Computational Modeling Approaches to Structure−Function Analysis of G Protein-Coupled Receptors

Francesca Fanelli*,† and Pier G. De Benedetti‡

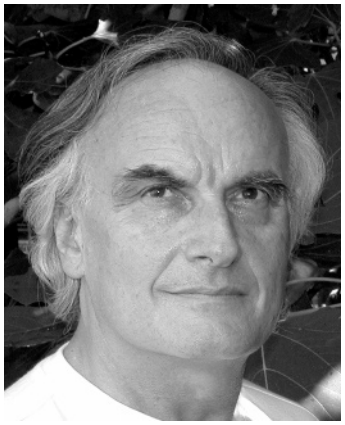
Dulbecco Telethon Institute and Department of Chemistry and Advanced Scientific Computing Laboratory, University of Modena and Reggio Emilia, via Campi 183, 41100 Modena, Italy

Received December 13, 2004

Contents

* To whom correspondence should be addressed. Telephone: +39-059-205-5114. Fax: +39-059-373-543. E-mail: fanelli@unimo.it.
† Dulbecco Telethon Institute and Department of Chemistry.
‡ Department of Chemistry and Advanced Scientific Computing Laboratory.

Francesca Fanelli graduated in Medicinal Chemistry (1989) and received her Ph.D. education in Computational Biophysical Chemistry (1994). She is now an Assistant Scientist of the Dulbecco Telethon Institute. Her research since 1991 has been almost entirely devoted to developing computational protocols and molecular models to unravel the chemical communication mechanisms involving G protein-coupled receptors.



Pier G. De Benedetti is Professor of Biophysical Chemistry at Modena and Reggio Emilia University, where he received his education in Chemistry. His research is mainly devoted to modeling the structure−property and structure−function relationships in complex molecular (bio)-systems. He loves the wind and the sea.

1. Introduction

Any aspect of cell activity is regulated by extracellular signals that are recognized, decoded, and transduced inside the cell via different classes of plasma membrane receptors.[1−3] G protein-coupled receptors (GPCRs) constitute the largest family of signal transduction membrane proteins, which mediate responses of a variety of bioactive molecules, including biogenic

10.1021/cr000095n CCC: $53.50
Published on Web 08/25/2005

amines, amino acids, peptides, lipids, nucleotides, and proteins (reviewed in refs 1 and 4−7) As a result, GPCRs play a crucial role in many essential physiological processes as diverse as neurotransmission, cellular metabolism, secretion, cell growth, immune defense, and differentiation (reviewed in refs 1 and 4−7). Not surprisingly, GPCRs are the most privileged targets for the drugs currently used in clinics and for the wealth of drug candidates that high throughput methods promise to deliver in the immediate future. This reflects not only the broadness of potential applications in all therapeutic fields but also the fact that GPCRs are by far the easiest targets that can be obtained through the synthesis of small organic molecules. In fact, they are naturally built to recognize a single structure among the widest variety of extracellular chemicals, and unlike many enzymes, they are exposed and reachable on the cell surface (reviewed in refs 1 and 4−8).

Although varying considerably in molecular sizes, any GPCR polypeptide sequence contains seven hydrophobic α-helices that span the lipid bilayer and dictate the typical architecture of the macromolecule: seven transmembrane (TM) domains bundled up to form a polar internal tunnel and expose the N-terminus and three interconnecting loops, to the exterior, and the C-terminus with a matching number of loops, to the interior of the cell. This structural information, initially based on low resolution electron diffraction studies and on predictions from bioinformatics and molecular modeling,[9−12] was recently confirmed by the resolution of the crystal structure of rhodopsin.[13] In many cases, ligands bind within the three-dimensional (3D) cavity of the TM bundle and form multiple connections with amino acid residues located in different domains. There is plenty of variation, however. In some cases the binding site is entirely located on a large and highly structured amino-terminal domain, as for glutamate[14] or glycoprotein hormone receptors,[15] or it involves both extracellular loops and residues in the transmembrane core, as often observed for neuropeptide receptors (reviewed in refs 1 and 4−7).

According to sequence analyses, GPCRs have been clustered in a number of families/classes.[4,16−19] The different classification systems include the A to F system,[19] the 1 to 5 system,[4] and the GRAFS system.[17] Thus, A (named 1 or rhodopsin in the 1 to 5 or the GRAFS system, respectively) is the rhodopsin-like class/family; B (or 2 or secretin) is the secretin-like class/family; C (3 or glutamate) is the metabotropic glutamate and pheromone class/family; D (or 4) is the fungal pheromone class/family;[20] E is the cAMP receptor class/family; and F (or 5 or frizzled) is the frizzled/smoothened family.[4,17,19] Family A receptors, the topic of this review article, are by far the largest and the most studied. The overall homology among all family A receptors is low and restricted to a small number of highly conserved key residues distributed in each of the seven helices.[4,17,19]

All GPCRs also share a common molecular strategy of signal transduction, to which they owe their name. Agonist binding promotes allosteric interactions between the receptor and one or more members of the family of heterotrimeric guanidine triphosphate (GTP) exchanging proteins or G proteins. These are specialized GTPases that act as signal transducers and broadcast the signal to a host of intracellular effectors, either enzymes, such as adenilyl cyclase, cGMP phosphodiesterase, and phospholipases, or ion channels, such as potassium and voltage gated calcium channels.[21,22] Direct observation that GPCRs, including rhodopsin and the δ opioid receptor, can couple to G protein even in their inactive states comes from plasmon-waveguide resonance (PWR) spectroscopy, thus providing support to the hypothesis that receptors and G proteins are "constitutively coupled" or precoupled.[23,24]

G proteins account for the majority of signals that GPCRs can evoke into the cell. However, a multitude of relatively recent data shows that the signaling pattern of GPCRs is more dynamical than what was originally believed. On many occasions, that reflects the enhanced knowledge on the network of intracellular signaling pathways, in which G proteins are implicated. Thus, each GPCR can generate a secondary "wave" of signal transduction that essentially depends on the type of G protein α-subunits that are activated and, typically, involves signaling pathways, which are primary targets of growth-factor and cytokine receptors.[25] In others cases, however, there are clear indications that GPCRs can generate signals bypassing G protein intervention.[26] Although the exact mechanisms are not entirely elucidated, it is becoming increasingly apparent that direct receptor−receptor interactions, leading to a dimeric or multimeric quaternary structure, play a role in G protein-independent signaling.[26−28]

Cells have counter-regulatory mechanisms that attenuate signaling by activated GPCRs (reviewed in refs 1 and 29). These mechanisms include acute desensitization involving GPCR-specific protein kinases and arrestins (homologous desensitization) and second messenger-activated protein kinases, such as protein kinases A and C (heterologous desensitization) (reviewed in refs 1, 29, and 30).

The classical idea that GPCRs function as monomeric entities has been unsettled by the emerging concept of GPCR dimerization (reviewed in refs 31−41). Recent findings have indicated not only that many GPCRs exist as homodimers and heterodimers but also that their oligomeric assembly could have important functional roles. Several studies have shown that dimerization occurs early after biosynthesis, suggesting that it has a primary role in receptor maturation (reviewed in refs 38 and 41). For many proteins, oligomeric assembly has an important function in endoplasmic reticulum (ER) quality control because it masks specific retention signals or hydrophobic patches that would otherwise retain the proteins in ER.[42] G protein coupling, downstream signaling, and regulatory processes, such as internalization, have also been shown to be influenced by the dimeric nature of the receptors (reviewed in refs 38 and 40). The question whether dimerization influences ligand-induced activation/regulation of GPCRs still remains to be answered. In fact, some studies suggest that ligand binding can regulate the

dimer by either promoting or inhibiting its formation, whereas many others conclude that homodimerization and heterodimerization are constitutive processes that are not modulated by ligand binding (reviewed in ref 38). In any case, the structural data available strongly suggest that at least some GPCRs can form dimers in the absence of ligand stimulation. A clear evidence, in this respect, comes from atomic force microscopy (AFM) measurements, which showed that rhodopsin and opsin form a constitutive dimer in dark-adaptive retinal membrane.[39,43] Recent studies have suggested that heterodimerization could affect agonist-promoted GPCR endocytosis, a well-characterized process classically involved in signal attenuation (reviewed in ref 38). In the case of adenosine/dopamine and somatostatin/opioid receptors, the cointernalization was also associated with a cross-desensitization of the signaling activities (reviewed in ref 38).

Thus, regulated protein−protein interactions are key features of many aspects of GPCR function and there is now increasing evidence for GPCRs acting as part of multicomponent units comprising a variety of signaling and scaffolding molecules, organized in supramolecular signaling assemblies (signalsomes or transducisomes).[1,2]

Usually with native GPCRs, activation is initiated by agonist binding. However, GPCRs can achieve the active states independently of agonists; that is, they can become constitutively active.[44] It is now evident that a number of native GPCRs exhibit constitutive signaling activity, but the role of agonist-independent activity in normal physiology is not known (reviewed in refs 45 and 46). Constitutively active GPCRs also are invaluable tools to discover the signal transduction pathways of the hundreds of orphan GPCRs, which are potential targets of novel drugs (reviewed in ref 47). On the other hand, a number of constitutively active GPCR mutants have been found, which are involved in the pathogenesis of human disease (reviewed in refs 29, 48, and 49). Given the large number of GPCRs encoded within the human genome, additional examples of this pathogen mechanism are likely to be uncovered. Furthermore, the spectrum of diseases caused by constitutively active GPCRs is expanding to include diseases caused by infectious agents (reviewed in ref 29). A more complete elucidation of the roles of constitutively active GPCRs in human disease and an understanding, at the molecular level, of how these pathogenic GPCRs could be inactivated may allow rational development of specific compounds as therapeutic agents. Diseases are caused not only by constitutively active mutations ("gain-of-function" mutations) but also by mutations of an endogenous GPCR, which cause the receptor to lose the ability to bind agonist or to signal ("loss of function" mutations). A number of pathologies have been, hence, found to be related to mutations of GPCRs.[29,48,50] Most of these pathologies are related to obvious clinical manifestations, such as blindness, X-linked diabetes insipidus, and hypo- or hyperthyroidism, precocious puberty, obesity, cancer, etc. Some undiscovered mutations, providing nonobvious phenotypes, are also likely to be responsible for pathologies such as psychiatric or neurological disorders. Some inherited mutations may never be detected because they are incompatible with life.

Despite the enormous biomedical relevance of GPCRs, high resolution structural information on their active and inactive states is still lacking. So far, the only GPCR whose structure has been resolved with atomic detail is rhodopsin in its dark state.[13,51−54] This frustrating situation leaves unanswered too many questions on GPCR function. Some of these are the following: Which is the precise structural basis of ligand specificity for a particular receptor, and how can the basic seven-helical structure be tuned to bind such a large and chemically diverse spectrum of ligands? Which is the precise molecular mechanism of ligand-dependent and ligand-independent GPCR activation? Which are the architectures of the different supramolecular assemblies of GPCRs? Which is the role of receptor dimerization/oligomerization in GPCR function? How is ligand- and mutation-induced chemical information transferred within a molecular network of GPCRs? Which are the G protein contact sites on the receptor and which is the stoichiometry of the receptor−G protein complexes? Which is the atomic pathway of signal transduction from the ligand-binding side on the receptor to the nucleotide-binding side of the G protein? Which are the structural bases of G protein-independent signaling displayed by many GPCRs? How do GPCRs dimers/oligomers and intracellular proteins organize themselves into a functional unit?

Experimental data obtained on chemically complex biosystems, like GPCRs, often contain more information than we need for a specific answer to a well conjectured hypothesis. Sound chemical/molecular models and data analysis techniques can help with decoding and describing intriguing experimental data. The advent of information and computer technology, hence, allowed for an intriguing integration between computational modeling and in vitro experiments.

The present work is aimed at critically reviewing the results of computational experiments, which have attempted, over the last 16 years, to gain insights into different aspects of family A GPCR function.

2. Structural Features of Rhodopsin: The Founder of Family A GPCRs

2.1. Structural Models of the Dark State

Despite a growing appreciation of functional analogies between visual and hormonal signaling systems in the early 1980s, the discovery of the close structural relationship between rhodopsin and the β_2-adrenergic receptor (β_2-AR), and of the existence of a larger "superfamily" of such receptors, came as a total surprise.[6] Rhodopsin, thus, became the founder of family A GPCRs, which includes also the β_2-AR, and the best source of high resolution information on the homologous receptors. Therefore, we thought it right to summarize herein the insights gained so far into the structural features of active and inactive states of the photoreceptor. These pieces of information have been, indeed, widely used in computational modeling of the homologous GPCRs

Rhodopsin is involved in the molecular transformation of light energy into a neuronal signal transmitted to the secondary neurons of the retina and ultimately to the brain.[13] In the case of rhodopsin, the signal is made up of two components: the bound chromophore, which undergoes cis−trans photoisomerization, and the trigger of such photoisomerization, a photon.[13]

The first highly resolved structure of rhodopsin, deposited in the Protein Data Bank (PDB: http://www.rcsb.org/pdb/) under the identifier 1F88, showed all major structural features as predicted from years of biochemical, biophysical, and bioinformatics studies on wild type and mutated proteins.[13] Bovine rhodopsin contains 348 amino acids and folds into seven TM helices, varying in length from 19 to 34 residues, and one cytoplasmic helix, H8 (the letter "H" stands for helix). The seven TM domains contain a mix of canonical α- and 3.10-helices, and they possess a large number of kinks, twists, and bends (Figures1 and 2). Further refinements have not reduced the number of these conformational anomalies.[51,52,54] Such anomalies are also present in the rhodopsin structure recently released by Li and coworkers (PDB code 1GZM).[53]

In the rhodopsin structure, H1 is 44 Å long, it is tilted from the membrane normal, and it contains a bend, mostly due to the presence of P53(1.48). (The numbering in parentheses follows the arbitrary scheme by Ballesteros and Weinstein;[55] according to this scheme, every amino acid identifier starts with the helix number, followed by the position relative to a reference residue among the most conserved amino acid in that helix. That reference residue is arbitrarily assigned the number 50). H2 is tilted from the membrane normal about the same as H1, and it deviates from an ideal helix around G89(2.56) and G90(2.57). H3 is the longest (48 Å), the most tilted, and the most buried helix, bent at G120(3.35) and G121(3.36) and at S127(3.42).[13] The cytosolic extension of H3 is particularly important, because it contains the highly conserved E/DRY motif. In the rhodopsin structure, the arginine of this conserved motif, R135(3.50), is engaged in a double salt bridge with the adjacent glutamate, E134(3.49), and E247-(6.30).[13] Both the E3.49−R3.50 and R3.50−E6.30 interactions are suggested to contribute to keep the photoreceptor and the homologous GPCRs in their inactive states, on the basis of the results of in vitro and computational studies.[45,56−74] H4 is the shortest helix; it is almost perpendicular to the membrane and deviates from an ideal helix at its extracellular end, due to P170(4.59) and P171(4.60) (Figure 2). H5 is 35 Å long, it is tilted from the membrane normal, and it has two internal bends at residues F203(5.38) and H211(5.46) (Figure 2). H6 is the second longest helix. Its cytosolic half is almost perpendicular to the membrane plane, whereas its extracellular half is bent, because of P267(6.50), one of the most conserved residues in the rhodopsin family of GPCRs.[10,11,75] H7 shows a considerable distortion and elongation in the region around the retinal attachment site K296(7.43) and contains two prolines, P291(7.38) and P303(7.50). The latter belongs to the highly conserved NPxxY motif (Figures 1 and 2). The

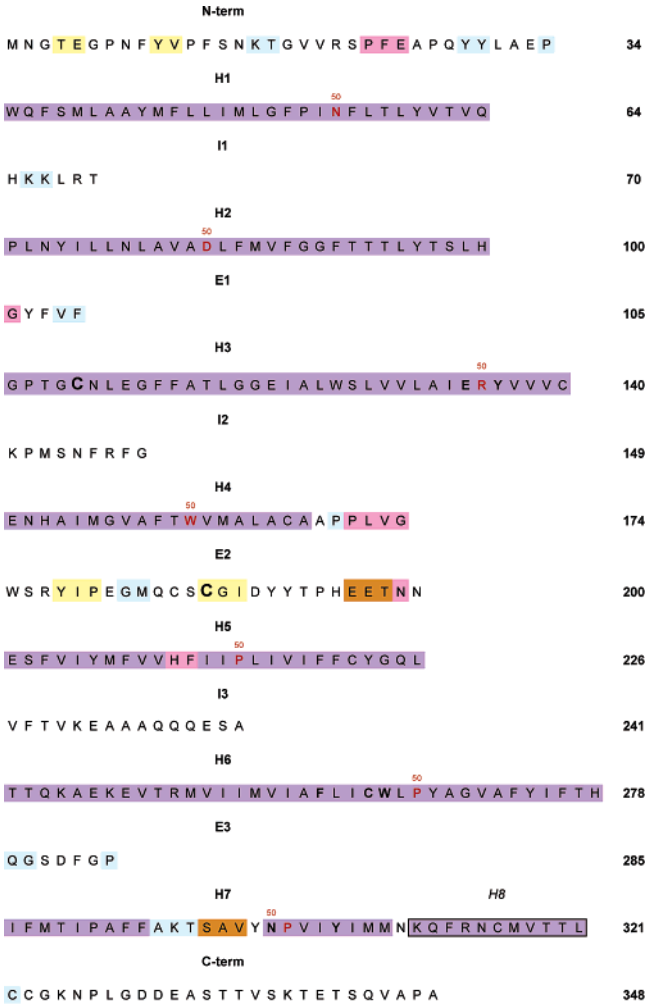


Figure 1. Amino acid sequence of bovine rhodopsin. Color highlights indicate the secondary structure computed on the latest structure (PDB code 1U19).[54] In detail, violet means canonical α-helices, and yellow stands for strand, whereas cyan, orange, and pink indicate, respectively, type 3-, 4-, and 5-turns. The most conserved amino acids in each helix are colored in red. These amino acids are at position 50 according to the arbitrary numbering scheme by Ballesteros and Weinstein.[55] Bold characters indicate the conserved members of the E/DRY, FxxCWxP, and NPxxY motifs in H3, H6, and H7, respectively. Boxed characters indicate the cytoplasmic H8.

tyrosine of this motif is involved in aromatic interactions with F313, a conserved residue in H8 (Figure 2). Recent experimental evidences suggest that the NPxxY(x)$_{5,6}$F and E/DRY motifs provide, in concert, a dual control of the activating structural changes in the photoreceptor.[76] In addition to these TM helices, another short helix in the cytoplasmic surface, termed H8, is located at the cytosolic end of H7 and it is almost perpendicular to the membrane normal (Figure 2).

The extracellular and intracellular regions of rhodopsin each consist of three interhelical loops (given the prefix E or I, for extracellular and intracellular, respectively) as well as of two tails (i.e. N-term and C-term, respectively).

A clear contrast exists concerning the packing of the intra- and extracellular domains; whereas the four extracellular domains associate significantly with each other, only a few interactions are observed

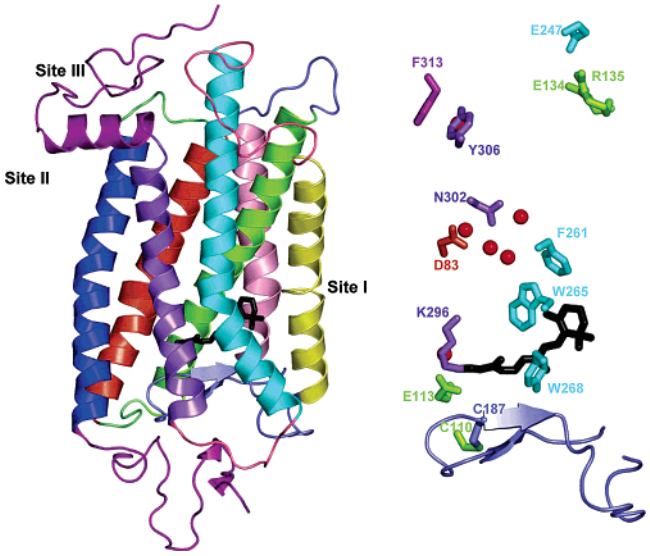


Figure 2. Side view, in a direction parallel to the membrane surface, of the rhodopsin structure encoded as 1U19.[54] On the left side, a cartoon representation of the whole structure is shown, including the 11-*cis*-retinal, represented by black sticks. On the right side, a stick representation of selected highly conserved amino acids in the seven-helix bundle and of C110(3.25) and of C187, which are engaged in a disulfide bridge, is shown. Red spheres represent the oxygen atoms of some of the water molecules resolved into the channel and close to the highly conserved amino acids in H2 and H7. A cartoon representation of E2 is also shown in the right panel. The intracellular side is at the top. H1, H2, H3, H4, H5, H6, and H7 are respectively colored in blue, orange, yellow-green, pink, yellow, cyan, and violet; the N- and C-termini, including H8, are in purple, I1 and E1 are in lime, I2 and E2 are in slate, and I3 and E3 are in salmon. Sites I, II, and III represent respectively the retinylidene binding pocket, the putative retinoid entrance site, and the putative retinoid exit site.[115] Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

among the cytoplasmic domains (Figure 2). The cytoplasmic loops are poorly determined in the structures. This is the region of the protein with the highest B-factors, and these loops are probably mobile in solution.[13] Residues missing from 1F88 are the stretches 236−240 in I3 and 331−333 in the C-term.

