley NJ, Bywater RP, Psaroudakis G, Brooks HD, Diaz C, et al. Toward the active conformations of rhodopsin and the beta2-adrenergic receptor. *Proteins* 2004;56:67–84. [PubMed: 15162487]
60. Fanelli F, De Benedetti PG. Inactive and active states and supramolecular organization of GPCRs: insights from computational modeling. *J Comput Aided Mol Des* 2006;20:449–461. [PubMed: 17009093]
61. Suel GM, Lockless SW, Wall MA, Ranganathan R. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat Struct Biol* 2003;10:59–69. [PubMed: 12483203]
62. Ye K, Lameijer EW, Beukers MW, Ijzerman AP. A two-entropies analysis to identify functional positions in the transmembrane region of class A G protein-coupled receptors. *Proteins* 2006;63:1018–1030. [PubMed: 16532452]
63. Rader AJ, Anderson G, Isin B, Khorana HG, Bahar I, Klein-Seetharaman J. Identification of core amino acids stabilizing rhodopsin. *Proc Natl Acad Sci U S A* 2004;101:7246–7251. [PubMed: 15123809]
64. Rader AJ, Hespenheide BM, Kuhn LA, Thorpe MF. Protein unfolding: rigidity lost. *Proc Natl Acad Sci U S A* 2002;99:3540–3545. [PubMed: 11891336]
65. Ballesteros, JA.; Weinstein, H. San Diego, CA: Academic Press; 1995. Integrated methods for the construction of three dimensional models and computational probing of structure function relations in G protein-coupled receptors.
66. Reggio PH. Computational methods in drug design: modeling G protein-coupled receptor monomers, dimers, and oligomers. *Aaps J* 2006;8:E322–E336. [PubMed: 16796383]
67. Muramatsu T, Suwa M. Statistical analysis and prediction of functional residues effective for GPCR-G-protein coupling selectivity. *Protein Eng Des Sel* 2006;19:277–283. [PubMed: 16565146]
68. Sreekumar KR, Huang Y, Pausch MH, Gulukota K. Predicting GPCR-G-protein coupling using hidden Markov models. *Bioinformatics* 2004;20:3490–3499. [PubMed: 15297294]
69. Sgourakis NG, Bagos PG, Papasaikas PK, Hamodrakas SJ. A method for the prediction of GPCRs coupling specificity to G-proteins using refined profile Hidden Markov Models. *BMC Bioinformatics* 2005;6:104. [PubMed: 15847681]
70. Vakser IA. Protein docking for low-resolution structures. *Protein Eng* 1995;8:371–377. [PubMed: 7567922]
71. Soulier JL, Russo O, Giner M, Rivail L, Berthouze M, Ongeri S, et al. Design and synthesis of specific probes for human 5-HT4 receptor dimerization studies. *J Med Chem* 2005;48:6220–6228. [PubMed: 16190749]
72. Dell'Orco D, Seeber M, Fanelli F. Monomeric dark rhodopsin holds the molecular determinants for transducin recognition: insights from computational analysis. *FEBS Lett* 2007;581:944–948. [PubMed: 17300784]
73. Ciarkowski J, Witt M, Slusarz R. A hypothesis for GPCR activation. *J Mol Model* 2005;11:407–415. [PubMed: 15889287]
74. Nikiforovich GV, Taylor CM, Marshall GR. Modeling of the complex between transducin and photoactivated rhodopsin, a prototypical G-protein-coupled receptor. *Biochemistry* 2007;46:4734–4744. [PubMed: 17397191]
75. Filipek S, Krzysko KA, Fotiadis D, Liang Y, Saperstein DA, Engel A, et al. A concept for G protein activation by G protein-coupled receptor dimers: the transducin/rhodopsin interface. *Photochem Photobiol Sci* 2004;3:628–638. [PubMed: 15170495]
76. Crozier PS, Stevens MJ, Forrest LR, Woolf TB. Molecular dynamics simulation of dark-adapted rhodopsin in an explicit membrane bilayer: coupling between local retinal and larger scale conformational change. *J Mol Biol* 2003;333:493–514. [PubMed: 14556740]
77. Faraldo-Gomez JD, Forrest LR, Baaden M, Bond PJ, Domene C, Patargias G, et al. Conformational sampling and dynamics of membrane proteins from 10-nanosecond computer simulations. *Proteins* 2004;57:783–791. [PubMed: 15317024]