The extracellular domains of rhodopsin consist of a folded N-term, from residue 1 to 33, which overlays the rest of the loops (Figure 2). A β-hairpin in this domain interacts with a β-hairpin in E2, forming a β-sheet. E2 contains C187, which is involved in a disulfide bridge with C110(3.25) (Figure 2). Very recently, simulations of thermal unfolding of rhodopsin, by the recently developed floppy inclusion and rigid substructure topography (FIRST) method,[77] combined with fast mode analysis of rhodopsin, using the Gaussian network model,[78] identified the C110(3.25)−C187 disulfide bond and the retinal ligand binding pocket as part of a core region, which is assumed to be important for the formation and stability of folded rhodopsin.[79] This region was the most rigid one in rhodopsin. Experiments confirmed that 90% of the amino acids predicted by the FIRST method to be part of the core cause misfolding upon mutation.[79] The fundamental role of E2 in stabilizing the inactive state of the receptor has also been

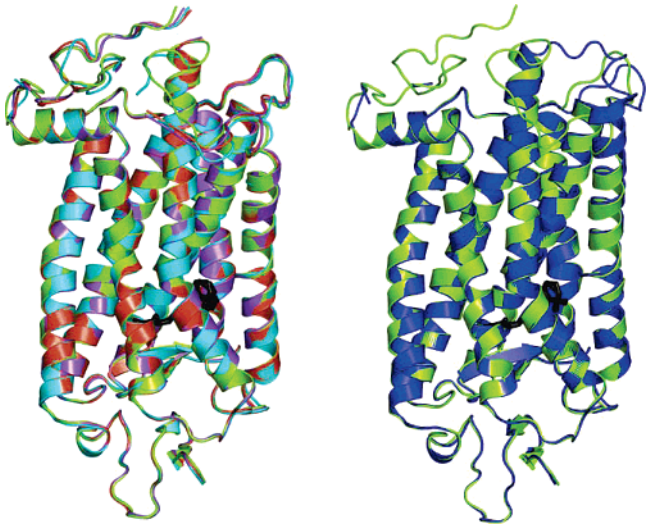


Figure 3. (left) Cartoons of the superimposed structures of 1U19 (yellow-green),[54] 1L9H (violet),[52] 1HZX (orange),[51] and 1F88 (cyan).[13] (right) Cartoons of the superimposed structures of 1U19 (yellow-green)[54] and 1GZM (blue).[53] The helix bundles are seen in a direction parallel to the membrane surface, with the intracellular side being at the top. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

demonstrated for the complement factor 5a receptor (C5aR),[80] and it might serve a similar role in other GPCRs.[81]

The tight packing between the N-term and E2 and their secondary structure elements, as shown by the crystal structure, had not been predicted by bioinformatics approaches. Other unpredicted features were the irregularities in H1, H2, H3, and H5 and the short helix named H8 (Figure 2). The susceptibility of H7 to deviation from the canonical conformation had been instead already predicted prior to the release of the first crystal structure.[82−84]

Crystallographic refinement of the two molecules in the asymmetric unit generated the model of rhodopsin at 2.8 Å resolution, deposited in the PDB under the identifier 1HZX.[51] Differences between 1F88 and the refined 1HZX structures are located mainly in the cytosolic domains (Figure 3). The root mean square deviation (RMSD) between the Cα-atoms (i.e. Cα-RMSD) of 1F88 and 1HZX is equal to 1.05 Å. Major structural differences between 1F88 and 1HZX concern I3, which was rebuilt in 1HZX, lacking residues 236−240. In contrast, the extracellular loops and the chromophore were changed minimally. Furthermore, in 1HZX, additional amino acids in the C-tail (i.e. residues 328−330) are missing as compared to 1F88, whereas the amino acid stretch 334−348, which in 1F88 was filled with Ala residues, carries all the side chains in 1HZX.[51] Improved resolution was obtained concerning the successive structural models, 1L9H (2.6 Å resolution)[52] and 1U19 (2.2 Å).[54] 1L9H is changed to a much lesser extent when compared to 1HZX (i.e. the Cα-RMSD is 0.36 Å; Figure 3). No additional amino acids were added. The latest and most resolved rhodopsin structure, 1U19 , completes the description of the protein backbone and is in general agreement with earlier diffraction studies.[13,51,52,54] The main differences be-

tween the 2.6 and 2.2 Å structural models essentially concern the completion of I3 and the C-term in the latter structure in addition to the consequent conformational differences in these cytosolic domains. The structures of the 11-*cis*-retinal chromophore and its binding site have been defined with greater precision than ever before in the 2.2 Å resolution structure, demonstrating a significant pretwist of the C11−C12 double bond, which is suggested to be critical for the function of rhodopsin.[54] The position of water molecules in 1L9H had already been defined with high precision.[52] In this respect, some of the highly conserved residues within family A GPCRs, including D83(2.50), N302(7.49), and Y306(7.53), are found to form binding sites for these water molecules (Figure 2). The latest structure by Okada and coworkers confirms the water molecule topography found in 1L9H but also adds new molecules, which leave no cavity in the protein.[54]

Li et al. have resolved a rhodopsin structure at 2.65 Å (PDB code 1GZM),[53] by using untwinned native crystals in the space group *P*31, by molecular replacement from the 2.8 Å model solved in space group *P*41 (i.e. 1F88).[13] Like 1U19, 1GZM also resolves all the interhelical loops. The most significant main chain differences (i.e. Cα-RMSD above 1 Å) between 1GZM and 1U19 reside in I2, I3, and the C-terminal tail following H8. In detail, I2 has a similar "L-shape" in both structures but with a different orientation relative to the helix bundle. The difference may be described as a hinge movement about the junctions with the cytoplasmic ends of H3 and H4 (Figure 3). I3 and the cytoplasmic ends of H5 and H6 show the most striking difference between the two structures. In 1U19, H5 terminates at residue L226(5.61), following which the polypeptide chain re-enters the bilayer and is nonstructured between residues V227 and A241 (Figure 3, right). In contrast, in 1GZM, H5 extends to residue V230(5.65), and then the Cα-trace continues in a helixlike spiral path away from the membrane to Q236, where it changes direction to run parallel to the membrane and joins the cytoplasmic end of H6 at A241(6.24) (Figure 3, right). Differently from the case of 1U19, in 1GZM, the C-tail is disordered, except for a dipeptide suggested by the density in contact with I1.[53,54] Other differences between 1U19 and 1GZM concern the conformation of the retinal ring, which, in the latter, resembles the conformation held in the first structural model, 1F88.[13,53,54]

2.2. Supramolecular Organization of Rhodopsin

Very recently, the presumed higher order oligomeric state in native membranes has been demonstrated for rhodopsin.[43] AFM experiments revealed the native arrangement of rhodopsin, which forms paracrystalline arrays of dimers in mouse disk membranes. Indeed, at higher magnification, almost all rhodopsin molecules are organized in rows of dimers, with a few monomers and some single rhodopsin pairs that have broken away from the rows.[43] These experimental evidences were challenged by Chabre and co-workers.[85] The criticisms of Chabre and coworkers were promptly addressed by the authors through additional experiments.[86−88] Whether oligomerization of rhodopsin is a constitutive feature or an artifact is still in debate.[39,89]

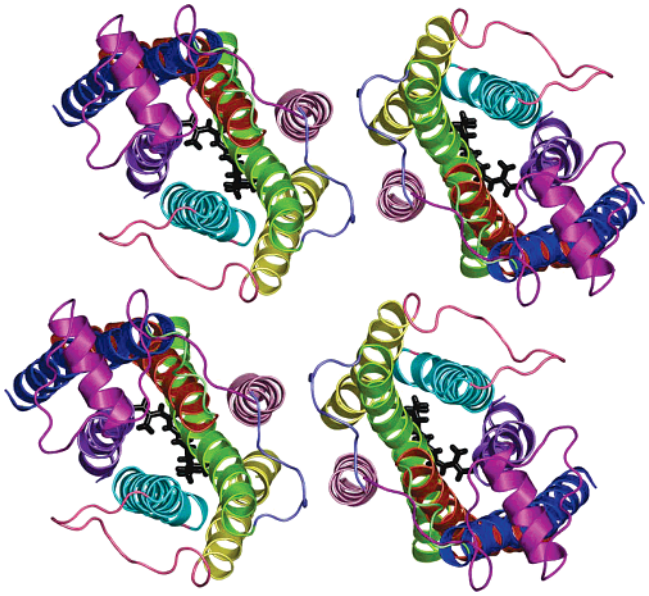


Figure 4. Cartoon representation of the semiempirical model of the rhodopsin tetramer 1N3M,[90] seen from the intracellular side in a direction perpendicular to the membrane surface. The extracellular loops are not shown. Color coding follows the same criteria as in Figure 2. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

On the basis of the AFM experiments, a semiempirical model of a higher order rhodopsin structure in native membranes was built (PDB code 1N3M, Figure 4).[90] Such an oligomeric model is made of repeats of the same monomeric unit. The latter was obtained by completing the structural model 1HZX,[51] by means of the MODELLER program.[90,91] According to this supramolecular model, two monomers of rhodopsin interact with each other at the extracellular (intradiscal) side (i.e. at E2, from both the monomers), at the cytoplasmic side (i.e., at I2 from both monomers), and also within H4 and H5 (Figure 4). On the basis of this model, the authors suggested that only the extracellular interactions involving E2 can transmit information about ligand binding from one receptor monomer to another, because of the close proximity of E2 to the ligand binding pocket. Contacts between dimers involve I3 and both I1 and the C-tail (Figure 4). The latest structural model of rhodopsin, i.e., 1U19,[92] which differs from the monomer in 1N3M essentially for the conformation of the 228−244 amino acid stretch constituting I3, and for a portion of the C-tail (i.e. 323−335), can effectively substitute for the original monomers in 1N3M.[93] Ciarkowski et al., very recently, proposed an alternative arrangement of the 1N3M monomers similarly compatible with the geometrical constraints from the AFM measurements[43] and judged by the authors more consistent with the mechanistic hypothesis of rhodopsin activation than the original organization in the 1N3M oligomer.[94]

Evidences for ordered alignment of squid rhodopsin in the membrane have also been derived from structure determinations by cryo-electron microscopy.[95]

Consistent with previous findings from the projection map of squid rhodopsin, it appeared clear that the crystal lattice is formed by rows of rhodopsins with less protein−protein contacts between the rows. A different oligomerization model of rhodopsin arises from such studies, compared to the semiempirical model 1N3M.[90,95] In fact, docking the Cα-atoms of the structural model 1F88 into the 3D map of squid rhodopsin suggested that the contacts between adjacent rhodopsins along the lattice rows are made by H1, H8, H5, and I2. In detail, on the cytoplasmic side, positively charged amino acids in I2 or I3, of one monomer, would interact with negatively charged amino acids in the C-tail, of the other monomer.[95] Cross sections through the center of the membrane show contacts between H5 and H1, whereas interdimer contacts appear to be mediated by H4.[95]

Assuming that the dimeric/oligomeric state of rhodopsin is a constitutive feature of living cells, the functional role of oligomerization of the photoreceptor is still obscure. Functional characterization of rhodopsin monomers and dimers in detergents has recently demonstrated that monomeric rhodopsin can activate transducin, though the oligomeric form is more active.[87] This evidence seems to be in line with the idea that the receptor monomer holds the structural determinants for G protein activation and, in this respect, is the functional unit.[89]

Rhodopsin oligomerization may be essential for the ontogeny and/or desensitization of the photoreceptor and, hence, in the control of light signaling. Addressing these aspects may also have important implications in unraveling the molecular determinants of retinal degenerative diseases.

2.3. Insight into the Active States of Rhodopsin from Biophysical Experiments

The high resolution structures of rhodopsin refer to the inactive state of the photoreceptor. So far, information on the active states has been essentially inferred from biophysical and biochemical experiments (reviewed in ref 96).

Photochemical experiments allowed definition of the reaction coordinates of rhodopsin activation.[97,98] Absorption of a photon provides rhodopsin with the energy to form the active state. Three phases of the activation process can be distinguished: (1) light-induced cis−trans isomerization of the retinal; (2) thermal relaxation of the retinal−protein complex; and (3) the late equilibria that are affected by the interaction of rhodopsin with the G protein (reviewed in refs 97 and 98).

Photobleaching of rhodopsin involves different intermediates, identified by low temperature and time-resolved spectroscopic experiments (Figure 5).[98−102] Fourier transform infrared spectroscopy (FTIR) experiments suggest that D83(2.50) and E122(3.37) are both protonated in the dark rhodopsin state.[103] Following photon absorption and electronic excitation, fast isomerization of the chromophore leads to formation of bathorhodopsin (BATHO, 529 nm) (Figure 5) through photorhodopsin (PHOTO, 570 nm). This event, which utilizes two-thirds of the energy taken up by light absorption, would require

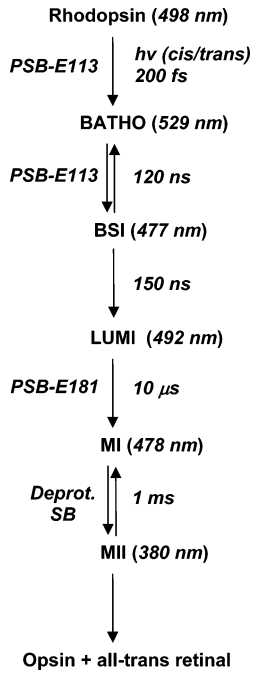


Figure 5. Rhodopsin photobleaching.

200 fs to occur. BATHO is in equilibrium with the blue-shifted intermediate (BSI, 477 nm), which decays to lumirhodopsin (LUMI, 492 nm) in 150 ns. The BATHO to BSI transition only involves conformational changes of retinal, whereas the BATHO to LUMI transition is accompanied by a large motion of the β-ionone ring away from W265(6.48), which is in close proximity to the β-ionone ring in the dark and BATHO states. LUMI undergoes a transition to metarhodopsin-I (META-I, MI, 478 nm) in 10 μs. A shift of the protonated Schiff base (PSB) from E113(3.28) to E181 (in E2) accompanies the transition from the dark to the MI state. It has been proposed that the switch is accomplished by transferring a proton from E113(3.28) to E181, through a structurally evolving H-bond network.[100] In the LUMI state, the PSB group has shifted away from E113(3.28) and the identity of a formal counterion has been lost because the negative charge becomes delocalized along the H-bond chain, which involves two water molecules and S186 in E2. Raman studies support the idea that LUMI is the transition state in the counterion-switching process, revealing that the H-bond of the Schiff base dramatically weakens in the BSI to LUMI transition before a more normal H-bond is formed with some residues in the LUMI to MI transition. Finally, in the MI state, the Schiff base group has moved toward E181. Because the distance from the PSB to E181 in rhodopsin is only ∼7 Å compared to ∼3 Å for E113(3.28), only a modest conformational change that alters the spatial relationship between E2 and H3 is predicted to be required for the formation of the new salt bridge between E181 and the PSB in MI.[100] Very recently, a density map of a photostationary state highly enriched in MI, to a resolution of 5.5 Å in the membrane plane, has been determined by electron crystallography.[104] The map shows density for H8, the cytoplasmic loops, the extracellular domains, all tryptophan residues, an ordered cholesterol molecule,

and the β-ionone ring. Comparison of this map with the X-ray structures of the ground state reveals that MI formation does not involve large rigid-body helix movements, but there is a rearrangement close to the bend of H6, at the level of the retinal chromophore. This evidence suggests that there is no gradual buildup of the large conformational change known to accompany formation of MII, the signaling state capable of activating the G protein.[104] These results provide evidence for the first time that rhodopsin remains in a conformation similar to that of the ground state until late in the photobleaching process, the gross helix movements, and the conformational changes occurring in the MI to MII transition.[104] The transition from MI to MII would, indeed, occur in 1ms. Formation of MII is also linked to proton uptake from the cytoplasm. The reaction depends on and probably involves the protonation of E134(3.49), a member of the highly conserved E/DRY motif.[98,105,106] In the dark state, this residue forms a salt bridge with the adjacent R135(3.50). Protonation of E134-(3.49), which requires the presence of transducin,[106] would destabilize the charge-reinforced H-bond between E134(3.49) and R135(3.50), releasing an important constraint of the inactive state. This hypothesis is supported by the findings that a mutation eliminating the negative charge at E134(3.49) induces constitutive activity of the opsin.[56] Electron paramagnetic resonance (EPR) analyses of the cytoplasmic surface of the E134(3.49)Q constitutively active mutant showed a local conformational change around H3 and H7, whereas only a small change was seen in H6, which is suggested to undergo a dominant motion upon photoactivation.[107] These data suggest that the H6 motions induced by photoactivation are essentially triggered by the changes in the interaction pattern between the retinal and the opsin, following the cis−trans isomerization of the chromophore. Mutation of E134(3.49) abolishes light-induced proton uptake but leaves the proton transfer to E113-(3.28) unchanged,[108] suggesting that the two proton translocations can be decoupled, corresponding to the MIIa and MIIb states.[98,105] According to this model, the transition of MI to MIIa is accompanied by translocation of the Schiff base proton to E113(3.28).[98] In this stage, the *all-trans*-retinylidene group has the characteristics of a ligand agonist, in that it facilitates the MI−MIIa transition by elevation of the free energy (ΔG) of MI through a scaffold function for proton translocation.[109] Proton uptake from the cytoplasm leads to formation of MIIb, with a pH-dependent ΔG.[98] The positive enthalpy of MII formation indicates that molecular interactions built up in MI are lost upon transition to MII. To drive the conversion, the entropy and, thus, the overall disorder in the protein must increase.[98] This observation would be consistent with the idea that formation of the active state is merely a release of constraints in the helix bundle, thus exposing cytoplasmic binding sites.

Initial deactivation of MII begins with the interaction of active rhodopsin with its receptor kinase, phosphorylation of the receptor, and a tight binding of arrestin to the still activated phosphorylated form of the receptor.[110,111] Full deactivation occurs when rhodopsin is regenerated. This requires the hydrolysis of the *all-trans*-retinylidene linkage and release of *all-trans*-retinal from the active site.[112,113] Critical steps include the nucleophilic the attack of water on the retinylidene bond within the hydrophobic binding site of rhodopsin and the diffusion of the hydrolyzed chromophore out of the binding pocket. Formation of opsin accompanies a significant increase in intrinsic tryptophan fluorescence after release of *all-trans*-retinal from the active site.[112,113] The retinal remains associated with opsin membranes and is converted by endogenous NADPH-dependent retinol dehydrogenase (RDH, reviewed in ref 97) to *all-trans*-retinol without further change in the intrinsic protein fluorescence. The RDH activity is suggested to exert an influence on the stability of the complex between arrestin and phosphorylated Meta II. In addition, during the Meta II decay, a storage form of rhodopsin, metarhodopsin III (Meta III, MIII), is generated. The formation of MIII can be triggered by blue light absorption in MII, passing through the anti−syn isomerization form or "reverted-Meta" intermediate MIII, and the subsequent reprotonation of the Schiff base.[102,114] Recently, it has been inferred that, in addition to the retinylidene pocket (site I), there are two other retinoid binding sites within opsin (Figure 2). Site II, involved in the uptake signal, is an entrance site, while site III is the exit site that is occupied when retinal remains bound after its release from site I.[115] Support for a retinal-channeling mechanism comes from the crystal structure of rhodopsin, which unveiled two putative hydrophobic binding sites in the cytosolic domains, i.e., close to the C-terminal end of H8 (site II) and on the solvent-exposed surface of the C-tail (site III) (Figure 2). The storage form is suggested to be characterized by the photolyzed *all-trans*-retinal bound in the exit site (site III). In this state, the retinal does not quench intrinsic tryptophan fluorescence. Opsin eventually returns to the ground state via a transiently formed opsin−11-*cis*-retinal complex, which contains both retinal isomers bound to site II and site III. RDH has access to its retinal substrate while bound to the site III of opsin.[115]

The activity of ligand-free opsin is equal to 10^{-6} of the activity of the *all-trans*-retinal-bound active MII state.[98,116] However, the 11-*cis*-retinal-bound rhodopsin ground state exhibits an even lower level of activity against transducin (Gt), suggesting that the 11-*cis*-retinal acts as an inverse agonist and imposes further structural constraints. Besides interactions with the chromophore, numerous intramolecular interactions are found in the crystal structure, which stabilize the ground state, and most of these are mediated by highly conserved residues in GPCRs.[75]

Elegant cysteine cross-linking, site-directed spin labeling, scanning accessibility, and NMR determinations on rhodopsin provided invaluable information on the structural rearrangements associated with MII formation. Indeed, the results of the experiments on the dark rhodopsin state were consistent with the rhodopsin structure, providing also a model for rhodopsin activation (reviewed also in ref 96).

Cross-linking experiments aimed at inferring the structural changes in the retinal binding site, as induced by photoactivation, revealed a cross-link between retinal and A169(4.58) in the light-activated rhodopsin.[96,117] Since A169(4.58) in the crystal structure of rhodopsin faces outward from the helix bundle, such cross-linking is inconsistent with the retinal binding mode in the inactive state of rhodopsin. In fact, for the β-ionone ring of retinal to be linked to A169(4.58), the extracellular end of H3 should move away from the core of the helix bundle, and H4 and H7 should tilt or rotate to open a path between K296(7.43), which is covalently bound to retinal, and A169(4.58).[96,117] Insights into the configurational differences between dark and MII states in the retinal binding site come also from solid-state NMR determinations.[118−120] In detail, ^{13}C-labels on the retinal chromophore and specific ^{13}C-labels on tyrosine, glycine, serine, and threonine residues in the retinal binding site indicate that the essential aspects of the isomerization trajectory are a large rotation of the C20 methyl group toward E2, and a 4- to 5-Å translation of the retinal chromophore toward H5.[118] The observed shifts of the chromophore are predicted to trigger the motions of W265(6.48) accompanied by outward rotation of H6 using the conserved P267(6.50) as flexible hinge.[118] Changes in the interaction pattern of W265(6.48) upon rhodopsin activation are also indicated by changes in the chemical shifts of the ^{15}N-labeled tryptophan consistent with the indole nitrogens of W265(6.48) becoming less hydrogen bonded on going from dark rhodopsin to MII. NMR measurements also indicated breakages of the interhelical H-bonding interaction between E122(3.37) and H211(5.46) as a consequence of the establishment of intermolecular interactions with the β-ionone ring of the chromophore.[119] In contrast with the observations and conclusions of Patel et al., Spooner et al., on the basis of NMR determinations on ^{13}C-labeled C16 and C17 of retinal, concluded that the β-ionone ring becomes even more strongly restrained on activation.[120] Retaining the initial position on cis−trans isomerization, the chromophore ring would increase its steric interactions with the receptor, forcing the protein to adjust its structure around the ligand.[120]

NMR determinations also targeted the whole photoreceptor, providing the first evidence that activation of rhodopsin may involve differential dynamic properties of side chain versus backbone atoms.[121] In fact, NMR studies of the α-^{15}N-lysine-labeled receptor revealed large backbone motions in the inactive dark state. In contrast, indole side chain ^{15}N groups of tryptophans showed well resolved, equally intense NMR signals, suggesting restriction to a single specific conformation.[121] These results indicate that tryptophan side chains are more restricted in conformation than their backbone, suggesting that the indole side chain contacts, in part, contribute to restricting the conformation in a "locked" dark state, without fully restricting motional fluctuations in the overall molecule including the helix bundle itself.[121] Other high resolution information comes from solution and solid-state NMR spectroscopy of the intact

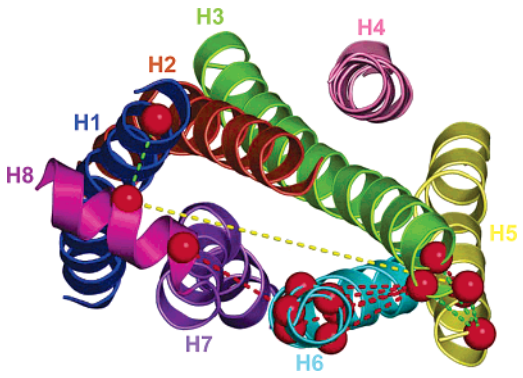


Figure 6. Cartoon representation of the seven-helix bundle and H8 of 1U19. The bundle is seen from the intracellular side in a direction perpendicular to the membrane surface. Red spheres indicate the α-carbons of the amino acids targeted by cysteine cross-linking experiments (reviewed in ref 96). Dashed red lines indicate inhibitory cross-links, and dashed green lines indicate cross-links compatible with activation, whereas the yellow dashed line indicates a cross-link that occurs only in light-activated rhodopsin (reviewed in ref 96). Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

mammalian photoreceptor rhodopsin in detergent micelles. These experiments suggest that the C-tail conformation observed in the crystal structure is no longer maintained upon phosphorylation, with the C-tail becoming disordered.[122]

A large number of intramolecular links, either cysteine disulfide bridges or Zn^{2+} chelated by substituted histidine side chains, were also engineered to connect the cytosolic ends of the helices, including H8. These cross-links were investigated in their ability to permit or inhibit activation (reviewed in ref 96). The observation that all the cross-links between H3 and H6 inhibit activation (dashed red lines in Figure 6) suggests that rhodopsin activation would require a change in the relative position of these two helices, most likely a separation of their cytoplasmic ends. In contrast, activation would require a change in the relative position of H3 and H5 without a significant separation. This was suggested by the findings that two inhibitory cross-links between these two helices were found close to activation permissive cross-links between the same helices (dashed green and red lines in Figure 6).[96] Cross-linking experiments also suggested that H3 might become less tilted upon activation, that is, more perpendicular to the plane of the membrane, causing its end to protrude more into the cytoplasm. Interestingly, one disulfide bond connecting residues 140(3.55) and 316 (in H8) was found only after photoactivation of rhodopsin (dashed yellow line in Figure 6).[123] This cross-link connected Cα-atoms that are 29 Å apart in the ground-state crystal structure, whereas the distance between Cα-atoms of disulfide-bonded cysteine residues is in the range 3.8−6.8 Å. One possibility is that light activation causes a significant conformational change that brings the two positions together.[96,123]

Site-directed spin labeling has been extensively used to explore the cytosolic domains of the photoreceptor.[96] This biophysical approach was used to assess the following changes induced by photoacti-

Table 1

residues[a]	dark[b]	inter-Cα X-ray[c]	light[d]	inter-Cα light[e]	1LN6[f]	Gouldson[g]	1ov0[h]	1BOJ[i]
139−248	**12−14**	8.75	**23−25**	19.75	16.59	15.35	11.83	10.02
139−249	**15−20**	11.89	**15−20**	11.89	16.20	13.62	11.28	12.96
139−250	**15−20**	10.80	**12−14**	6.30	12.41	13.98	8.04	12.39
139−251	**12−14**	8.83	**23−25**	19.83	13.63	16.89	10.53	12.22
139−252	**15−20**	12.17	**23−25**	18.67	15.96	19.31	13.32	15.41

[a] Amino acids in bovine rhodopsin. [b] Distances (Å) between spin labels determined for dark rhodopsin (reviewed in ref 96). [c] Distances (Å) between the Cα atoms measured in the crystal structure of dark rhodopsin (PDB code 1U19).[54] [d] Distances (Å) between spin labels determined light-activated rhodopsin (reviewed in ref 96). [e] Distances (Å) between the Cα atoms estimated for light-activated rhodopsin, based upon the differences determined between spin labels and measured between the α-carbons in dark rhodopsin (1U19). [f] Distances (Å) between the Cα atoms measured in the computational model 1LN6.[137,138] [g] Distances (Å) between the Cα atoms measured in the computational model by Gouldson and co-workers.[139] [h] Distances (Å) between the Cα atoms measured in the computational model 1OV0.[140] [i] Distances (Å) between the Cα atoms measured in the computational model 1BOJ.[320]

vation: (a) changes in the hydrophobic/hydrophilic nature of the environment of the labeled amino acids, (b) changes in their solvent exposure, (c) changes in the distances between pairs of spin labels, and (d) increase in their mobility. Increases in the mobility of the targeted side chains were interpreted as increases in their solvent exposure. The estimated distances between the spin labels in dark rhodopsin and MII suggest that an increase in distance of more than 10 Å between the Cα-atoms of V139(3.54), on one hand, and those of K248(6.31) and T251(6.34), on the other one, would characterize the transition from dark to MII (Table 1).[96] Our conversion of inter-spin distances into inter-Cα-atom distances is based on the comparisons of the results of experimental determinations on dark rhodopsin with measurements on the crystal structure (Table 1). All together, the observed mobility changes suggest that activation opens a cleft on the cytoplasmic side of the helix bundle.[124−127] More specifically, the postulated rigid-body tilt or translation of H6, moving its cytoplasmic end out from the bundle, would simultaneously increase exposure at the cytoplasmic end of H3 and decrease the exposure of some positions near the end of H5.[124,125] Mobility changes in the H1−I1−H2 region upon activation are relatively minor.[128] Very recently, site-directed spin labeling has also been used to explore the role of the salt bridge between the protonated Schiff base at K296(7.43) and its counterion E113(3.28). In detail, molecular sensors were placed at selected positions in the cytosolic domains to monitor the structural changes in such domains triggered by the breakage of the salt bridge in the retinal binding site.[129] All together, the results provide structural evidence that the salt bridge is a key constraint maintaining the resting state of the receptor and that disruption of the salt bridge is the cause, rather than a consequence, of the H6 motion that occurs upon activation.[129]

In summary, experiments on rhodopsin produced a crude qualitative picture of the structural changes that occur during receptor activation. In this scenario, the helical bundle opens at its cytoplasmic end, exposing various regions for interaction with the G protein; the strongest evidence indicates activation-induced separation of H3 and H6.[127,130] Overall movement of H6 probably exceeds that of H3, which is more constrained by its central position in the helix bundle (Figure 6). The primary result of photoisomerization of retinal, which triggers H6 motion, is suggested to be the breakage of the salt bridge between K296(7.43) and E113(3.28).[129] Other helices probably also adjust their positions upon activation as well. The activation model includes outward (away from the bundle) movements of the cytoplasmic ends of H3, H6, and H7. The inner faces of H2, H3, H6, and H7 become more exposed, whereas the cytoplasmic ends of H4 and H5 become less exposed (reviewed in ref 96). Recent in vitro experiments on rhodopsin−transducin recognition suggest that key interactions occur between the C-terminus of transducin and the inner face of H6 that contributes to form a solvent-exposed hydrophobic cleft.[127]

2.4. Computational Experiments on Rhodopsin

The crystal structure of rhodopsin is perhaps the most representative of a collection of states, which constitute the native conformational ensemble.

MD simulations on rhodopsin in an explicit membrane/water environment have attempted to investigate the conformational space around the native state of dark rhodopsin.[131] In detail, all-atom rhodopsin in a lipid/water environment was simulated for 15 ns. Different rhodopsin structures, including the incomplete A, C, and E chains of 1F88[13] and of 1HZX,[51] as well as a completed structure were used as input of calculations. Both the protonated and deprotonated forms of D83(2.50), E123(3.38), and H211(5.46) were probed. Simulations with and without internal water molecules were also carried out. Analysis was carried out over the last 5 ns trajectory. The computational study provided useful information on the membrane topology of the photoreceptor. A relevant outcome of this study is the finding that the cytoplasmic loops and the C-terminal tail, containing the G protein recognition and protein sorting sequences, exhibited high mobility, in marked contrast to the extracellular and transmembrane domains.[131] In fact, the average RMSD from the crystal structure during the last 5 ns of the 15 ns MD simulations is 1.86 Å for all the Cα-atoms of the protein. In contrast, the C-terminal tail, I3, and I2 exhibited, respectively, Cα-RMSDs up to 13 Å, 9.0 Å and 5.0 Å, whereas the central parts of the TM helices revealed Cα-RMSD values as low as 0.5 Å. Also, H8 underwent significant motions. The local deviations from the crystal structure and fluctuation amplitudes revealed correspondences between flexible and functional domains of the protein.[131]

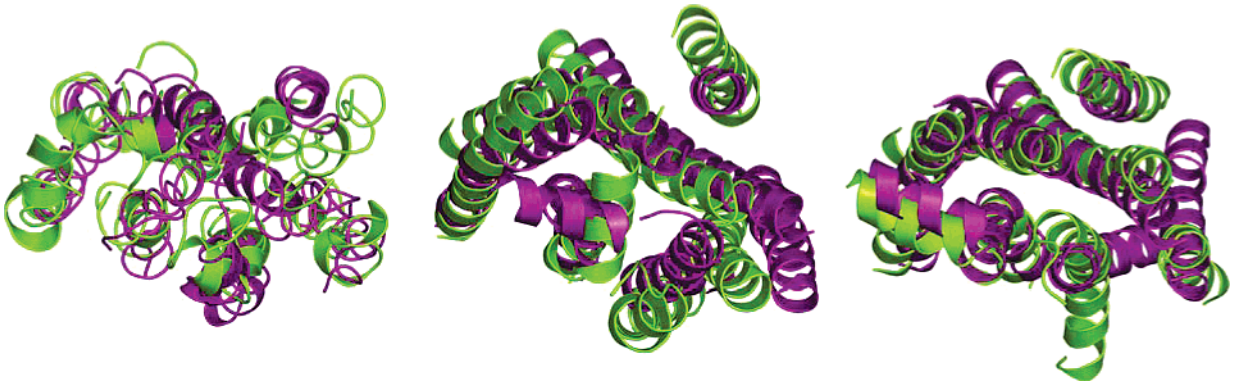


Figure 7. Cartoon representation of the seven-helix bundle and of H8 of computational models of the inactive (violet) and active (yellow-green) states of rhodopsin. In detail, on the left side, the models from Yeagle and co-workers deposited in the PDB as 1JFP (violet) and 1L6N (green) are shown;[136−138] in the central panel, the structures computed by Gouldson and co-workers are shown;[139] finally, in the right panel, the 1U19 structure (violet) is superimposed on the computational model of active rhodopsin, deposited in the PDB as 1ov0.[140] The helix bundles are seen from the intracellular side in a direction perpendicular to the membrane surface. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

In an analogous study, the analysis of a 40 ns MD trajectory of dark rhodopsin in an explicit membrane/water environment highlighted changes in the water accessibility of I3 and the C-tail, likely driven by helix motion.[132] In an attempt to approach as much as possible the real composition of the retinal rod outer segment disk membranes, which are rich in polyunsaturated lipids and cholesterol, Pitman at al. embedded the rhodopsin molecule in a bilayer composed of a 2:2:1 mixture of 1-stearoyl-2-docosahexaenoylphosphatidylcholine (SDPC), 1-stearoyl-2-docosahexaenoylphosphatidylethanolamine (SDPE), and cholesterol (i.e. 1 protein, 49 SDPCs, 50 SDPEs, 24 cholesterols, and 7400 waters) and produced a 118 ns equilibrated MD trajectory in the NVE ensemble. Consistent with the observations from in vitro experiments, in silico experiments showed that the protein breaks the lateral and transverse symmetry of the bilayer. Lipids near the protein preferentially reorient such that the unsaturated chains interact with the protein, while the distribution of cholesterol in the membrane complements the variations in rhodopsin's transverse profile, thus suggesting that cholesterol stabilizes the dark state of the photoreceptor without interacting with the protein.[133] These elegant studies prove an enormous effort to improve the reliability of the system, providing insights into the effects of lipids on receptor conformation. However, the reliability of the explicit lipid/water models remains undermined by the membrane−compartment dependence of lipid composition and lipid−receptor stoichiometry, in light of the ever-increasing evidences that GPCRs, including rhodopsin, exist as constitutive oligomers.[38,43]

Explicit membrane/water models make it even more difficult to detect, in the nanosecond time scale, the protein motions, which accompany MII formation. The latter event has been, indeed, demonstrated to occur in the millisecond time scale and may require the presence of transducin.[98,106] These limitations clearly emerge in recent computational attempts to simulate the active states of the photoreceptor.[134,135] External forces were introduced in explicit membrane/water simulations to induce the clockwise rotation of H6 (when seen from the intracellular side), predicted to be associated with MII formation.[135] Experimental-derived distance restraints were, instead, employed by Yeagle and co-workers to transform a computational model of dark rhodopsin (PDB code 1JFP)[136] into MII (PDB code 1LN6) (Figure 7, left).[136−138] The model of dark rhodopsin, 1JFP, was obtained by combining NMR determinations of rhodopsin fragments with computational modeling.[136] The RMSD of the main chain atoms of amino acids 40−348 in 1JFP and in the latest rhodopsin structure (1U19)[54] is 7.3 Å. Such a deviation is still high, i.e., 5.7 Å, if only the seven TM domains are considered, as differences between rhodopsin structure and the computational model concern both the secondary structure elements and their packing. Both computational models of the inactive and active rhodopsin states, 1JFP and 1LN6, are of poor stereochemical quality, indicative of limitations in the modeling approach. Instead of using a computational model, Gouldson and co-workers employed the crystal structure of rhodopsin for restrained MD experiments.[139] Also in this case, restraints were derived by in vitro experiments on rhodopsin and homologous GPCRs.[139] The results of this study suggested that the main changes in the receptor conformation on activation involve H4, H5, H6, and H7. Such changes include (a) clockwise rotation of H4; (b) displacement of H6 away from H3, accomplished by a straightening of H6 itself; (c) increased flexibility in the intracellular half of H7; and (d) a general opening of the intracellular part of the structure (Figure 7, central panel).[139] In line with the attempts to overcome the computational and time-scale limitations in unforced and unrestrained simulations of rhodopsin activation in an explicit environment, Nikiforovich and Marshall recently presented the results of a computational approach aimed at generating the active state of rhodopsin.[140] The approach started with simplified energy calculations in an effort to find a set of sterically and energetically reasonable options for the TM helix arrangements with *all-trans*-retinal. Various 3D models of the TM helix packing found by computations were then compared to limited site-

directed spin-labeling experimental data on rhodopsin activation, to identify the most plausible model of the TM helix bundle in the active state. The experimental data mainly used to drive model selection were light-induced changes in distances between spin labels in positions 139(3.54) and 248(6.31)−252-(6.35) (Table 1).[140] In the next step, all non-TM structural elements (i.e. tails and loops) were reconstructed, and after the entire "MII" structure had been relaxed, all other currently available additional experimental data, both mutational and spectroscopic, on the structure of the MII state of rhodopsin were used to validate the resulting 3D model (PDB code 1ov0; Figure 7, right).[140] Contrary to what the authors say, the analysis of such a computational model does not reveal a striking agreement with site-directed spin-labeling experimental data. In fact, the distances between the Cα-atoms of V139(3.54), on one hand, and those of K248(6.31), T251(6.34), and R252-(6.35), on the other one, in the computational model are respectively 11.83, 10.53, and 13.32 Å, whereas, according to the site-directed spin-labeling data, such distances should be about 19.75, 19.83, and 18.67 Å, respectively (Table 1).[96] Thus, the computational model does not show the predicted separation of the cytosolic extensions of H3 and H6, as compared to the dark state, and it should be taken with caution the author's proposal that such a computational model can be used as a template for modeling the active states of homologous GPCRs.[140]

In vacuo weighted masses MD was applied to a truncated form of the crystal structure of rhodopsin,[13] lacking all the nonprotein molecules, lacking all the intracellular and extracellular domains, except for E2, and holding the 11-*trans*-retinal, to produce MII.[141] The structure averaged over the last 100 ps of a 1.2 ns equilibrated trajectory and minimized was considered consistent with the available experimental data and used as a template to achieve what the authors called an "in silico activated" form of the 5-HT_{2A} receptor. Such a receptor model was then used for docking experiments with known activating ligands, without any further MD simulation.[141]

An alternative approach to build rhodopsin's photointermediates lumirhodopsin, MI, and MII, starting from the crystal structure of the dark state, consisted of applying swings to selected TM domains according to in vitro experimental data coupled with optimizations by MD/energy minimizations.[142,143] The putative models of MI, MIb (opsin), MI_{380}, and MII were then used to construct the structural models of the putative inverse agonist-, antagonist-, partial agonist-, and full agonist-bound forms of aminergic GPCRs.[144] We have concerns about the approach employed to model the rhodopsin intermediates, as it is significantly biased by human intervention. We, therefore, are also concerned about the use of such rhodopsin intermediates as templates for modeling functionally different states of the homologous GPCRs.