78. Huber T, Botelho AV, Beyer K, Brown MF. Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure. *Biophys J* 2004;86:2078–2100. [PubMed: 15041649]
79. Kong Y, Karplus M. The signaling pathway of rhodopsin. *Structure* 2007;15:611–623. [PubMed: 17502106]
80. Saam J, Tajkhorshid E, Hayashi S, Schulten K. Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin. *Biophys J* 2002;83:3097–3112. [PubMed: 12496081]
81. Schlegel B, Sippl W, Holtje HD. Molecular dynamics simulations of bovine rhodopsin: influence of protonation states and different membrane-mimicking environments. *J Mol Model* 2005;12:49–64. [PubMed: 16247601]
82. Cordomi A, Perez JJ. Molecular dynamics simulations of rhodopsin in different one-component lipid bilayers. *J Phys Chem B* 2007;111:7052–7063. [PubMed: 17530884]
83. Grossfield A, Feller SE, Pitman MC. Convergence of molecular dynamics simulations of membrane proteins. *Proteins* 2007;67:31–40. [PubMed: 17243153]
84. Filizola M, Wang SX, Weinstein H. Dynamic models of G-protein coupled receptor dimers: indications of asymmetry in the rhodopsin dimer from molecular dynamics simulations in a POPC bilayer. *J Comput Aided Mol Des* 2006;20:405–416. [PubMed: 17089205]
85. Tirion MM. Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Phys Rev Lett* 1996;77:1905–1908. [PubMed: 10063201]
86. Liang Y, Fotiadis D, Filipek S, Saperstein DA, Palczewski K, Engel A. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *J Biol Chem* 2003;278:21655–21662. [PubMed: 12663652]
87. Guo W, Shi L, Filizola M, Weinstein H, Javitch JA. Crosstalk in G protein-coupled receptors: changes at the transmembrane homodimer interface determine activation. *Proc Natl Acad Sci U S A* 2005;102:17495–17500. [PubMed: 16301531] •The first evidence of conformational rearrangement of a dimer interface upon activation.
88. Periole X, Huber T, Marrink SJ, Sakmar TP. G protein-coupled receptors self-assemble in dynamics simulations of model bilayers. *J Am Chem Soc* 2007;129:10126–10132. [PubMed: 17658882]
89. Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, et al. New developments in the InterPro database. *Nucleic Acids Res* 2007;35:D224–D228. [PubMed: 17202162]
90. Bader GD, Cary MP, Sander C. Pathguide: a pathway resource list. *Nucleic Acids Res* 2006;34:D504–D506. [PubMed: 16381921]
91. Okuno Y, Yang J, Taneishi K, Yabuuchi H, Tsujimoto G. GLIDA: GPCR-ligand database for chemical genomic drug discovery. *Nucleic Acids Res* 2006;34:D673–D677.
92. Fang YC, Sun WH, Wu LC, Huang HD, Juan HF, Horng JT. RINGdb: an integrated database for G protein-coupled receptors and regulators of G protein signaling. *BMC Genomics* 2006;7:317. [PubMed: 17173697]
93. Kazius J, Wurdinger K, van Iterson M, Kok J, Back T, Ijzerman AP. GPCR NaVa database: natural variants in human G protein-coupled receptors. *Hum Mutat*. 2007
94. Horn F, Bettler E, Oliveira L, Campagne F, Cohen FE, Vriend G. GPCRDB information system for G protein-coupled receptors. *Nucleic Acids Res* 2003;31:294–297. [PubMed: 12520006]
95. Beukers MW, Kristiansen I, AP II, Edvardsen I. TinyGRAP database: a bioinformatics tool to mine G-protein-coupled receptor mutant data. *Trends Pharmacol Sci* 1999;20:475–477. [PubMed: 10603488]
96. Papasaïkas PK, Bagos PG, Litou ZI, Promponas VJ, Hamodrakas SJ. PRED-GPCR: GPCR recognition and family classification server. *Nucleic Acids Res* 2004;32:W380–W382. [PubMed: 15215415]
97. Yabuki Y, Muramatsu T, Hirokawa T, Mukai H, Suwa M. 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:W148–W153. [PubMed: 15980445]
98. Bhasin M, Raghava GP. GPCRclass: a web tool for the classification of amine type of G-protein-coupled receptors. *Nucleic Acids Res* 2005;33:W143–W147. [PubMed: 15980444]