The substantially different computational models of activated rhodopsin share the breakage of the charge-reinforced interaction found in the dark state between R135(3.50), of the E/DRY motif, and E247-(6.30).[138−140] The breakage of such an interhelical salt bridge has been suggested to be a feature of the active states of different family A GPCRs, by in vitro and computational experiments.[61,62,64−69,72]

3. Computational Approaches to GPCR Model Building

3.1. GPCR Databases

Modeling of GPCRs has enormously profited from the availability of molecular class-specific databases capable of dealing with highly heterogeneous data on these receptors. The G protein-coupled receptor database (GPCRDB; http://www.gpcr.org/7tm/) is a Molecular-Specific Information System (MSIS) for GPCRs aimed at the collection and dissemination of GPCR related data.[145−147] It holds sequences, mutant data, and ligand binding constants as primary experimental data. Mutation data within the GPCRDB (http://www.gpcr.org/7tm/mutation/) are automatically extracted from the scientific literature.[148] In detail, the mutation data extracted from the literature are validated by plausibility filters and integrated into the corresponding Molecular Class-Specific Information System, where they are combined with structural and sequence information already stored in the database.[148] The GPCRDB is linked to the SWISS-PROT protein sequence database (http://expcsy.hcuge.ch/cgi-bin/search-7tm/), to the Olfactory Receptor DataBase (http://senselab.med.yale.edu/senselab/ORDB/), specializing in olfactory receptors, and to the GRAP and tinyGRAP mutant databases (http://tinygrap.uit.no/ and http://tinyGRAP.uit.no/).[149] The GRAP database allows searches for specific amino acid substitutions in specific proteins within user-specified groups of receptors.[150−153] The database also contains searchable information on quantitative ligand-binding data and qualitative descriptions of the effect of the mutation on agonist binding and signal transduction. The GRAP database has not been updated recently. Mutant database development moved on to the tinyGRAP database, which holds only the most basic information on the receptor type, literature reference, and mutant type.[150−153]

Computationally derived data such as multiple sequence alignments, 3D models, phylogenetic trees, and two-dimensional (2D) visualization tools are added to the GPCRDB to enhance the database usefulness. The GPCRDB is also linked to a fingerprint database that determines if a given query sequence belongs to one of the characterized superfamilies, families, or receptor subtypes (http://www.biochem.ucl.ac.uk/bsm/dbbrowser/GPCR/). Another bioinformatics tool useful for GPCR modeling through the GPCRDB is the Viseur program, which allows one to interactively visualize and/or modify amino acid sequences, TM areas, alignments, models, and results of mutagenesis experiments in an integrated environment.[154] The GPCRDB is also linked to the SWISS-MODEL server for the automated modeling of the TM helix bundle of GPCRs (http://swissmodel.expasy.org//SWISS-MODEL.html).[155] Very recently, the developers of the GPCRDB have created the MSIS for GPCR Interacting Partners (G proteins and

RAMPs), the GPCRIPDB (http://www.gpcr.org/GPCRIP/).

A support vector machine (SVM)-based method, GPCRpred, has recently been developed for predicting families and subfamilies of GPCRs from the dipeptide composition of proteins.[156] The method classified GPCRs and non-GPCRs with an accuracy of 99.5% when evaluated using 5-fold cross-validation. The method is also able to accurately classify GPCRs. A server for recognition and classification of GPCRs based on multiclass SVMs has been set up at http://www.imtech.res.in/raghava/gpcrpred.[156]

A list of the URLs for sequence-based classification of GPCRs is provided and critically discussed in ref 157.

A database extremely useful for the study of G protein−GPCR interactions has recently been set up with a focus on G proteins and their coupling specificity with GPCRs (http://bioinformatics.biol.uoa.gr/gpDB).[158]

3.2. Comparative Modeling Using the Bacteriorhodopsin Structure as a Template

A high resolution structure of the light-driven proton pump from *Halobacterium halobium* bacteriorhodopsin (BRD) has been available since 1990.[159] Since BRD possesses seven TM α-helices and binds the retinal chromophore, it has been considered a bacterial homologue of vertebrate rhodopsin. Reports concerning comparative modeling of GPCRs using the BRD structure as a template have appeared in the years that preceded the release of the first electron density 2D map of rhodopsin,[9,160−170] in the years between the releases of the first 2D and the first 3D electron density maps of rhodopsin,[9,12,171−193] in the years between the releases of the 3D map and of the crystal structure of rhodopsin,[13,194−201] and even in the years after the release of the rhodopsin crystal structure.[202−205] However, BRD is a proton pump, is not coupled to a G protein, and does not even display remote sequence similarity with any GPCR. Two hypotheses were made to demonstrate the existence of a link between GPCRs and BRD.[206,207] One of these hypotheses was formulated by Pardo and co-workers, who suggested that the sequence homology in the helical region of BRD and GPCRs would be greater if the sequential ordering of the helices is ignored.[206] The authors, hence, proposed a mismatch of the helices, in which BRD H2, H3, H4, H5, H6, and H7 matched, respectively, with GPCR H4, H5, H6, H1, H2, and H3. Such a mismatch between BRD and GPCR sequences was suggested to be caused by an exon shuffling event that occurred during the evolution of GPCRs and BRD from a common ancestor.[206] The second hypothesis suggests that H5, H6, and H7 originated from H1, H2, and H3 as a result of ancestral gene duplication, leading to homologies between helices H1, H2, and H3 in BRD and helices H5, H6, and H7 in GPCRs.[207] An alternative helix-mismatching alignment between GPCRs and BRD was obtained by Metzger and co-workers, and this made the authors trust comparative modeling based upon the BRD structure.[208] The authors, indeed, built a model of the κ opioid receptor, which has been used in computational experiments until recently.[199,203,209,210] Sequence analysis studies led, however, to the conclusion that there is no significant evidence for similarities between BRD and GPCR, regardless of the ordering of the helices, and that, hence, BRD cannot be used as a template for comparative modeling of GPCRs.[211] These conclusions found strong support in the evidences from the first 2D electron density map of rhodopsin at 9 Å resolution, which showed clear differences from the electron density map of BRD resolved at comparable resolution.[9] These structural evidences, combined with the information from alignments of a significant number of GPCR sequences, suggested that GPCRs share a common architecture of the seven helices that differs from that of BRD, strongly underscoring the drawbacks in the use of BRD as a template for the construction of molecular models of GPCRs.[10]

Comparisons of the high resolution structures of BRD and rhodopsin,[54] in fact, confirmed the marked structural dissimilarities between the two photoreceptors, in line with the evidences from the low resolution electron density maps.[9] Indeed, the seven helices differ in length and location of the breakages. Furthermore, loops and tails are completely different in length, amino acid composition, and structure. In line with these data, the sequence alignment obtained from the superimposition of rhodopsin and BRD structures shows mismatches in the seven helices, concerning also the lysines in H7, which are covalently bound to the retinal chromophore in both the photoreceptors (Figure 8). Along the same line, the highly conserved amino acid motifs characterizing family A GPCRs (colored in red in Figures 1 and 8) are not present in BRD. This suggests that, even if proteins may share the same fold even in the absence of any sequence similarity,[212,213] BRD is not the proper template for modeling family A GPCRs, including rhodopsin. On the basis of these considerations, any attempt to overcome the drawbacks of using BRD as a template, by first modeling rhodopsin on BRD, then the β_2-AR on the rhodopsin model, and finally the target receptor on the β_2-AR model, is likely to fail.[179,181−183,189,205] Critical comparisons between a B_2 bradykinin receptor achieved through this approach and a model based upon the rhodopsin structure were, however, unable to distinguish a preference for one of the two models.[205] This could be due to the low resolution of the data employed for evaluating the models. In more recent works, the authors adjusted their original BRD-based models to incorporate high resolution structural information on rhodopsin.[214−216]

In summary, due to the lack of functional, sequence, and structural similarity between rhodopsin and BRD, comparative modeling of GPCRs by using the BRD structure as a template is likely to produce unreliable models and there is no way to change the alignment so as to improve the model. Despite this evidence, BRD has been used as a template for achieving 3D models of GPCRs until very recently and it was even concluded that the receptor model based upon the BRD structure is more consistent with site-directed mutagenesis data than the model

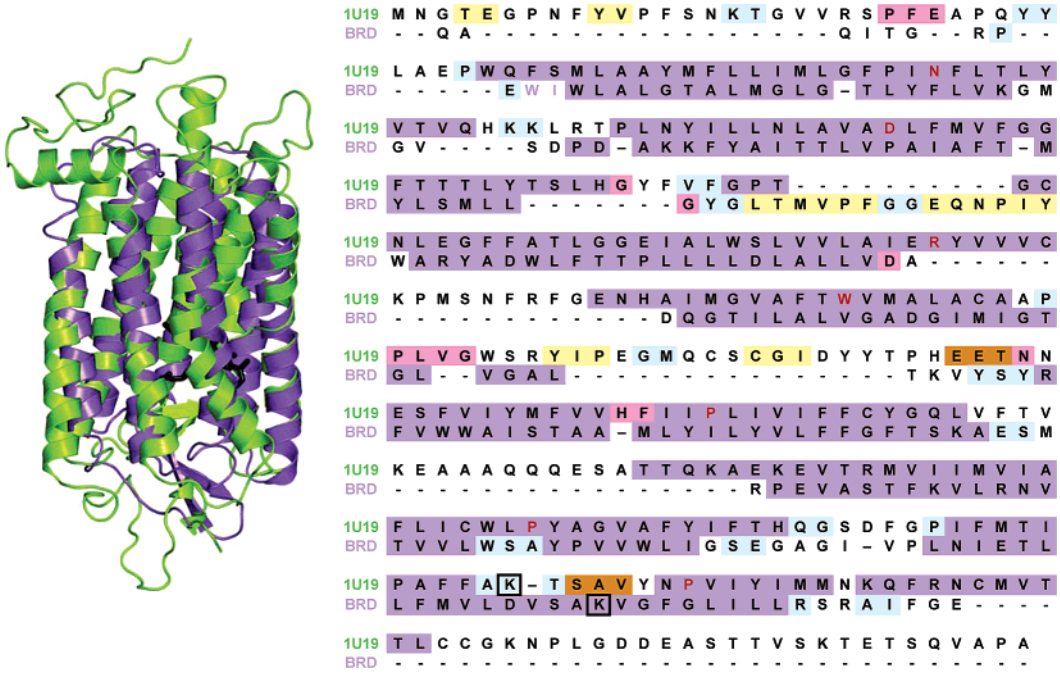


Figure 8. (left) Cartoons of the superimposed structures of rhodopsin (1U19, in yellow-green)[54] and of BRD (PDB code 1BRR, violet).[582] The retinal molecules are also shown by black sticks. (right) Sequence alignment which results from the α-carbon atom fit shown on the left side. In this alignment, color highlights indicate the secondary structure computed on the superimposed crystal structures of the two proteins. In detail, violet means canonical α-helices, and yellow stands for strand, whereas cyan, orange, and pink indicate respectively type 3-, 4-, and 5-turns. Black boxes enclose the lysine residues covalently linked to retinal in both the photoreceptors. As for the rhodopsin sequence, the most conserved amino acids in each TM helix are colored in red. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

based on the crystal structure of rhodopsin.[204] This exceptional conclusion may reflect the inadequacy of low resolution site-directed mutagenesis data for evaluating the reliability of atomistic models.

3.3. Ab Initio Modeling of GPCRs

Because of the lack of a proper homologous template for comparative modeling of GPCRs, various ab initio approaches have been developed for computational modeling of GPCRs (reviewed also in ref 217).

Very early ab initio modeling experiments on the D_2 dopamine, β_2-adrenergic, and 5-HT_2 serotonin receptors led to predictions of either clockwise (when viewed from the extracellular side) or nonsequential helix-packing arrangements, inconsistent with both the BRD and rhodopsin structures.[218−222] Lybrand and co-workers were the most enthusiastic in pursuing their ab initio "clockwise" model.[219,223−228] Their computational modeling approach is summarized as follows. The seven α-helices are built in a canonical conformation, oriented approximately perpendicular to the plane of the bilayer in a continuous bundle arrangement and bundled together, in a way that hydrophobic residues are exposed to the lipids. Then the loops are modeled by restrained MD and added to the helix bundle. Topological and physical properties and low resolution information from site-directed mutagenesis and biophysical studies are used as constraints for helping to reduce the number of helix-bundle candidate structures from about 1500 to only 10−20 structures.[219,223] These best candidate structures are then subjected to energy minimization and MD refinement, leading to the final selection of only one structure, based upon the available experimental data. In the first study, the clockwise helix arrangement was proposed as the most reliable, whereas in successive studies it was always flanked by the "counterclockwise" model.[219,223−228] The necessity of probing both the alternative arrangements was dictated by the author's judgment that the available experimental data were inadequate to drive selection of the proper stereochemistry and that, at least for the β_2-AR, the clockwise model appeared to better explain ligand selectivity data.[223] Despite the indications in support of the "counterclockwise model", inferred from computational[10,11,83] and in vitro experiments on several GPCRs,[229−232] the authors recognized the inconsistency of the "clockwise" model only when the first crystal structure of rhodopsin was released.[228] They, however, pursued their ab initio approach, until very recently.[233−236]

One of the most successful approaches to GPCR-structure prediction was that used by Baldwin and co-workers, which consisted of integrating the information from sequence analyses with that from the 3D electron density map of frog rhodopsin and that from biochemical and biophysical experiments on the members of the rhodopsin family.[10,11] This approach led to the building of a Cα-atom model of the seven-helix bundle of rhodopsin and a Cα-atom template for comparative modeling of GPCRs.[11] Retrospective analysis validates Baldwin's model of rhodopsin, as the Cα-RMSD between rhodopsin structure (i.e. 1U19) and the model is 2.96 Å (Figure 9, left). This deviation is due, at least in part, to the regularity of the seven helices in the model, which is different from

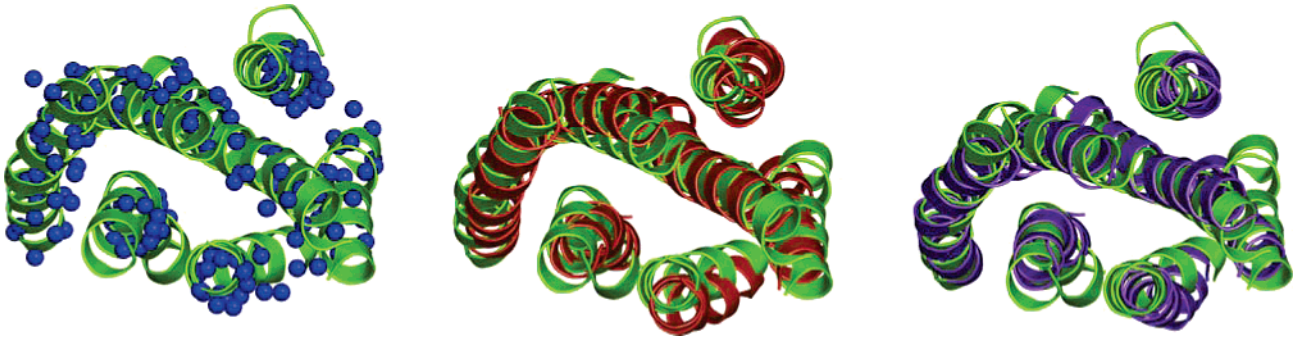


Figure 9. Cartoon representation of the seven-helix bundle and of H8 of the rhodopsin structure 1U19 (yellow-green)[54] superimposed on the computational models by Baldwin and co-workers (left side, in blue),[11] by Herzyk and Hubbard (central panel, in orange),[310] and by Pogozheva and co-workers (right side, in violet).[320] The helix bundles are seen from the intracellular side in a direction perpendicular to the membrane surface. Superimposition between rhodopsin and each of the three computational models has been done on the α-carbons of the following amino acid stretches: 38−63, 71−95, 111−140, 151−175, 204−223, 250−274, and 288−310. The Cα-RMSDs resulting from the three matches shown in the left, central, and right panels are respectively 2.96, 2.91, and 2.68 Å. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

the crystal structure (Figure 9, left). Indeed, the only helix bend correctly predicted by Baldwin's study was the kink at P267(6.50). The results of Baldwin's studies[10,11] and the electron density 2D and 3D maps of rhodopsin[9,12] have been milestones and invaluable sources of information for GPCR modeling in the years that preceded,[237−278] and even in those that followed, the release of the crystal structure of rhodopsin.[279−292]

The helical wheel projection models, which were inferred from an earlier Baldwin study,[10] represented the background of our ab initio building of the seven-helix bundles of several GPCRs of the rhodopsin family.[293] The study consisted of probing different helix arrangements by MD and then selecting the average arrangement more consistent with Baldwin's helical wheel projection model and the 2D electron density map of bovine rhodopsin at 9 Å resolution.[9,10,293] These computational experiments, which highlighted the highly conserved polar amino acids as drivers of the helix packing through H-bonding networks, constituted the first step of an iterative procedure, consisting of progressive domain additions and modifications in the receptor model, to incorporate the ever-increasing experimental evidences on rhodopsin and the homologous GPCRs.[294] The preferential targets of the approach were the α_{1b}-adrenergic receptor (α_{1b}-AR), the M_1- and M_3-muscarinic receptors, the oxytocin receptor (OTR), and the luteinizing hormone receptor (LHR).[57,58,60,294−305] The last updates of the ab initio models of these receptors held the structural information derived from the electron micrographs of 3D frog rhodopsin crystals[12] and from the last GPCR sequence analysis from Baldwin and co-workers.[11,60,294,302−305]

Weinstein and Ballesteros proposed an ab initio approach to GPCR model building based upon the integration of information about the primary, secondary, and tertiary structural properties of GPCRs as well as inferences from the experimental probing and biophysical analysis of TM proteins.[55]

Alkorta and Loew employed an ab initio approach to model rhodopsin and the δ opioid receptor, consisting of different steps, including (a) multiple sequence alignment, (b) calculation of a variability profile of the aligned sequences, (c) use of the variability profile to identify the boundaries of the TM regions, (d) prediction of their secondary structure, (e) helix bundling, (f) predictions of side chain conformations, and (g) structure refinement.[306,307] Helix packing into a bundle was based on the assumption that helices are packed in a sequential order and in antiparallel fashion except for H1 and H7, that they are all canonical, and that they are all perpendicular to the putative membrane surface.[306,307]

Donnelly and co-workers reported a 3D model of a GPCR based on the helix arrangement observed in the projection structure of rhodopsin.[308,309] The model was created using a method that detects helical periodicity in sequence alignments, using amino acid substitution data derived from protein structures. The method, which also compares the relative directions of the conserved and hydrophobic faces, was first tested on BRD and proved to be successful at orientating the helices. As the method detects the outside face of each helix, the positions of charged residues on this face were used to detect the points at which these helices contact the more polar regions of the phospholipid headgroup/aqueous interface. The projection structure of rhodopsin and the connectivity predicted by Baldwin were also considered in the model building.[308,309] The Cα-RMSD between the rhodopsin model obtained through this approach and available at the GPCRDB (http://www.gpcr.org/7tm) and the rhodopsin structure (i.e. 1U19) is 6.89 Å.

Herzyk and Hubbard developed a rule-based automated method for modeling the structure of the seven TM helices of family A GPCRs.[310] With this method, the structures are generated by using a simulated annealing Monte Carlo (MC) procedure that positions and orients rigid helices to satisfy structural restraints. The restraints are derived from analysis of experimental information from biophysical studies (including electron density maps) on native and mutant proteins, from analysis of the sequences of related proteins, and from theoretical consideration of protein structure. The four main steps in this approach are as follows: (a) analysis of the available experimental and theoretical data to derive geometrical restraints, (b) employment of a

protein representation adequate for applying the restraints, (c) construction of a penalty function that efficiently penalizes violations of the restraints, and (d) optimization of the penalty function to find a family of structures that best satisfies the restraints. The method was first validated by generating a model of BRD, characterized by a Cα-RMSD of 1.87 Å, from the structure determined by electron microscopy.[159] Calculations were then carried out by using experimental and theoretical information available for bovine rhodopsin to assign the helices to a projection density map[9,10] and to produce a rhodopsin model useful as a template for comparative modeling of homologous GPCRs.[310] The Cα-RMSD between the TM model of rhodopsin, generated by such a method and available at http://www.gpcr.org/7tm, and the crystal structure of rhodopsin (i.e. 1U19) is 2.91 Å (Figure 9, center). These results validate the approach by Herzyk and Hubbard in its ability to predict the architecture of TM α-helical proteins. Even if the seven helices in the computational model are shorter than those in the rhodopsin structure and are all canonical, the topography of the highly conserved amino acids is quite conserved in the computational model and in the crystal structure. The rhodopsin model by Herzyk and Hubbard has been used as a template for modeling several homologous GPCRs.[311−313] Moreover, this automated approach, implemented in the program PANDA, was used to achieve computational models of different GPCRs of the rhodopsin family, including the melanocortin-1 receptor (MC1-R).[312]

Peitsch and co-workers developed the automated protein modeling server, SWISS-MODEL.[314] With this approach, models are constructed in a two-stage process. In the first stage, the seven TM helices are represented as being idealized and rigid. Structural restraints derived from theoretical and experimental data are then used to fit the helices together. A penalty function is used to measure any violations to the structural restraints. This penalty function is then globally optimized using a MC simulated annealing procedure to generate an optimal model. In the second stage, the optimal model is converted into a full-atom model by the ProMod package.[155] The Cα-RMSD between the rhodopsin model generated by such an automated method and the rhodopsin structure (i.e. 1U19) is 4.52 Å. The SWISS-MODEL server has been widely used to build GPCR models aimed at mapping the ligand binding sites.[315−319]

An elegant and successful approach was that employed by Mosberg's team to predict the architecture of the seven-helix bundle of several GPCRs of the rhodopsin family.[320−322] This approach, based on the use of distance restraints, as in calculations of protein structures from NMR spectroscopy data, consists of iterative distance geometry refinements of an approximate initial receptor model, using an evolving system of H-bonds. The rationale for identifying the required constraints is based on the presence of numerous polar residues in the TM hydrophobic α-helices of GPCRs and on the knowledge that polar side chains of proteins buried from water have a strong tendency to form H-bonds.[320] In TM α-helices, backbone peptide groups are already paired, whereas the polar side chains must interact with each other to form intra- or interhelical H-bonds. The H-bonding pairs can be identified from the analysis of sequence alignments as polar residues in TM segments, which appear and disappear in a correlated manner, and by using approximate receptor models to exclude all spatially distant residues from the list of possible correlations. H-bonds thus identified can be applied as distance restraints for the packing of the TM α-helices using the distance geometry algorithm. In detail, the approach consists of the following stages: (a) construction of the initial "crude" receptor model using electromicroscopy and a few site-directed mutagenesis and cross-linking data; (b) calculation of the average seven-helix bundle model for rhodopsin-like GPCRs, by using an iterative distance geometry refinement of the initial model with an evolving system of interhelical side chain H-bonds formed by various GPCRs and collectively applied as distance constraints; and (c) distance geometry calculations of the TM domain of the target receptor from its own H-bonds and using the "average" GPCR model to restrain the relative positions of the helices.[320] The approach was used to build a model of rhodopsin and of 26 rhodopsin-like GPCRs, including the opioid receptors.[320−322] Retrospective analysis supports the validity of the approach, as the Cα-RMSD between the rhodopsin model (PDB code 1BOK) and the rhodopsin structure (i.e. 1U19) is 2.68 Å (Figure 9, right). The model 1BOK was used as a template to achieve computational models of the oxytocin and vasopressin receptors.[323]

Another good prediction by Pogozheva and co-workers was that E2 of the opioid receptors would assume a β-hairpin structure that would partially cover the ligand binding cavity between H3 and H7, coming from H4 toward H3 and returning back to H5.[321] The predicted loop also adopts a nonregular structure at its N- and C-terminal extremities. A significant sequence similarity exists concerning this loop between the opioid receptors and rhodopsin, thus suggesting structure similarity. In the rhodopsin structure, E2 adopts a structure similar to that predicted by Pogozheva and co-workers.[13,321] The authors also inferred that both branches of the loop forming a U-like shape are too short to form any additional α-helices.[321] Collectively, these structure predictions, which are expected to be correct in light of the significant sequence similarity concerning E2 between rhodopsin and the κ opioid receptor, disagree with predictions by Paterlini and co-workers[210] and by NMR structure determination on the isolated E2 of the κ opioid receptor.[324] Pogozheva and co-workers employed their distance geometry-based approach also for predicting the structure of MII.[320] Inconsistent with the evidences from in vitro experiments (reviewed in ref 96), the model of the active state of rhodopsin, deposited in the PDB as 1BOJ, does not show any significant deviation in the backbone conformation as compared with the model of the dark state (PDB code 1BOK), obtained by the same approach (Table 1). Thus, the method appears inadequate for predicting the active state of the photo-

receptor and the predicted MII model should not be used as a template for modeling the active states of homologous GPCRs, as recently reported.[325]

A program for packing the TM helices of GPCRs from the electron density map of rhodopsin, named BUNDLE, was also proposed.[326,327] The approach consisted of the following steps: (a) location of the centers of the helices according to the low resolution electron density map; (b) calculation of the tilt angle of each helix on the basis of the elliptical shape observed by each helix in the map; (c) definition of a local coordinate system for each of the helices; (d) packing of the helices in an antiparallel orientation; (e) rotation of each helix through the helical axis in such a way that its hydrophobic moment points in the same direction of the bisector formed between three consecutive helices in the bundle; (f) rotation of each helix through an axis perpendicular to the helical one to assign a proper tilt; and (g) translation of each helix to its center deduced from the projection map.

A different approach, based upon the integration of sequence analysis and computational molecular modeling, has been very recently proposed to build TM α-helix bundles from low resolution electron density maps.[328] The approach is based on the observation that, typically, the lipid-exposed faces of TM proteins are evolutionarily more variable and less charged than the core. On the basis of this rule, the authors developed score functions and automated methods for orienting TM helices, for which locations and tilt angles have been determined using cryo-electron microscopy data (cryo-EM). The method was parametrized with the aim of retrieving the native structure of BRD from among near- and far-from-native templates. It was then tested on proteins that differ from BRD in their sequences, architectures, and functions, such as the acetylcholine receptor and rhodopsin. The predicted structures were within 1.5−3.5 Å of the native state in all cases. In particular, for rhodopsin, the Cα-RMSD between the highest scored conformation and the native state was 1.5 Å.[328] These encouraging results led to the conclusion that such a computational method can be used in conjunction with cryo-EM data to obtain approximate model structures of TM domains of proteins, for which a sufficiently heterogeneous set of homologues is available.[328]

The MembStruk protocol for predicting structures of GPCRs has been recently proposed.[329−331] Such a quite elaborate protocol consists of the following steps: (1) prediction of the TM regions; (2) construction and optimization of individual helices; (3) assembly of the seven-helical TM bundle; (4) coarse grain optimization of the TM bundle; and (5) addition of interhelical loops and optimization of the full structure. Step 1 is accomplished by means of the TM2NDS program that determines the TM regions in GPCRs using hydropathicity analysis, combined with input from multisequence profiles. In step 2, the canonical right-handed α-helices are then built with extended side chains and subjected to torsion angle optimization. In step 3, each helical axis is oriented according to the 7.5-Å electron density map of bovine rhodopsin.[332] The hydrophobic moments of the optimized helical bundle are aligned so that the net hydrophobic moment of each helix would be pointing outward toward the membrane from the center of mass. Step 4 consists of coarse-grain rotations of the helical orientations, starting with the directions of the net hydrophobic moment of each helix from step 3. Each helix is rotated through a grid of rotation angles about its helical axis. The total energy of this helix in the field of all of the other helices (fixed) is minimized using conjugate gradients. After finding the optimum configuration for each specific helix, a second cycle is initiated (seven such optimizations) and continued until the energy converges. Layers of explicit lipid molecules (52 molecules of dilauroylphosphatidylcholine lipid) are then added and optimized with the current configuration of the seven helices. Then, to achieve proper packing of the TM helices, the seven helix-bilayer complex is further optimized with rigid-body MD of the seven helices and lipid for 100 ps. In step 5, following the rigid-body dynamics, loops are added to the helices by using the WHATIF software.[333] The possible disulfide bridges are added. Then, addition of the side chains for all of the residues is carried out, followed by a full-atom MD optimization of the structure, with the explicit lipids. The protocol has been benchmarked on the BRD and rhodopsin structures.[330] As for rhodopsin, the RMSD in coordinates (CRMS) of the α-carbons in the predicted structure of the photoreceptor for the residues in the TM domains is 3.1 Å compared with the crystal structure with a resolution of 2.8 Å (PDB code 1F88).[330] Including the loops, the overall CRMS is 8.3 Å.[330]

Another ab initio modeling approach for packing TM helical bundles, called PREDICT, has been recently developed.[334,335] Without relying on sequence similarity to any other protein of known structure, the algorithm predicts the native 3D conformation of a protein using only its amino acid sequence and the physicochemical properties of the membrane environment. The concept of structural "decoys" is employed, consisting of generating many alternative possible conformations and optimizing them simultaneously to ensure that the algorithm identifies the correct structure without risk of ending up trapped at a local minimum. The core of PREDICT consists of two main steps: generation of decoys, followed by optimization and scoring of the decoys. In the first step, the program automatically produces hundreds to thousands of possible TM conformations (decoys). In this step, a coarse 2D grid search is conducted over the receptor conformational space, followed by initial optimization of the hydrophobic moments and protein−protein interactions in each decoy conformation, and generation of the initial 3D structures. In the second step, all the decoys are optimized and the most stable structures are identified according to the PREDICT energy function. The final model is further optimized using MD simulations and virtual complexation with a known ligand. The PREDICT algorithm was used to create an all-atom model of bovine rhodopsin that was further optimized using 300 ps of MD simulation without the retinal. The Cα-RMSD

between the retinal-bound TM model of rhodopsin and the crystal structure was 2.9 Å.[334]

Very recently, a two-step approach to assemble the TM helices of integral membrane proteins, including rhodopsin, has been presented.[336] In the first step, the conformational space of membrane protein folds matching a set of distance constraints is explored to provide initial structures for local conformational searches.[336] In step 2, the helical bundles from step 1 are refined using a Monte Carlo Simulated Annealing (MCSA) protocol designated for local minimization of an empirical structure-based penalty function. Applying this method to the seven rhodopsin helices, by using 27 distance constraints from biophysical experiments on the photoreceptor, reduced the approximately 7.0×10^{11} possible bundle configurations to 87 helical bundles with Cα-RMSD ranging from 4.3 to 9.5 Å. After MCSA refinement, the bundle with the lowest penalty function (i.e. with a penalty function of 3.3) had a Cα-RMSD from the known structure of 4.1 Å. It was, however, noticed that the best nativelike bundle, i.e., with a Cα-RMSD of 3.2 Å, was characterized by a penalty function of 1003.3. It was, therefore, concluded that stopping the refinement when the penalty function is between 1000 and 2000 would be worthy and more effective for seven-helix bundles.[336]

Most of the methods described above, which all require some input from structural determinations such as electron density maps or from biophysical and biochemical experiments, have proven to be effective in predicting the architectures of the seven TM helices in rhodopsin. Some of these methods have been invaluable in the years that preceded the release of the high resolution structure of rhodopsin. In fact, the models of the TM domains of rhodopsin achieved by these approaches, in particular Baldwin's Cα-atom model, have been used as templates in most of the comparative modeling studies done from 1993 to 2004. Hence, electron microscopy data on bovine and frog rhodopsin have been the main experimental foundation for ab initio and comparative modeling of GPCRs over one decade. Now that high resolution structures of rhodopsin are available, it is worth using those structures as templates for achieving initial models of the homologous GPCRs, instead of ab initio approaches, which necessarily incorporate low resolution information.

3.4. The Functional Microdomain Approach to GPCR Modeling

In the years 1997−2005, an approach consisting of extensive conformational searches was presented, based on simulated annealing MC runs with the conformational memories (CM) method or MD simulations on isolated receptor domains, either a helix or a loop. This was done to overcome limitations inherent in both the low resolution level of site-directed mutagenesis results and computational modeling of the whole receptor structure.[59,61,84,337−349] The assumption of the approach is that GPCRs are made of structural microdomains characterized by discrete function and, hence, treatable as independent units. In this context, extensive calculations on such microdomains are assumed to provide a reliable picture of the microdomain itself in the context of the whole receptor molecule.[342]

The approach was used to investigate different aspects of GPCR function and to interpret the results of in vitro experiments. In particular, it was used to predict the conformational behavior of H7 in the 5-HT_{2A} serotonin receptor.[84] In detail, two different initial models of H7 of the receptor were built: one incorporating the structural features of the (N/D)P motif inferred from a database search and the other holding a regular Pro-kink. The two structures were used as starting points for two identical sets of CM simulations. For the simulations, H7 was divided into three regions: (1) an entirely flexible region (A7.47−N7.49); (2) two flanking semiflexible regions (i.e. S7.45−S7.46 and P7.50−Y7.53) where the ϕ and ψ dihedral angles were constrained ±20° around the values −63.0 and −41.6, respectively; and (3) the intracellular and extracellular ends that were kept fixed. The CM procedure consisted of four steps: in step 1, a classical Metropolis MC simulation was performed in torsion space at 10 000 K for 1 000 000 steps, collecting 64 structures. In step 2, a simulated annealing from 10 000 K to 582 K was performed on both sets of 64 structures, providing two separate sets of dihedral angle maps. In step 3, two biased Metropolis MC simulations were done using the set of populations obtained from all data. At each biased Metropolis MC step, three randomly selected dihedral angles were assigned a new value, by using a biased temperature annealing method, leading to two collections of 100 structures. The fourth step consisted of cluster analysis of the two sets of 100 structures. The main outcome of this study was that the conserved N/DPxxY motif in this helix is the major determinant for deviation of H7 from ideal helicity.[84]

CM simulations of isolated TM helices of the D2 dopamine receptor (D_2R) were extensively used to rationalize the results of the substituted-cysteine accessibility method applied to H2, H4, and H6 of the D_2R.[61,339,340,344] Moreover, CM simulations of H6 done in combination with in vitro experiments were instrumental in inferring the helix motions associated with activation of the β_2-AR and the $5HT_{2A}$ receptor.[61,66,337,338,345]

MC simulations on the isolated H3 of the gonadothropin-releasing hormone receptor (GnRH), in combination with in vitro site-directed mutagenesis experiments, highlighted the intrahelix charge-reinforced H-bond between D3.49 and R3.50, of the E/DRY motif, as a feature of the inactive state of the receptor. Interpreting the results of computations in the context of a TM model of the GnRH receptor led to the speculation that, in the active states, the E/DRY arginine would lose the interaction with the adjacent aspartate, establishing new interactions with the highly conserved D2.50.[59]

A similar computational approach was employed to infer the structure/dynamics differences between the wild type and a mutated form of I2 of the 5-HT_{2C} receptor, which might correlate with the lower basal activity of the mutant. Computations revealed differences in the conformational space explored by the

two different forms of I2, and it was speculated that this would imply differences in G protein recognition.[346]

CM calculations on isolated H6 of the CB1 cannabinoid receptor highlighted the conformational behavior of the highly conserved W6.48 as responsible for the functional state of the receptor.[349] The results of computations suggested that the presence of the constitutive activity in the CB1 receptor was due to the absence of aromatic amino acids at positions i−4 (i.e. F6.44) and i+3 (i.e. F6.51) with respect to W6.48. This would provide W6.48 with higher conformational mobility, with a W6.48 trans χ_1 preferred. The χ_1 g+ → trans transition has been indicated to be linked to rhodopsin activation by spectroscopic studies.[350] In rhodopsin, the presence of F6.44 and F6.51, forming an aromatic cluster with W6.48, would restrict the conformational freedom of the conserved tryptophan, thus contributing to the lack of constitutive activity in the photoreceptor. It was also inferred that the W6.48 χ_1 g+ → trans transition is correlated with the degree of kinking of H6, with the bending being smaller with W6.48 χ_1 trans, thus suggesting that H6 straightens upon activation.[349] The inferences of these computational experiments were found consistent with similar calculations on H6 of the β_2-AR.[345] Interpreting the results of CM simulations on H6 in the context of a TM model of the CB1 receptor based upon the rhodopsin structure led to the hypothesis that F3.36 χ_1 trans helped maintain W6.48 χ_1 g+, hence stabilizing the inactive state of the receptor. It was, thus, suggested that the W6.48/F3.36 interaction may act as the "toggle switch" for CB1 activation, with W6.48 χ_1 g+/F3.36 χ_1 trans representing the inactive state and W6.48 χ_1 trans/F3.36 χ_1 g+ representing the active state of the CB1 receptor.[349] The CM method was also used to provide an explanation for the apparent deviation from the ideal helicity of H2 in the CB2 receptor.[351] Computational modeling pinpointed S2.54 as the location responsible for possible conformational differences concerning H2 between CB2 and rhodopsin.[351]

MD simulations of model α-helices in a nonpolar environment were done to investigate the structural effect of the TXP motif in H2 of the CCR5 chemokine receptor and, hence, rationalize the functional effects of mutating this motif.[343] A similar approach was used to compare the dynamic behavior of H3 holding the rhodopsin sequence with that of the same helix with the 5-HT_{1A} sequence.[348] The results of simulations on the isolated helix in a hydrophobic environment were interpreted in the context of the rhodopsin structure, and it was inferred that the 5-HT_{1A} H3 tends to bend toward H5, whereas the rhodopsin H3 does not. This structural/dynamic divergence, which would allow H3 and H5 to be properly bridged by the cationic neurotransmitter serotonin, was attributed to the presence of the conserved C(3.36)T(3.37) motif, a feature of cationic neurotransmitter GPCRs and not of rhodopsin.[348]

We find interesting the use of an extensive conformational sampling method to infer the structure/dynamics features of helices, which carry conserved amino acids known to play important structural/functional roles. One must, however, be cautious in interpreting the results of simulations on isolated helices or loops in the context of the whole helix bundle. In fact, the degrees of freedom and the dynamics of an isolated helix or loop may be significantly different from those of the same helix or loop in the folded protein. Along the same lines, the inferences on interhelical interaction patterns based upon extensive simulations on a single helix inserted, *a posteriori*, in a low resolved static model of the receptor risk being too speculative. In summary, we think that extensive simulations on isolated receptor domains can provide useful information, unless the information involves receptor domains not included in the simulation.

3.5. Comparative Modeling of GPCRs Using the Rhodopsin Structure as a Template

In the years that preceded the release of the first crystal structure of rhodopsin,[13] both comparative and ab initio modeling approaches strongly relied on information from low resolution electron density maps of rhodopsin[9,12] and on the results of sequence analyses predicting that the members of family A share the same architecture of the seven helices.[10,11] Paradoxically, as soon as the structural information on rhodopsin reached atomic detail, several papers were published raising the question whether and to which extent the other members of the rhodopsin family share with rhodopsin the structure and the mechanism of functioning.[75,342,352−355] This repeated and not yet definitely addressed question originates from the low sequence similarities among the members of family A. The length of GPCRs from family A, in fact, may vary between less than 300 and more than 900 amino acid residues, with the majority of receptors having a length around 310−470 residues.[75]

Sequence analysis suggested that family A GPCRs could share the same arrangement of the seven helices, also due to the presence of a few but significantly conserved residues and motifs in each of the seven helices.[10,11,75] The amino acids, which resulted in having 80−100% conservation from a recent study on the alignment of 270 members of family A, are N1.50, L2.46, D2.50, C3.25, E/D3.49, R3.50, W4.50, F6.44, W/F6.48, P6.50, P7.50, and Y7.53 (Figure 1).[75] In detail, the analyses of the frequencies of individual amino acids in particular positions of the seven helices revealed that H1, in addition to the invariant asparagine, N1.50, holds G1.49, L1.52, and V1.53 that are highly conserved (68, 60, and 66%, respectively). In several cases, basic residues are found at the beginning of H1 and at the end of H2, possibly stabilizing the interaction of these helices with phospholipids. Frequently, a proline residue is at the beginning of H2, likely without perturbing the helical structure (Figure 1). H2, in addition to the almost invariant D2.50, has six conserved aromatic/hydrophobic residues toward the cytoplasmic surface, while residues toward the extracellular domain are more divergent.[75] H3 is characterized by the presence of highly conserved amino acids at its extracellular and intracellular ends, i.e.,

C3.25 and the E/DRY motif, respectively. This cysteine in the rhodopsin structure forms a disulfide bond with cysteine residue 187 of E2, and its very high conservation (90%) could be suggestive of potential structural similarities in E2 within family A. Interestingly, with few exceptions, the residue immediately following C3.25 corresponds to the ligand type for the receptor. If the residue is basic, the ligand for the receptor is most likely a peptide (34% K, 19% R). If it is acidic (D, E), it is a biogenic amine.[75] The highly conserved residues at the cytosolic extension of H3 are E/D3.49 and R3.50 (86% and 96% conservation, respectively) of the E/DRY motif (Figures 1 and 2). H4 contains invariant aromatic residues, mostly tryptophan. H5 contains two aromatic residues, F5.47 and Y5.58, which are common in GPCRs (70 and 77%, respectively). As for H6, the highly conserved aromatic amino acids F6.44 and W5.48, together with P6.50, constitute the FxxCWxP functionally important motif (Figure 1). Interestingly, the glutamate/aspartate at position 6.30, suggested to be involved in stabilizing the inactive state of rhodopsin and other homologous GPCRs,[61,62,64−66,68−70] has a low conservation (32%), being mainly a feature of rhodopsin, the amine, and some non-peptide subfamilies. Three highly conserved amino acids in H7, i.e., N7.49, P7.50, and Y7.53, form the functionally important NPxxY motif, where x's are mostly hydrophobic L, V, and I residues (Figure 1).