Abbreviations

3D, Three-dimensional
CGMD, Coarse Grain Molecular Dynamics
CSI, Cell Signaling Information
dbCFC, Data Base of Cytokine Family Class
EC, Extracellular
eGPCRL, Endogenous GPCR List
ENM, Elastic Network Model
FIRST, Floppy Inclusion and Rigid Substructure Topography
GLIDA, GPCR-Ligand Database
GPCR, G-protein Coupled receptor
GPCRDB, GPCR Database
GPCRIPDB, GPCR interacting proteins Database
GPCR-NaVa, GPCR Natural Variants
GPCR-OKB, GPCR Oligomerization Knowledge Base
GPCRPR, GPCR Pattern Recognition
gpcrPRINTS, Gpcr Protein fingerPRINTS
GPCRsclass, GPCRs Classification
gpcrTASSER, GPCR Threading ASSEmbly Refinement
gpDB, G-Protein Database
GRIFFIN, G-protein Receptor Interacting Feature Finding INstrument
hGPCRdb, Human GPCR Database
HMMs, Hidden Markov Models
IC, Intracellular
IR, InfraRed
IUPHAR-RD, International Union of basic and clinical PHARmacology Receptor Database
LHRmdb, Lutenizing Hormone Receptor Mutation Database
MC, Monte Carlo
MD, Molecular Dynamics
MPKS, Membrane Proteins of Known Structure
NMA, Normal Mode Analysis
ORDB, Olfactory Receptor Database
ORMD, Olfactory Receptor Microarray Data
PDSP, Psychoactive Drug Screening Program
POPC, Palmitoyl-Oleoyl-Phosphatidyl Choline
PRED-GPCR, Prediction of GPCRs
RINGdb, INtgrated Database for Rgs and Gpcrs
RMSD, Root Mean-Square Deviation
SCV, Scaled Collective Variables
SDPC, 1-stearoyl-2-docosaheptaenoyl-phosphatidylcholine
SDPE, 1-stearoyl-2-docosaheptaenoyl-phosphatidylethanolamine
SVM, Support Vector Machine
TASSER, Threading ASSEmbly Refinement
TM, Transmembrane

Figure 1.

Cartoon representation of a GPCR structure embedded in the cell membrane. Lipid bilayer is colored in grey, TM1 in light blue, TM2 in green, TM3 in yellow, TM4 in orange, TM5 in red, TM6 in dark blue, TM7 in purple, and H8 in pink. Intracellular (IC) and extracellular (EC) loops as well as N- and C-termini are represented by thick black lines.

Figure 2.

Classification of dedicated GPCR databases and web server tools according to the information they manage. This information can be divided into seven main categories (depicted as circles): Sequences, Mutations, Ligands, Structures/Models, Signaling Pathways, Interacting Proteins, and Specialized Repositories.

Figure 3. Information flow between experimental data and computational methods in research on GPCR structure and dynamics. Relationships between different modules are indicated by either directional or bi-directional arrows.

Table 1

GPCR Databases and Servers.