Mapping the structural information from biochemical experiments on GPCRs of family A (reviewed in refs 96 and 342) into the rhodopsin structure led to the predictions that several of the highly unusual structural features of rhodopsin are also present in amine GPCRs, despite the absence of amino acids that might have been thought to be critical to the adoption of these features.[342] These conclusions were based on the hypothesis that different amino acids or alternate microdomains can support similar deviations from regular α-helical structure, thereby resulting in similar tertiary structure. This phenomenon has been defined as "structural mimicry", in other words a mechanism by which a common ancestor could diverge sufficiently to develop the selectivity necessary to interact with diverse signals, while still maintaining a similar overall fold.[342]

The availability of the high resolution structure of rhodopsin made it also possible to extend the prediction of structural commonalties between rhodopsin and the homologous GPCRs to the intracellular and extracellular domains.[75] Combining the information from sequence analysis with the structure of bovine rhodopsin, the beginning and ending of the N- and C-termini and of the three extracellular and intracellular loops were predicted for all the 270 aligned receptors. The study predicted that the amino acid length of the N-terminal region is highly variable, containing from as little as four to as many as over 50 amino acid residues in length.[75] Among the three extracellular loops, E1 has the most consistent loop size; in fact, 144 GPCRs have the same number of amino acids in E1 as rhodopsin (i.e. 5 amino acids, Figure 1), whereas, in the remaining members, this loop ranges from only 3 amino acids to as many as 18 amino acids. The other two extracellular loops (E2 and E3) have more variable sizes. Similarly to the extracellular loops, the intracellular loops can vary in size, with the most conserved loops being I1 and I2. There are 198 GPCRs that have the same number of amino acids in I1 as rhodopsin (i.e. 6 amino acids, Figure 1), whereas the remaining members of family A have either five or seven amino acids. For I2, over 150 receptors have a loop size of 10−12 amino acids (i.e. 11 amino acids in rhodopsin, Figure 1). I3 and the C-terminus have the highest variations in amino acid lengths among all the considered GPCRs.[75]

Collectively, the results of in vitro experiments aimed at structurally probing GPCRs of family A[96,342] and those of sequence analyses[11,75] suggest that comparative modeling of the seven-helix bundle of GPCRs using rhodopsin structure as a template is likely to produce reliable results, and this has been, indeed, the assumption of high throughput computational modeling experiments, very recently reported, which targeted the TM domains of 235 GPCRs.[356]

It is, however, possible that the beginning and ending of a TM helix may vary among the members of different subfamilies. Therefore, in those cases in which a given helix in the target receptor is expected to be longer than the corresponding one in the rhodopsin structure, extra α-helix restraints could be required while modeling the target receptor. Predictions of the beginning and ending of each helix are difficult but essential to characterize the nonstructured parts of the individual intracellular and extracellular domains and, hence, to estimate the potential structural similarity in such domains between the target receptor and the rhodopsin structure. Difficulties are, indeed, expected to reside in modeling such hydrophilic domains, which vary both in amino acid composition and in length. In this respect, we have estimated whether each of the extracellular and intracellular domains in 163 selected GPCRs of family A can be modeled using the corresponding domain in the rhodopsin structure as a template. We have considered the human sequences of 163 members of family A, including 35 amine, 79 peptide, 3 hormone, 11 olfactory, 8 prostanoid, 9 nucleotide, 2 cannabinoid, 1 platelet activating factor, 1 gonadotropin-releasing hormone, 3 thyrotropin-releasing hormone and secretagogue, 4 melatonin, 5 lysosphingolipid, and 2 leucotriene B4 receptors. The evaluation has been based on sequence alignments between rhodopsin and one or more members of a given subfamily. For each receptor, a pairwise sequence alignment with rhodopsin has finally been obtained, which has been manually adjusted to incorporate additional information, including the results of multiple sequence alignments. As stated above, the intracellular and extracellular domains are characterized by lack of sequence similarity among the members of the rhodopsin family and rhodopsin. However, initial models of such domains, based upon the rhodopsin structure, could be achieved in those cases in which there is a significant similarity in length, either in the presence or absence of conserved amino acids. Since our reference program for com-

parative modeling is MODELLER, which is based on the satisfaction of stereochemical restraints,[91] we have also estimated those cases in which extra α-helical restraints would be needed to impose an α-helical conformation to insertions in the target receptor proximal to the N- or C-terms of each TM helix. In these cases, deletions in the template structure would be needed as well, at the junctions between the tail/loop and the helix. As for the extracellular domains, the N-term, in a few cases, including the glycoprotein hormone, the endothelin, and the proteinase-activated-like subfamilies, in which it is significantly longer than that of rhodopsin, structural restraints are unlikely to be transferred from the rhodopsin structure to the target receptor. As for the glycoprotein hormone subfamily, the N-term is predicted and demonstrated to fold into a regular structure;[15,70,357−361] thus, it must be modeled separately and then incorporated into an ad hoc modified rhodopsin template to achieve the receptor model. E1 can almost always be modeled on the basis of the rhodopsin structure; exceptions include the glycoprotein hormone receptor subfamily, whose E1 is significantly longer than that of rhodopsin. In these cases, E1 modeling should be attempted following energy-based ab initio approaches, such as that implemented in the MODELLER program.[69,362] An energy-based computational protocol for ab initio modeling of GPCR loops has been described in ref 363. Alternatively, attempts to model the exceeding sequence as an extracellular extension of H2 and/or H3 should be done, by adding extra-α-helical restraints during comparative modeling, following deletion of one or two amino acids at the helix/loops junction in the template structure. As for E2, in all the rhodopsin structures released so far, the first 18 amino acids in this loop (i.e. the segment 174−191) form a β-hairpin whose second strand begins with a conserved cysteine, C187, which is involved in a disulfide bridge with the highly conserved cysteine at position 3.25 (Figures 1 and 2).[13,51−54] In the majority of the analyzed sequences, matches with the rhodopsin's β-hairpin can be achieved with a few insertions and/or deletions, allowing C187 to align with a cysteine residue in the sequence of the target GPCR. For some of the amine receptors, such matches would require deletions in the nonstructured C-terminal part of the rhodopsin loop. Therefore, comparative modeling of E2 can be attempted for the majority of family A GPCRs. As for E3, it can be comparatively modeled for the amine subfamily, whereas, for the PEPTIDE receptors, in which such a loop is predicted to be longer than that of rhodopsin, it is frequently difficult to achieve on the basis of the rhodopsin structure. In this respect, the rhodopsin structure can be used as a template only if the required insertions can be modeled as one/two-turn extensions of H6 and/or H7 and not as part of the loop. In those cases, extensions of H6 and/or H7 could be eventually achieved during comparative modeling by applying extra α-helical restraints to the target amino acid segment. This would require deletion of one or two amino acids at the helix/loops junction in the template structure.

As for the intracellular domains, I1 and I2 can almost always be modeled on the basis of the rhodopsin structure. In contrast, comparative modeling of I3 based upon the rhodopsin structure is not possible for many GPCRs, in particular for the amine subfamily. In fact, in this case, I3 is significantly longer as compared to that of rhodopsin, and it is also predicted to hold elements of secondary structure and, hence, a fold. Structure predictions of these huge loops could be attempted by alternative approaches such as fold recognition, leading to chimeric rhodopsin templates in which rhodopsin's I3 has been replaced by the corresponding loop of the target receptor in the predicted fold.[294,364] In contrast to the case of the amine GPCR subfamily, for the remaining subfamilies, there may occur two situations: (a) the loop can be modeled on the basis of that of rhodopsin, eventually introducing one or two α-helix turns at the N-term of H6, as reported for subtypes 1 and 2 of the melanin-concentrating hormone receptor (MCHR1 and MCHR2);[365] (b) I3 is significantly shorter than that of rhodopsin, and its modeling should be attempted following energy-based ab initio approaches, such as that implemented in the MODELLER program[69,362] (see ref 363 as an example of alternative computational protocols), or a database loop search; and (c) I3 is significantly longer than that of rhodopsin, and it should be built separately, as seen above for the amine subfamily. As for case (a), it should be considered that profound disagreement is found between the $P4_1$ (1U19)[54] and $P3_1$ (1GZM)[53] rhodopsin structures concerning the ending of H5, the beginning of H6, and the structure of their interconnecting loop. Such mismatch would require probing both the two different rhodopsin structures as templates. As for case (c), this situation occurs only in a minority of cases, including the vasopressin-like subfamily.[60,71]

For almost all GPCRs, the C-tail cannot be entirely modeled on the basis of the rhodopsin structure. In general, the N-terminal half of the C-tail can often be achieved, differently from the C-terminal half. In a few cases, the C-tail is predicted to fold in a regular structure and it should be modeled on the basis of alternative templates and then merged into an ad hoc modified rhodopsin template.[364]

In summary, comparative modeling of the entire sequences of GPCRs, based upon the rhodopsin structure, is largely far from being trivial. We have estimated that full automated modeling of an entire GPCR sequence is never feasible, but rather human intervention is always necessary for modifying ad hoc the rhodopsin template, for introducing extra restraints or for ab initio modeling of selected domains. Moreover, the models produced by the MODELLER program should be considered as the starting points of further energy refinements and MD simulations. Average structures produced by such calculations show RMSDs (in the TM domains) from the rhodopsin structure more reliable than those of the models prior to calculations.[62,366] In fact, the structure deviation, in terms of main-chain RMSDs, for pairs of protein which, such as rhodopsin and the homologous GPCRs, are ~22−24% identical, is expected to be

close to 2 Å instead of 0 Å, as is the case of the initial receptor model produced by MODELLER.[367]

Despite the difficulties in using the crystal structure of rhodopsin as a template for modeling the homologous GPCRs, we think that, at the moment, comparative modeling remains preferable to the most effective ab initio approaches, and it has, indeed, been widely used since the year 2000.[62,64,65,67−70,72,139,141,347,356,361,363−365,368−441]

Sequence-based methods aimed at predicting residue−residue contacts such as those based on covariation analysis have been proposed, which could eventually improve comparative models of GPCRs and detect functional sites such as those intended for ligand binding.[442]

4. Computational Experiments on Family A GPCRs

4.1. Thermodynamic Models of GPCR Function

Until 1995, GPCR ligands were classified as antagonists, and full or partial agonists, depending on whether they produce null, full, or partial stimulus upon binding to the receptor, respectively.[443] However, recent advances in receptor theory and experimental technology for ligand screening led to the discovery of many additional types of receptor ligands.[443] Some of these are (a) inverse agonists, which inhibit constitutive activity of GPCR systems (i.e. agonist-independent spontaneous activity that emanates from the system itself); (b) allosteric agonists, which function as agonists by interacting with a site distinct from that of the endogenous agonist (usually a non-peptide ligand for a peptide receptor); (c) allosteric modulators (antagonists), which block receptor function but do not necessarily interfere with ligand−receptor interaction (receptor occupancy); and (d) allosteric enhancers, which potentiate the agonist effect on the receptor (reviewed in ref 443).

Concepts regarding the mechanisms by which drugs activate receptors to produce biological and physiological response have progressed beyond considering the receptor as a simple on−off switch (reviewed in refs 443 and 444). Current evidence suggests that the idea that agonists produce only varying degrees of receptor activation is obsolete and must be reconciled with data showing that agonist efficacy, i.e., the ability of a molecule to produce some observable physiological response, has texture as well as magnitude. Thus, GPCR ligands can block a system's constitutive response (inverse agonists), behave as positive and inverse agonists on the same receptor (protean agonists), and differ in the stimulus pattern they produce in physiological systems (ligand-selective agonists). The underlying general mechanism for this seemingly diverse array of activities is the same, namely, ligand-dependent changes in the conformational states of the receptor (reviewed in refs 443 and 444).

GPCRs are allosteric proteins designed by nature to transmit extracellular signals to the intracellular milieu of the cell. They accomplish this task by changing their interaction with large intracellular proteins (G proteins) upon binding of extracellular ligands. The most general model that formulates this ligand-mediated process is the ternary complex model (TCM) (Figure 10).[445,446] According to this view, the interactions of ligand (H), G protein (G), and receptor (R) can be described by the equilibrium scheme shown in Figure 10, where K and M are the unconditional affinities for the formation of ligand−receptor and G protein−receptor complexes, respectively, and α represents the reciprocal effect that ligand and G protein impart on the binding of each other when they form the ternary complex (HRG) (Figure 10). It measures the standard free energy transferred from the binding of ligand to the binding of G protein (or vice versa). The equivalence between ligand efficacy and α is obvious: α greater than, equal to, or less than 1 means that the ligand enhances, leaves unchanged, or reduces the tendency of R to bind G; thus, it has positive (agonist), null (neutral antagonist), or negative (inverse agonist) efficacy, respectively. Efficacy defined by α depends on H and R but also on G. Therefore, the scheme is extendible to describe the interaction of R with G_1, G_2, ..., G_n, which yields α_1, α_2, ..., α_n for the same ligand. This means that a given ligand H interacting with the same receptor R will have different efficacies when R binds to different types of G proteins. Note also that the receptor and G protein influence each other and they can interact even in the absence of agonist ligands. The system can be forced to produce a response by changing the stoichiometry of the reactants, namely, R and G. Thus, the constitutive activity (as defined by elevated levels of [RG]) can be increased by raising the receptor concentration or by increasing the concentration of G protein. Another way in which constitutive activity can be produced is through alteration of M. Although this idea summarizes the entire phenomenology mentioned above, it addresses the question of ligand-induced conformational change only implicitly; changes in the conformational state of the receptor are hidden in the parameter α, and the model does not specify their nature but only their consequences on the interactions of H and G with R.

The extended ternary complex (ETC)[447] or cubic ternary complex (CTM) models,[448−450] which are based on the same thermodynamic principles as the TCM, consider the active (R*) and inactive (R) conformational states of the receptor explicitly (Figure 10). Those conformations of the receptor that hold an arrangement in the cytosolic domains that can activate the G protein are referred to as the "active state", and those that do not are referred to as the "inactive state". Accordingly, the two conformations R and R* exist in equilibrium with each other, which is governed by an "allosteric constant" (denoted as J and defined as [R*]/[R]). Under normal circumstances, J is a unique molecular constant for a given receptor (i.e. the energy barrier to formation of spontaneous active states for some receptors is lower than it is for others), but experimental methods such as the removal of sodium ions[451,452] or point mutation (reviewed in refs 29, 45, 48, and 49) can affect J and make receptors more constitutively active. In principle, agonists can induce response by causing enrichment of the active state by selectively binding to

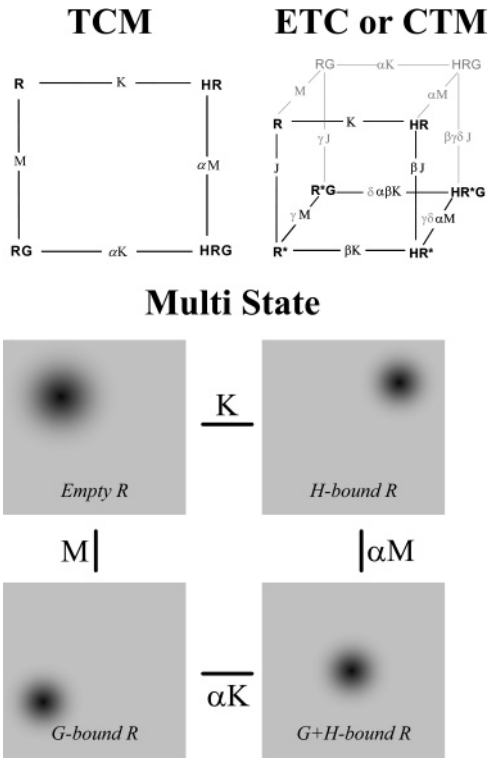


Figure 10. Representation of thermodynamic models of GPCR interactions. In the ternary complex model (TCM), the receptor is assumed to hold four ligation states: empty, G protein (G)-bound, ligand (H)-bound, and bound to both ligand and G protein. Two unconditional equilibrium constants (*K*, *M*) and one equilibrium coupling constant (α) describe the entire equilibrium between the possible ligation states. α describes the reciprocal effect that ligand and G protein impart on each other's binding. In the extended (or cubic) ternary complex model (ETC or CTM), the active and inactive states of the receptor (R* and R) are explicitly included in the equilibrium scheme. An additional equilibrium constant *J* governs the unconditional isomerization equilibrium of the receptor (R−R*). As in the case of the ternary complex scheme, greek letters represent the coupling constants between relevant equilibria. Note that δ is a "second order" constant stating the reciprocity of the three chemical events occurring in the receptor (two ligand binding processes and the isomerization). The dark paths and the associated species indicate how the model simplifies if we assume that there is no binding of G to the R form of the receptor (which represents ETC). The picture of the multistate model is a schematic representation of the model, where the receptor is assumed to have an unlimited number of states in equilibrium with each other and representing the entire conformational space of the receptor (schematized as gray squares in the picture). A macroscopic state of the receptor is defined as a probability distribution over this conformational space. This probability distribution, schematized as a fuzzy spot in the conformational space, is subjected to change depending on the ligation state of the receptor. Again, the receptor has four ligation states and the equilibrium between them is governed by three macroscopic equilibrium constants (*K*, *M*, α), as in the TCM. In this case, however, the macroscopic constants are determined by the relationship between the probability distributions that the receptor assumes at different lagation states. For example, the allosteric constant α results from the discrepancy between the probability distributions that arise when the receptor is empty, G protein-bound, or ligand-bound.

R*. Under these circumstances, efficacy is a matter of the quantity of the active state produced by the agonist. More generally, selective affinity of ligands for various receptor conformations will change the overall distribution of the species in the GPCR ensemble and, thus, either induce or inhibit response. This is the basic mechanism of ligand efficacy and the basis for the molecular nature of inverse, protean, and ligand-selective agonism according to the allosteric models such as the ETC and CTM models (Figure 10; reviewed in refs 443 and 444). There is no theoretical constraint on the number of active states of the receptor. Even though the ETC and CTC models have both been referred to as "two-state" models, this is a misnomer in that there is the capability within both to be multistate models (reviewed in refs 443 and 444).

In light of the ever-increasing evidence of GPCR dimerization/oligomerization, thermodynamics models have been recently proposed to account for either the lack of cooperative ligand binding or the presence of negative or positive cooperative binding generated by the crosstalk between protomers within a dimer/oligomer (reviewed in refs 39, 453, and 454).

It is now clear that efficacy cannot be considered as a strictly receptor-based ligand property, but rather a function of the entire GPCR system, involving both the receptor and its cellular environment (reviewed in refs 444 and 455). GPCRs have an extremely wide spectrum of behaviors, which include pleiotropic coupling (i.e. the ability to activate more than one type of G protein), dimerization, oligomerization, internalization, desensitization, and interaction with other numerous cellular coupling proteins (reviewed in refs 444 and 455). Studies of these behaviors indicate that some molecules can induce some of these effects but not others and that, occasionally, a drug might not produce an overt physiological response but might, nevertheless, change receptor behavior and, hence, have efficacy. So, there are different "efficacies" for different behaviors beyond the activation of G proteins. The question is whether a general molecular mechanism can account for all of these expressions of efficacy. One approach to describing the numerous other manifestations of efficacy that do not necessarily produce a cellular response is to consider receptor proteins in terms of ensemble theory.[444,456,457] This theory is based on the fact that receptors exist as collections of ensembles of numerous conformations. In this respect, the probabilistic multistate model that assumes an unlimited number of conformational states for the receptor predicts that the relationships between conformational states and protein function are stochastic rather than deterministic as the allosteric models imply (i.e. the multistate model in Figure 10). The latter statement amounts to the idea that the active state of the receptor cannot be attributed to individual conformational states, but rather to an ensemble of the states in the conformational space of the receptor. Each ensemble is associated with a given function such as G protein activation, interaction with cellular proteins, dimerization/oligomerization, etc. At any given instant, the various conformations in an ensemble can be represented as a Gaussian distribution, and there could be intersect-

ing ensembles for a range of GPCR behaviors (see ref 457 for a more detailed discussion of the idea).

Ligands influence receptor behavior through selective affinity for the various conformations in the receptor reference ensemble. Although some ligands will bind preferentially to some receptor conformations over others, the weighted average affinity that a ligand has for a receptor ensemble is known as the "macroaffinity" of the ligand for the receptor, that is the concentration of ligand that is bound to 50% of the receptors at any one instant.[444] Even if common in vitro experiments are inadequate to address the question whether ligands select or induce various receptor conformations, the restructuring of a collection of receptor conformations through selective binding affinities emphasizes the idea that affinity is not a passive phenomenon, whereby a ligand simply binds to a protein but does not change it.[444] Both ligand and receptor are dynamics entities. Their selective interaction is governed by dynamic cooperative events encoded in their molecular structure and described by statistical thermodynamics. The idea that ligands, whatever their efficacy is, may alter receptor conformational states appears to be supported by fluorescence spectroscopy experiments on the β_2-AR[458−460] and on the human leukotriene B_4 receptor.[461] Fluorescence spectroscopy studies aimed at monitoring catecholamine-induced conformational changes in purified β_2-AR provided evidence for a multistep process of agonist binding where contacts between the receptor and key moieties on the agonist stabilize a succession of conformational states with distinct cellular functions.[462,463]

4.2. Computational Modeling of Mutation-Induced Active States

We have pioneered the use of computational modeling to infer the structural features of the agonist-independent active states of different GPCRs.[45,57,58,60,62,64,65,69,72,302−305,366,464−466]

The first studies were done on an ab initio model of the α_{1b}-AR holding all the domains but the N-term, I3, and the C-tail.[57,58] The idea was that comparing the structure/dynamics features of the wild type α_{1b}-AR with those of receptor mutants, characterized by different degrees of constitutive activity or inability to activate the G protein, would help identify some of the molecular changes correlated with the transition from R to R*, independently of the presence of agonist.[57,58]

The computational approach consisted of generating a large number of average configurations following MD simulations of the wild type and the mutated forms of a common input structure. A comparative analysis of such average arrangements was then carried out, focusing on a few but significant structural features, which were shared in common by the majority of the mutant structures with similar functionality and which made the difference between active and nonactive forms. This strategy is thought to overcome, at least in part, the drawbacks related to the low resolution of the computational models and the approximations and simplifications in the computational setup.

Simulations were carried out in vacuo, using a distance-dependent dielectric and intrahelix distance restraints between the backbone oxygen and nitrogen atoms of all amino acids in the helix, except for prolines. The application of these intrahelical distance restraints, developed in early computational simulations of the packing arrangements of seven-helix bundles,[293] was instrumental in (a) reducing the system degrees of freedom, (b) inferring the structure/dynamics role of prolines, and (c) letting the helices move as rigid bodies, consistent with the experimental evidences on rhodopsin activation.[130] A fundamental step in this approach is the choice of the input structure and of the computational conditions that produced divergent average behaviors for the active and nonactive receptor forms, consistent with the experimental information available thus far.

Computational modeling of all the 19 possible substitutions for A293(6.34) in the α_{1b}-AR, characterized by variable levels of constitutive activity,[467] highlighted the role of the E/DRY conserved motif in regulating the agonist-independent receptor transition from the inactive to the active receptor states.[57] In the inactive states, represented by the wild type and the nonactive A293(6.34) mutants, R143(3.50), of the conserved E/DRY motif, was found engaged in H-bonding and van der Waals attractive interactions with the amino acids of a highly conserved polar pocket in the cytosolic halves of H2 and H7. These studies were, therefore, suggestive of an H-bonding network of conserved amino acids as a constitutive structural feature of the nonactive receptor states, consistent with the results of previous computations on more simplified models of the same receptor.[295,296] The release of the H-bonding network involving R143(3.50) and the cluster of conserved polar amino acids was found to be the common feature to all the constitutively active mutants (CAMs) at A293(6.34).[57] Computer simulations suggested also a potential mechanism of regulation of GPCR function via changes in the protonation state of the aspartate of the E/DRY motif, with the protonated form being associated with the active states.[57] This hypothesis was inferred from the observation that protonation of D142(3.49) conferred to the α_{1b}-AR the same average structural features shown by the highly active mutants of A293(6.34). The idea that protonation of D3.49 was the perturbation rather than mutation able to trigger agonist-independent active states of wild type α_{1b}-AR was also supported by the results of in silico mutagenesis showing that replacements of D142(3.49) with neutral amino acids would give the structural features of the active states to the receptor.[57,58] Predictions of computational modeling were validated by the experimental findings that the irreversible reprotonation, following mutation, of D142(3.49) led to constitutively active forms of the α_{1b}-AR.[57,58] In this respect, the D142(3.49)A mutant was the first example in the literature of computational design of a constitutively active GPCR mutant.[57] The hypothesis that reprotonation of the aspartate/glutamate of the E/DRY motif could favor the active states was successively strengthened by in vitro and in silico experiments on the GnRH and

μ opioid receptors.[59,63] In line with these evidences and consistent with the knowledge that the homologous glutamate in rhodopsin is involved in proton uptake from the cytosol (reviewed in ref 98), very recent pK_a calculations on the crystal structure of rhodopsin proved the susceptibility of E134(3.49) to undergo significant pK_a shifts and, hence, change its prototropic state depending on its environment.[468]

Upgrading and complicating the α_{1b}-AR model, to incorporate advances in structural determination of rhodopsin, strengthened the hypothesis that the E/DRY motif might play a role in maintaining the inactive state of the receptor, while introducing novel structural hallmarks of the active and nonactive states, such as the degree of solvent accessibility of selected cytosolic domains, including I2, I3, and the cytosolic extensions of H3 and H6. These domains, in fact, underwent solvent exposure on going from the nonactive to the active receptor forms and were, hence, suggested to hold potential recognition points for the G protein.[294,302,303] These results found consistency with the results of the automatic docking simulation of the wild type α_{1b}-AR and two CAMs (i.e. D142(3.49)A and A293(6.34)E) with different heterotrimeric G proteins.[302,303]

The same computational approach was used to build an agonist-independent activation model of the human LHR,[304] which is particularly susceptible to spontaneous pathogenic activating and loss-of-function mutations.[48,49,466] The studies of Lin and coworkers and by Fanelli represented the first attempts reported in the literature aimed at inferring the molecular determinants of mutation-induced activation of the LHR.[251,304] The LHR model by Lin et al., limited to the TM domains, which were packed by fitting the peaks in the 9 Å electron density map of bovine rhodopsin, suggested that single activating mutations perturb the specific interactions of H6 with H5 and H7, either by disrupting the hydrophobic packing between H5 and H6 or by weakening the H-bonds between H6 and H7.[251] In the study by Fanelli, MD simulations were carried out on the wild type and on the majority of the spontaneous active and nonactive LHR mutants known thus far.[304] Similarly to the case of the α_{1b}-AR, the mutation-induced active states of the LHR shared the release of the charge-reinforced H-bonding interactions involving the conserved arginine of the E/DRY/W motif in the wild type form and the opening of a crevice between I2 and I3.[304] The latter effect was found to be properly accounted for by the solvent accessible surface area of W465(3.51), of the E/DRY/W motif (i.e. $SAS_{W3.51}$). This index was indeed found to be close to 0.0 Å^2 in the nonactive forms and above 32.0 Å^2 in the active ones. The $SAS_{W3.51}$ index, together with the distance between D405(2.50) and R464(3.50), was used for predicting the functional behavior (i.e. active or nonactive) of 48 novel LHR mutants, constituting the first attempt in the literature of in silico functional screening of GPCR mutants.[304]

Computational modeling of both mutation- and agonist-induced activation of the human OTR also suggested the weakening of the interaction pattern of the E/DRY/C arginine and the opening of a cytosolic crevice as features of the active receptor states.[60]

As soon as the first crystal structure of rhodopsin was released, computational modeling of agonist-independent activation was carried out on novel models of the α_{1b}-AR, LHR, and OTR, which were achieved by comparative modeling, using the rhodopsin structure as a template.[62,64,65,69,71,72,366,464,465]

As for the α_{1b}-AR, MD simulations of more than 100 single and double mutants of the receptor, done in parallel with in vitro experiments, made possible the definition of virtual structures, representatives of the agonist-independent active and nonactive states, both for the ab initio and homology models.[62,64] The virtual structure of the ground state was obtained by making an average over the average structures of the wild type and of the receptor mutants showing a wild type-like phenotype, whereas the virtual structure representative of the active state was obtained by making an average over the average structures of the most active α_{1b}-AR mutants.[62,64] Critical comparison between the "ground state" models achieved by ab initio and comparative modeling highlighted overall similarities in the length and arrangement of the TM helices.[62] Both the ab initio and, perhaps better, the homology models predicted that the majority of amino acids susceptible to activating mutations [D142(3.49), Y144(3.51), V137(3.44), R143(3.50), E289(6.30), and A293(6.34)] belong or are close to the interface between the cytosolic extensions of H3 and H6. A common feature to all these residues is that they contribute to the environment of R143(3.50) of the E/DRY motif on H3. Critical comparisons also highlighted differences in the amino acids which participate in the H3/H6 interface and, thus, constitute the environment of R143(3.50) of the DRY motif. In the homology model, R143(3.50) makes a charge reinforced H-bond with both the adjacent D142(3.49) and E289(6.30) and, hence, D142(3.49) is suggested to exert a constraint on the motion of R143(3.50) through a charge-reinforced H-bond rather than through long-range electrostatic interactions as suggested by the ab initio model.[62] A feature of the inactive state within the homology model, not shared with the previous ab initio model because of slightly different rotation of H6, is the charge-reinforced H-bond between the E/DRY arginine and E289(6.30), predicted to be an additional constraint to the R143(3.50) motion and a link between the cytosolic extensions of H3 and H6. For both the ab initio and the comparative models, the virtual structures representative of the active state are characterized by the weakening of the charge-reinforced H-bonds involving R143(3.50) in the inactive state.[62,64] The hypothesis, raised by the homology model, that both D142(3.49) and E289(6.30) contribute to stabilize the inactive state of the α_{1b}-AR found support in the results of in vitro experiments, which showed that replacing either D142(3.49) or E289(6.30) with a neutral or a cationic amino acid that would break the salt bridge with R143(3.50) leads to agonist-independent receptor activation.[57,58,62,64] An equivalent role of the E/DRY arginine, as important switch of mutation-induced receptor activation, has been inferred

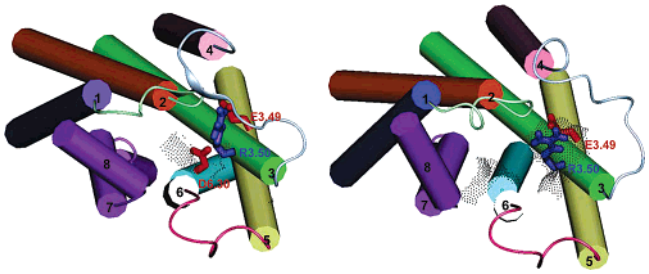


Figure 11. Average minimized structures of the wild type LHR (left) and of the D564(6.30)G constitutively active mutant (right).[65,69] The receptors are seen from the intracellular side in a direction perpendicular to the membrane surface. The seven helices are represented by cylinders whereas the three intracellular loops are represented by thin ribbons. The extracellular loops are not shown. H1, H2, H3, H4, H5, H6, and H7 are respectively colored in blue, orange, green, pink, yellow, light blue, and violet, whereas I1, I2, and I3 are respectively colored in light green, gray, and purple. The amino acid stretch corresponding to rhodopsin's H8 is colored in violet as well. The side chains of E463(3.49), R464(3.50), and D564(6.30) are represented by sticks and colored according to their polarities. The composite solvent accessible surfaces computed over the amino acids R464(3.50), T467(3.53), I468(3.54), and K563(6.29) are also shown, represented by gray dots.

by very recent computations on OTR.[71] The role of the charge-reinforced H-bond between R3.50 and E6.30 in maintaining the inactive states of GPCRs has been overemphasized by a number of computational and in vitro experiments.[61,62,64−70,72] However, the significantly lower conservation of the glutamate/aspartate at position 6.30 (i.e. 32%) compared to the glutamate/aspartate at 3.49 (i.e. 86%)[75] makes its potential role valid only for a few GPCRs. Computations on the μ opioid receptor,[469] on the MCHRs,[365] and on OTR[71] suggested that, in the absence of a conserved glutamate/aspartate at position 6.30, other amino acids in the cytosolic extension of H6 may contribute to create H-bonding interactions with the E/DRY arginine. These amino acids include T6.34, for the μ opioid receptor,[469] T6.30, for MCHR1,[365] and T6.33, for OTR.[71]

As for the LHR, simulations on the novel homology model of most of the spontaneous pathogenic activating and inactivating mutations discovered so far strengthened the role of the second and third intracellular loops and of the cytosolic extensions of H3 and H6 as the targets of the structural modifications induced by the different spontaneous activating mutations.[65,69,72,366,464,465] Indeed, the average minimized structures of the constitutively active mutants share the increase in solvent accessibility of selected amino acids at the cytosolic interface between H3 and H6. This effect, already observed in the previous ab initio model, is properly described by the solvent accessible surface area (SAS) computed over selected amino acids, including the E/DRY arginine. Such a computational index proved to be an effective hallmark of the functional receptor state, being lower than 50 Å^2 in the inactive forms and higher than 50 Å^2 in the active ones (Figure 11).[65,69,366] It was, therefore, successfully challenged in its ability to predict the functional behavior (i.e. active and nonactive) of a large number of artificial mutants of the LHR.[65,69,72] Another feature of the most active LHR mutants was the weakening of either one or both the charge-reinforced H-bonding interactions found in the wild type between R464(3.50) of the E/DRY/W motif and both E463(3.49) and D564(6.30) (Figure 11). The correlation between weakening of such interactions and an increase in the basal receptor activity has been proposed also for the thyrotropin receptor (TSHR).[70] However, according to our computational models, the interaction pattern of the E/DRY/W motif is a less effective hallmark of the functional state of LHR than the SAS index.[65,69,72]

In summary, the extensive in vitro and computational experiments on substantially different GPCRs, such as the α_{1b}-AR, LHR, and OTR, allowed us to infer hypotheses on the requirements for a GPCR site to be susceptible to activating mutations, highlighting commonalties and differences among the distinct mutation sites. One inference from computations was that, despite the topological and structural differences between them, the activating mutation sites are structurally connected with peculiar portions of the cytosolic domains, including the E/DRY motif. Such a highly conserved stretch of amino acids is particularly susceptible to undergo structural modifications in response to activating mutations. In fact, activating mutations tend to weaken the ground state interactions of R143(3.50) and increase the solvent accessibility of selected amino acids at the cytosolic extensions of H3 and H6. Calculations show that this structural effect is mediated by highly conserved polar amino acids in the seven-helix bundle. Whether the main role of the E/DRY arginine is to maintain the inactive state of the receptor or to recognize the G protein is not clearly understood and may depend on the receptor system (critically analyzed in refs 73 and 470). For some GPCRs, such as the α_{1b}-AR,[305] OTR,[60] and the novel receptor ORF74-EHV2,[74] in which the ad hoc engineered or spontaneous absence of the E/DRY arginine is associated with constitutive activity that is abolished in the presence of the DR pair, the main role of the conserved arginine might be to maintain the inactive state of the receptor and drive receptor isomerization into different functional states (reviewed also in ref 73).

Critical comparisons of the interpretative and predictive abilities of the previous ab initio and novel homology models of the α_{1b}-AR, LHR, and OTR seem to suggest that the homology models can be considered advancements over the previous ab initio models, which have, however, been extremely useful for developing the computational approach. However, these conclusions are essentially based on low resolution experimental data, which are not properly adequate for validating atomistic models. Therefore, perhaps it would be better to say that homology models are expected to be more reliable than the ab initio ones (at least in the portions where significant sequence similarity exists between template and target proteins), on the basis of the results of critical assessments of protein structure prediction methods.[471]

The major problems with in vacuo simulations, such as those that we have carried out to study mutation-induced GPCR activation, are associated

with the rough evaluation of the electrostatic energies, including the dielectric constants that represent the effect of the protein environment. One consequence of the lack of the proper screening effect of the solvent is an overall overestimation of the water-exposed salt bridges and charge-reinforced H-bonding interactions. However, active and nonactive receptor forms have been simulated by using the same computational setup. Therefore, the selective breakage or formation of electrostatic interactions with changes in the functional receptor state may be indicative of a significant effect. Along the same line, the inferences from in vacuo simulations were recently corroborated by MD simulations on mutation- and ligand-induced activation of OTR, by using the IMM1 implicit membrane/water model[472] without any intrahelical distance restraints. Convergence was, indeed, found between in vacuo and IMM1 computations, concerning the structural peculiarities of the inactive and active forms of the receptor.[71] This may be due to the fact that the structural hallmarks of the inactive and active states preferentially involve the helix bundle rather than the loops, which have lower resolution and are more sensitive to differences in the electrostatic models than the seven helices. In this respect, consistency between the two different electrostatic models has been found although the implicit representation of the protein environment provides a better screening effect than in vacuo simulation.

The strength of our computational approach is that it relies on an extensive comparative analysis, which is aimed at inferring similarities/differences within the same approximations. One of the main inferences of this comparative approach is that a receptor exists in different active states that, however, share a few key structural features, which presumably determine the macroscopic functional receptor state. This inference is in line with the ensemble theory.[444,456,457]

4.3. Computational Modeling of Ligand–Receptor Interactions

Heterogeneity in the functional properties of GPCR ligands is nothing compared to the heterogeneity in their structural features. Indeed they comprise small molecules, such as ions, organic compounds, amino acids, nucleotides, and lipids, or large ones, such as peptides, proteins, and even viral particles.[1,4−7] Chemical diversity on the ligand side corresponds to diversity in the receptor binding site and, hence, in the ligand−receptor interaction modes, as also highlighted by in vitro mutagenesis and biochemical experiments.[1,4−7] According to an oversimplified and broad view, small organic ligands are thought to bind into the TM portions of the helix bundle, whereas peptide compounds would dock primarily into the extracellular loops. The latter situation is particularly true for glycoprotein hormone receptors, for which the N-terminal domains have been predicted as being highly structured and holding the major recognition points for their natural agonists.[70,357−361] The very recently released crystal structure of the follicle-stimulating hormone (FSH) in complex with the ectodomain (i.e. the N-term) of the FSH receptor confirmed structure predictions that the major receptor binding site for the hormone is a portion of the N-term, holding a leucine-rich repeat (LRR) topology.[15] However, in the crystal structure, the LRR topology belongs to the class β,[15] rather than to the class $\alpha-\beta$ as predicted by early studies.[70,357−359,361]

Most of the computational experiments done so far on GPCRs concerned ligand binding site identification and predictions of the ligand interaction modes for either qualitative or quantitative structure affinity/selectivity relationships and drug design (reviewed also in refs 217 and 473−479). In contrast, very few computational studies have been aimed at investigating the structural changes induced by agonists into receptor portions more or less distal from the ligand binding site.[60,68,71,83,220,267,272,283,286,295,296,301,303,365,480] The results of these studies are reviewed in a separate section.

The literature is full of examples of more or less sophisticated computational approaches aimed at achieving interaction models between GPCRs and their ligands. Attempts to unravel the mechanism of ligand entry into the putative binding site have been done by a few computational experiments, based upon either Brownian dynamics[481] or MD coupled with quasi-harmonic analysis.[256] In the first case, Brownian dynamics on a model of the extracellular domain of the β_2-AR suggested that helical movement must accompany movements of the extracellular loops to allow the ligand to pass through.[481] The second case concerns computational experiments on a model of the extracellular loops of the thyrotropin-releasing hormone (TRH) receptor (TRHR).[256] The static analysis based on curvature and electrostatic potential on the surface of TRHR suggested the formation of an initial recognition site between TRH and the surface of its receptor. A quasi-harmonic analysis of the vibrations of the extracellular loops suggested that the low frequency motions of the loops could aid the ligand to access its transmembrane binding pocket. It was, hence, suggested that all small ligands may bind sequentially to the TM pocket by first interacting with the surface binding site and then may be guided into the transmembrane binding pocket by fluctuations in the extracellular loops.[256]

The difficulties in simulating the ligand entry into the receptor binding site are also linked to indetermination in the arrangements of the extracellular loops, which are expected to play a relevant role in ligand recognition. Extracellular domains are predicted to form a tightly bound canopy that makes ligand entry difficult. This is evident in the crystal structures of rhodopsin, in which the N-terminal tail forms a very compact β-sheet structure with E2, thus completely shielding the retinal chromophore from the extracellular environment. Such an arrangement of the extracellular domains, if conserved in a number of homologous GPCRs, suggests that dramatic conformational changes should occur to let the ligand enter the receptor binding site, or that interhelical locations may constitute the entrance. For rhodopsin, a potential entrance site for the chromophore has been, indeed, hypothesized to reside in the cytosolic domains (i.e. site II in Figure 2).[115] However, this

feature must be unique to rhodopsin and not to the homologous GPCRs, which recognize ligands from the extracellular side.

Thus, all the intermolecular interaction models reported so far were achieved following either manual or automatic docking of the ligand into the putative binding site of the target receptor model. Given the enormous number of reports on the matter, we prefer to broadly classify and quote them according to the approach used to bundle the TM helices, i.e., (a) by ab initio approaches strongly based on the 2D electron density maps of rhodopsin[321,329,482−484] or on the integration between 2D and 3D electron density maps of rhodopsin and inferences of Baldwin's analysis[60,83,238,240−244,246−248,250,255,258,260,262−267,269−278,280−292,298−301,303] or (b) by comparative modeling, employing either BRD[160,161,163,165−168,170−190,192−194,197,198,200−202,204,205,209,210,441,485−491] or ab initio models of rhodopsin,[311−313,325] or, since very recently, the crystal structure of bovine rhodopsin as a template.[68,71,141,324,347,361,365,369,370,373,376,378−394,397,399−411,415−422,424−438,440,441,480,492−503] Alternative approaches consisted of modifying helix-bundle models previously achieved using BRD as a template to incorporate the available structural information on rhodopsin, i.e., the cryo-microscopy 3D map,[12] the outcomes of Baldwin's study,[11] and the crystal structure.[13,214−216,504,505] A number of computational modeling studies on the GPCR ligand binding sites are also reviewed in refs 217 and 473−479.

Differences in the modeling approaches, however, concerned not only the ways of packing the TM helices but also the criteria for determining the beginning and ending of each TM domain, the absence or the more or less complete inclusion of the extracellular and intracellular domains, and the ways of modeling such hydrophilic domains. Differences also concerned the energy refinements of the ligand and the receptor before and after docking. In fact, a few steps of energy minimization or MD simulations were employed either to achieve the input receptor structure for manual or automatic docking or to produce the final ligand−receptor complex. Extremely variable computational setups were employed. Differences resided in the force field; the more or less implicit treatment of the hydrogen atoms; the protonation state of titratable amino acids or of ligand functional groups; the approaches for choosing the proper ligand conformation and charge distribution; the ways of capping the terminal amino acids; the ways of approximating the lipid/water environment; the ways of truncating the nonbonded potentials; the employment or not of intramolecular and intermolecular restraints; the type of restraints and the ways of applying them; the algorithms used for energy minimizations; the number of minimization steps; the molecular simulation methods (i.e. simple MD or mixed Monte Carlo/Stochastic Dynamics (MC-SD)); the integration steps; the length and organization of the different simulation stages (i.e. heating, equilibration, and production phases); the simulation ensemble; etc. A significant number of cases also exist in which no energy refinement has been done before and after the ligand−receptor docking. Overall, the receptor systems were subjected to short MD simulation time periods, which only in a few cases exceeded 500 ps (i.e. refs 68, 69, 71, 267, 281, 282, 302, 304, 365, 387, 394, 398, 414, 416, 421−424, 429, 430, 437, 438, 480, and 506 among the articles on the free and/or ligand-bound receptor forms reviewed in this article). In some cases, the lengths of MD simulations were exceedingly short (i.e. 250−500 ps unrestrained production phases) for the molecular systems under study (i.e. all-atoms TM-receptor model in explicit methane[397,494,495,507] or lipid/water[431,435,508]).