Name [*]	Description	URL	Date of Last Update/Release
PRED-GPCR	A bioinformatics classifier of GPCRs at the University of Athens that uses signature profile hidden markov models to screen sequences against well characterized GPCR families.	http://bioinformatics.biol.uoa.gr/PRED-GPCR/	2003
GPCRPR	Classification server that analyzes a query sequence against ~120 fingerprints.	http://www.biochem.ucl.ac.uk/bsm/dbbrowser/GPCR/	
gpcrPRINTS	A bioinformatics resource at the University of Manchester that profiles a query sequence against the PRINTS fingerprint database to determine most similar families of receptor subtypes.	http://umber.sbs.man.ac.uk/dbbrowser/gpcrPRINTS/	Nov 2001
IUPHAR-RD	It includes information on GPCR name synonyms, structure, functional assays, agonists and antagonists, transduction mechanisms, distribution, tissue function, and phenotype.	http://www.iuphar-db.org/iuphar-rd/index.html	
PDSP	It provides published and unpublished Ki, or affinity, values for a large number of drugs and drug candidates at GPCRs, ion channels, transporters and enzymes.	http://pdsp.med.unc.edu/indexR.html	
GLIDA	It integrates information on both GPCRs and their known ligands.	http://pharminfo.pharm.kyoto-u.ac.jp/services/glida/	Oct 2007
GPCRDB	Information system for GPCRs at the CMBI, in The Netherlands. It contains information about sequences, alignments, 3D models, mutations, and ligand binding constants.	http://www.gpcr.org/7tm	Jun 2006
gpcrTASSER	It contains structural models of all 907 GPCRs of the human genome.	http://cssb.biology.gatech.edu/skolnick/files/gpcr/gpcr.html	Feb 2006
MPKS	It contains atomic coordinates for membrane proteins of known structures.	http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html	Nov 2007
GPCRsclass	A specialized server for the prediction of the classification of the amine type GPCRs, based on dipeptide composition and support vector machines.	http://www.imtech.res.in/raghava/gpcrsclass/	2005
hGPCRdb	The human GPCR database at the University Louis Pasteur of Strasbourg. It allows to search for chemogenomics analyses and binding cavity domains.	http://bioinfo-pharma.u-strasbg.fr/gpcrdb/gpcrdb_form.html	2005
dbCFC	Specialized Cytokine Family Receptor database.	http://cytokine.medic.kumamoto-u.ac.jp/CFC/CK/Chemokine.html	Apr 2006
eGPCRI	List of endogenous GPCRs expressed in various cell lines.	http://www.biomedcomp.com/GPCR.html	Nov 2007
ORDB	Olfactory Receptor Database at Yale University. It supports sequencing and analysis of these receptors.	http://senselab.med.yale.edu/senselab/ORDB/	Jan 2007
ORMD	Repository for olfactory receptor microarray data.	http://neurolab.med.yale.edu/ormd	2007
LHRmdb	Specialized Lutenizing hormone receptor mutation database.	http://www2.eur.nl/fgg/endov/lhr.html	
GPCR-NaVa	Repository for sequence variants within the family of human GPCRs.	http://nava.liacs.nl/	Nov 2007
tinyGRAP	GPCR mutant database.	http://www.cmbi.ru.nl/tinygrap/credits/	1999
Arcadia	Repository of mutants of GPCR and transporter proteins, and associated information.	http://icb.med.cornell.edu/services/arcadia/start	
gpDB	A relational database of G-proteins and their interactions with GPCRs. It contains information about known coupling selectivity of GPCRs to G-protein α -subunits.	http://bioinformatics.biol.uoa.gr/gpDB/	Apr 2007

Name [*]	Description	URL	Date of Last Update/Release
CSI	Cell Signaling information for cellular signaling database of pathways and molecules involved in GPCR-mediated and other signal transduction.	http://www.signaling-gateway.org/	Nov 2007
RINGdb	An integrated biological database providing information about GPCRs and Regulators of G-protein Signaling (RGS).	http://ringdb.csie.ncu.edu.tw/ringdb/	May2007
GPCR-OKB	Information system for GPCR oligomers. Currently under construction.	http://www.gpcr-okb.org	
GRIFFIN	It predicts GPCR-G-protein coupling selectivity using a support vector machine and a hidden Markov model.	http://griffin.cbrc.jp	2005
GPCRIPDB	Information system for GPCR interacting proteins. Currently under construction.	http://www.gpcr.org/GPCRIP/	Jun 2005

* CSI, Cell Signaling Information; dbCFC, Data Base of Cytokine Family Class; eGPCRL, Endogenous GPCR List; GLIDA, GPCR-Ligand Database; GPCR, G-protein Coupled receptor; GPCRDB, GPCR Database; GPCRIPDB, GPCR interacting proteins Database; GPCR-NaVa, GPCR Natural Variants; GPCR-OKB, GPCR Oligomerization Knowledge Base; GPCRPR, GPCR Pattern Recognition; gpcrPRINTS, GPCR Protein fingerprints; GPCRclass, GPCRs Classification; gpcrTASSER, GPCR Threading ASSEmblY Refinement; gpDB, G-Protein Database; GRIFFIN, G-protein Receptor Interacting Feature Finding Instrument; hGPCRdb, Human GPCR Database; IUPHAR-RD, International Union of basic and clinical PHARMacology Receptor Database; LHRmdb, Lutenizing Hormone Receptor Mutation Database; MPKS, Membrane Proteins of Known Structure; ORDB, Olfactory Receptor Database; ORMD, Olfactory Receptor Microarray Data; PDSP, Psychoactive Drug Screening Program; PRED-GPCR, Prediction of GPCRs; RINGdb, Integrated Database for RGS and GPCRs.