Despite the enormous diversity in the computational approaches to GPCR modeling, it frequently happened that substantially different models of the same ligand−receptor complex were found consistent with the available experimental data, independently of the accuracy and reliability of the computational model. This is a consequence of the low resolution of the experimental data, such as the results of site-directed mutagenesis experiments, which were generally used both to drive the computational experiments and to evaluate the reliability of the models. In this respect, it is not surprising, for example, that an early model and a recent significantly different model of the melanin-concentrating hormone (MCH) interacting at its target receptor were both consistent with site-directed mutagenesis data.[200,365] In fact, the ligand−receptor interaction, suggested to be crucial for the MCH binding (i.e. the charge-reinforced H-bond between Arg11 of the hormone and D3.32 of the receptor), was present in both the models regardless of the different stereochemistries of the complexes.[200,365] It is noteworthy that the early model included only the TM helices in a BRD-like architecture, and the extracellular loops.[200] In contrast, the novel model was achieved by using the crystal structure of rhodopsin as a template and included all the domains but the C-tail.[365] The relevant differences between the two models originate also by the fact that the early model was simply subjected to a few steps of energy minimization, whereas the novel one was the product of a large series of MD simulations.[200,365] This example, that is quite representative of the many ones reported in the literature, suggests that the results of site-directed mutagenesis experiments are inadequate to assess or improve the stereochemical quality of an atomistic model. Significant insight into mapping the ligand−receptor interactions could be provided by photoaffinity labeling and FRET experiments, which could estimate spatial proximity between photolabile residues within a ligand and its receptor.[224,227,228,233,236,509,510] The resolution level of the information inferred by such approaches, although higher than that achieved by common in vitro site-directed mutagenesis experiments, is, however, still not high enough for help in assessing the stereochemical quality of atomistic models. This is clearly demonstrated by a novel report, showing that two substantially different models of E2 in the human NK1 receptor, one obtained by comparative modeling, using the crystal structure of rhodopsin as a template, and the other achieved by NMR determinations on the isolated peptide, both were consistent with distance estimations by photoaffinity labeling experiments.[376]

Computational models of GPCRs, which do not incorporate structural information from the high resolution structure of rhodopsin, have been useful in driving and interpreting the results of low resolution experiments. However, their indetermination is too high for de novo drug design purposes, which would require knowledge of the exact stereochemistry of the ligand binding site. This is particularly true for the interaction models involving huge peptide ligands, which are expected to establish more contact points with the receptor, as compared to small cationic compounds, and many of these contacts are predicted to involve the extracellular loops, whose conformation and architecture are ill-defined in the majority of the computational models not based upon the rhodopsin structure. The best examples of lead optimization based on 3D models of GPCRs relied on robust and extensive SAR analyses on highly informative sets of compounds, such as, for example, the ligands of the adenosine receptors.[476]

The high resolution structure of rhodopsin can now be used as a template for achieving acceptable starting models of the empty receptor, including almost all the extracellular domains, and, hence, of the ligand−receptor complexes. In fact, especially for peptide GPCRs, the majority of the extracellular domains, first of all E2, which is expected to be heavily involved in the interaction with the natural ligands, can be modeled with acceptable reliability, by using the homologous domain in the rhodopsin structure as a template. In this respect, we are confident that comparative modeling would produce more reliable results than NMR determinations on the isolates peptides, as largely demonstrated by the inconsistency between the cytosolic domains of rhodopsin determined by NMR and those in the crystal structures.[54,511−514]

Once different starting configurations of the ligand−receptor complexes fulfilling the ligand−receptor shape and electrostatic complementarities and the key interaction requirements from the experimental data have been obtained, these must necessarily be subjected to molecular simulation protocols aimed at finding suitable different local minima of the potential energy surface. This has to be done independently of the functional state of the considered ligand (i.e. full, partial, inverse agonist, neutral antagonist, etc.), essentially because GPCRs exist in different conformational states already in their empty forms and the shape of their potential energy surface changes on the basis of the structural features of the interacting ligands. An intriguing inference from recent docking simulations between agonists and antagonists and different GPCRs, including the 5-HT_{1A} serotonin receptor, MCHR1, and MCHR2 and OTR, is that the same ligand, depending on its interaction modes, can generate different average configurations of the same receptor. These configurations differ from those of the empty receptor forms. On the other hand, comparing a large number of average configurations obtained for the different agonist− and antagonist−receptor complexes, following different computational protocols, it has been possible to infer similarities in the interaction modes of the different agonists at their

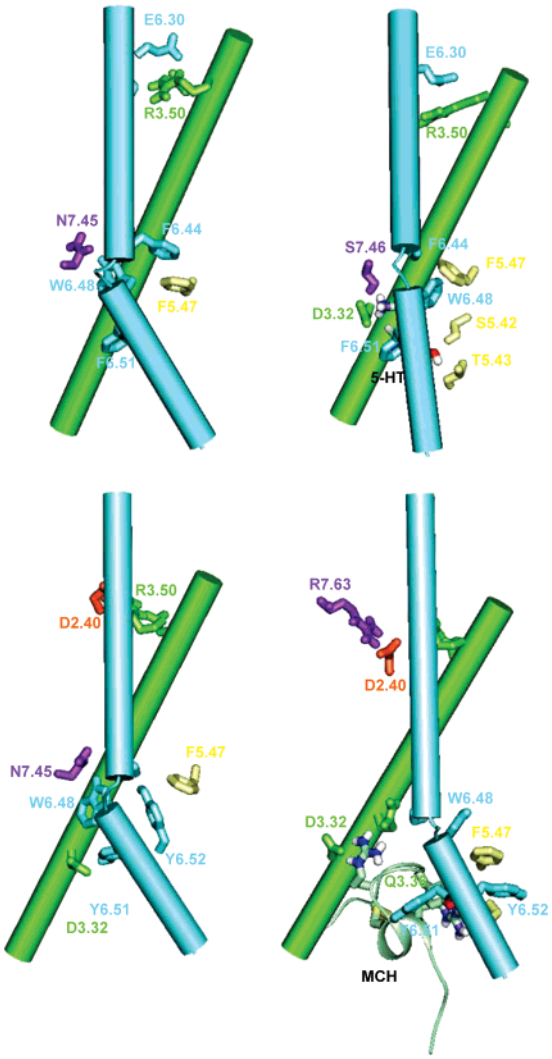


Figure 12. (top) Structures of the empty (left panel) and of the 5-HT-bound (right panel) forms of the 5-HT_{1A} receptor averaged over the 1000 structures collected during the last 500 ps of a selected 2 ns trajectory and minimized.[68] (bottom) Structures of the free (left panel) and MCH-bound (right panel) forms of the MCHR1 averaged over the 200 structures collected during the last 100 ps of a 2 ns trajectory and minimized.[365] The models are viewed in a direction parallel to the membrane surface, with the cytosolic side being at the top. Only H3 and H6 are shown, represented by cylinders. Details of the interactions made by R3.50, by the agonists, and by the members of the aromatic cluster in H6 are shown. The amino acid side chains are colored according to their location (i.e. the helix they belong to), following the color coding described in the legend to Figure 11. 5-HT and MCH are colored by atom type. 5-HT is represented by sticks (top right), whereas MCH is represented by cartoons (bottom right).

cognate receptors.[68,71,365] The receptor sites, in which most of such similarities occur, essentially concern selected positions in the extracellular halves of H3, H5, and H6 (Figure 12). Computational modeling of the agonist-bound forms of the 5-HT_{1A}, MCH, and OT receptors suggests that these receptor portions hold the key contact points for the ligand moieties responsible for the efficacy.[68,71,365] These receptor essential contact points overlap, at least in part, with those which have been recently proposed to constitute conserved GPCR binding pockets recognized by "privi-

leged structures".[400] The term "privileged structures" indicates an ensemble of structural motifs, shared in common by ligands, which bind different receptors, in other words, "a single molecular framework able to provide ligands for diverse receptors".[400,475] We have shown that the combination of rhodopsin-based receptor models with extensive comparative MD simulations could be able to find such privileged structures even in large peptide ligands such as MCH and OT.[71,365] These results would have been hard to achieve with static models. In our opinion, the most promising approaches to the search for a linkage between ligand−receptor atomistic models and biological responses (i.e. affinity, selectivity, and efficacy) will be those based on ever increasingly accurate and effective configurational sampling methods.

4.3.1. Computational Approaches to Virtual Screening of GPCR Ligands

A challenging task in computational modeling of GPCRs is to define intermolecular interaction descriptors and computational protocols, which could be used for virtual functional screening (reviewed also in ref 515). The effectiveness of these computational descriptors will rely on the reliability of the ligand−receptor complex, on the accuracy of the computational protocols, and on the availability of highly informative training sets of compounds with accurately determined biological responses. Pioneers in this respect have been computational experiments aimed at defining intermolecular interaction descriptors on the ligand−receptor complexes, which could correlate linearly with biological data such as binding affinities or efficacies.[172,173,242,295,299−301,383,516] In this respect, an intermolecular interaction descriptor, somewhat related to the binding energy and, hence, called BE, was used to rationalize and predict the binding affinities and selectivities of different GPCR ligands.[165,173,242,300,383] Such a descriptor, which ill-defines all the entropic effects, is computed on the ligand−receptor energy-minimized complexes according to the following formula: $BE = IE + E_R + E_L$, where IE is the ligand−receptor interaction energy (i.e. the total energy of the ligand−receptor complex minus the energy of the ligand and the receptor in the complex) and E_R and E_L are the distortion energies of the receptor and of the ligand, respectively, calculated as the differences between the energies of the bound and of the free optimized molecular forms. For a series of selective and nonselective antagonists of the α_{1a}-, α_{1b}-, and α_{1d}-AR subtypes, consistency was found between the affinity/selectivity predictions by the intermolecular interaction descriptor BE[300] and those by a ligand-based approach, i.e. the supermolecule approach.[517−519] The supermolecule approach assumes that the volume obtained by superimposing the most structurally different ligands showing the highest affinities for the same receptor (supermolecule) might reflect the overall shape and conformational flexibility of the high affinity receptor binding site. Therefore, size and shape descriptors can be defined ad hoc (that is, on a specific molecular series and in connection with a specific bioactivity) with respect to the supermolecule, and this constitutes the main advantage over molecular descriptors defined and performed for a single structure and a single conformation. The most effective indices in this respect were V_{in} and V_{out}, which are respectively the intersection and the outer van der Waals volumes of the ligand considered with respect to the volume of the reference supermolecule, and V_{dif}, which is computed according to the following formula: $V_{dif} = (V_{in} - V_{out})/V_{sup}$, where V_{sup} is the molecular volume of the reference supermolecule. According to the definition of size and shape descriptors, higher affinities are realized by maximizing V_{in} and minimizing V_{out}.[517−519]

Convergence was found concerning the inferences from the ligand−receptor interaction models and from the supermolecule approach.[300,517−519] It was, in fact, inferred that, while the protonated nitrogen atom is an essential pharmacophoric element for the long-range electrostatic recognition and productive interaction with the aspartate of the α_1-AR binding site, its contribution to the interaction energy is relevant but almost constant.[300] In contrast, short-range forces modulate both ligand affinity and selectivity. Thus, the modulation of the binding affinity by a wide noncongeneric series of α_1-AR ligands can be described and explained by the variation of the ligand size−shape features that are related to the short-range acting forces. In this respect, the intermolecular interaction descriptor, BE, is rather independent of the reliability of the ligand−receptor complexes.

Assuming that differences in entropic and solvation/desolvation terms are small within a series of structurally similar ligands, Shim et al. approximated the binding free energy of a series of CB1 ligands to the interaction energy. The latter is the total energy of the ligand−CB1 complexes minus the energies of the unbound ligand and receptor forms.[396] This index was used to evaluate the best binding conformations for a series of nonclassical CB1 agonists.[396]

Very recently, an intermolecular-interaction descriptor has been proposed for its qualitative agreement with the in vitro binding affinity data concerning selected GPCR ligands.[329,483,484] This descriptor, which merely relies on the ligand energetics, being the difference between the potential energy of the free solvated and the receptor-bound forms of the ligand, has been proposed as potentially useful for in silico screening.[329,483,484] The AutoDock 3.0.3 scoring function computed on the ligand−receptor complexes was, instead, employed as computational binding free energy in linear correlations with the in vitro-determined binding affinities of CCR5 antagonists.[420]

One of the first attempts to build computational models for in silico functional screening of GPCR ligands is represented by computational experiments on 34 functionally different ligands (i.e. antagonists as well as partial and full agonists of the M_1-muscarinic receptor).[172] In this respect, the percent of the van der Waals contribution (IE_{VDW}%) or the percent of the summation over the H-bonding and electrostatic contributions (IE_{HB+EL}%) to the total interaction energy of the ligand−receptor complexes

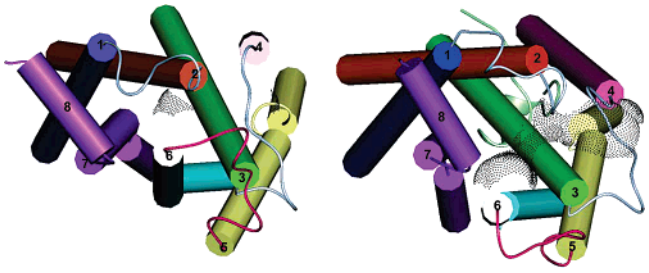


Figure 13. Structures of the empty (left) and MCH-bound (right) forms of MCHR1, averaged over the 200 structures collected during the last 100 ps of the selected 2 ns trajectories and minimized.[365] The extracellular loops are not shown. H1, H2, H3, H4, H5, H6, and H7 are respectively colored in blue, orange, green, pink, yellow, light blue, and violet, whereas I1, I2, and I3 are respectively colored in light green, gray, and purple. The amino acid stretch corresponding to rhodopsin's H8 is colored in violet as well. The hormone is colored by atom type. The structures are seen from the intracellular side in a direction perpendicular to the membrane surface. The solvent accessible surface computed over N76(2.37), R141(3.50), K153 (in I2), I247(6.26), and T251(6.30) is shown represented by gray dots.[365]

proved effectiveness in predicting the functional behavior of the targeted ligands.[172] In fact, the IE_{HB+EL}% index was below or close to 20% (IE_{VDW}% above or close to 80%) for antagonists, above 30% (IE_{VDW}% below or close to 70%) for full agonists, and between the previous two values for partial agonists.[172] These computational indices were successfully challenged in their ability to predict the functionality of novel muscarinic ligands docked into an upgraded model of the human M_1-muscarinic receptor.[299] Thus, these intermolecular interaction descriptors appeared effective, independently of the stereochemistry of the ligand−receptor complexes. This may be due, at least in part, to the fact that they are correlated with intrinsic properties of the ligands such as the size−shape and the propensity to perform H-bonding interactions. In fact, for the same training set of compounds, considered in the same bioactive prototropic forms, it was also possible to define molecular orbital (MO) derived descriptors, such as the electrophilic superdelocalizability, or size−shape descriptors computed through the "supermolecule approach", which correlated linearly with the ligand efficacy, though less efficiently than the intermolecular-interaction descriptors.[516,520]

A computational protocol, based upon extensive comparative MD simulations of the free as well as agonist- and antagonist-bound forms of MCHR1 and MCHR2, was recently proposed, which may be used for virtual functional screening of MCHR ligands.[365] This approach relies on size−shape descriptors computed on the average minimized structures representative of the different functional states of the receptor. In detail, the SAS computed over selected amino acids in the cytosolic domains proved to be a good hallmark of the functional receptor state, being close to 0 $Å^2$ in the nonactive states (i.e. empty and antagonist-bound forms) and significantly above 100 $Å^2$ in the active states (i.e. agonist-bound forms) (Figure 13). The amino acids employed for SAS calculations were defined ad hoc for each of the two MCHR subtypes. The predictive ability of the SAS indices was challenged on mutant forms of the hormone obtained by in silico alanine-scanning mutagenesis of all the 19 amino acids of the hormone, except for the two cysteines involved in a disulfide bridge.[365] A truncated form of the hormone, lacking the last three amino acids, was also considered. In detail, each modified form of the hormone was docked in both the MCHR1 and MCHR2 exactly in the same orientation and conformation selected as input for the wild type MCH. Docking was followed by MD simulations, according to the same computational protocol employed for the wild type MCH and for the antagonist, leading to the calculation of the SAS index. Consistency between in silico-predicted and in vitro-determined functionalities was obtained, as the SAS indices were below 100 $Å^2$ only for the MCH mutants with impaired functionality.[365] Interestingly, for both the MCHR subtypes, the SAS indices correctly predicted that deleting the last three amino acids of the hormone does not cause any impairment in functionality.[365] This, we think, is an encouraging example of in silico functional screening based upon a computational descriptor of allosteric structural modifications in distal cytosolic domains as induced by critical intermolecular interactions in the extracellular half of the receptor.

The AutoDock 3.0 scoring function was employed in virtual screening of $P2Y_1$ ligands.[497] The receptor model, which comprised only the TM domains, was achieved by fitting each helix into the corresponding one of the rhodopsin structure.[521] A database of 500 compounds was screened. Of the 30 top-ranked compounds, 21 were selected for in vitro functional screening. Twelve compounds were, finally, identified as agonists or antagonists of the $P2Y_1$ receptor.[497]

Combinations of the scores provided by different docking algorithms have been recently used by an in silico screening study aimed at finding lead compounds for target GPCRs.[385] In detail, Bissantz and co-workers constructed "antagonist-bound" molecular models of three human GPCRs (dopamine D3, acetylcholine M_1-muscarinic, and vasopressin V_{1a} receptors).[385] These models were achieved by manual docking of a known antagonist into a homology model (based upon the crystal structure of rhodopsin) of its cognate receptor, followed by energy minimization. Removing the ligand atoms from the minimized complexes finally yielded, for each of the three receptors, one set of coordinates of what the authors considered to be representative of the "antagonist-bound state".[385] The three models were evaluated in terms of their ability to identify known antagonists seeded into a database of randomly chosen "druglike" compounds. A virtual screening procedure based upon the combination of different docking algorithms in association with seven scoring functions was used.[385] Consensus scoring was then applied to generate small subsets (hit lists) comprising only the top scorers common to two or three scoring lists. Agonist ligands were also considered in a separate study. Different agonist-bound GPCR models were built to account for the knowledge that the "activated

states" of GPCRs are conformationally more flexible than the antagonist-bound inactive states. Two agonist-bound models were built for each receptor following the same refinement procedure as that used for the antagonist-bound models, but using full agonists rather than antagonists for receptor refinement, merely based upon energy minimization.[385] The conclusions drawn from this study are that rhodopsin-based models are suitable for virtual screening, since known antagonists seeded in the test databases could be distinguished from randomly chosen molecules. However, such models are not accurate enough for retrieving known agonists.[385] Better suited for agonist screening is a knowledge- and pharmacophore-based modeling procedure that might partly simulate the conformational changes occurring in the active site during receptor activation.[385] The authors speculated that receptor coordinates generated by such a new procedure are suitable for agonist screening and that two alternative strategies relying on different sets of receptor coordinates are needed for the virtual screening of GPCR ligands, depending on whether they are agonist or antagonists. We agree with the author's idea that for antagonist and agonist screening it is better to use, respectively, antagonist- and agonist-bound states of the target receptor. A simple energy minimization is, however, not sufficient to explore the configurational space of the ligand-bound states, whatever the functional behavior of the ligand is. Thus, despite the apparent success of the virtual screening done on static receptor models, we would expect that an exhaustive sampling of the complexes between selected ligands and the receptor, leading to the extraction of one or more representative average configurations of the empty receptor, would improve the results of the screening.

Recently, ab initio computational models of GPCRs, generated by the PREDICT method, have been used for in silico screening.[334,522] Also in this case, virtual screening was performed on the receptor coordinates extracted from a refined complex with an ad hoc selected antagonist. This strategy is instrumental in overcoming the drawback that all high throughput docking procedures must use a rigid protein structure, which would impair the results of in silico screening. To generate the ligand−receptor complex for virtual screening, a potent small-molecule ligand is placed inside the binding pocket in an initial orientation that agrees with proposed key interactions in the site. To overcome the bias introduced by the imposed initial orientation, a free-energy-perturbation-like simulation is done on the ligand−receptor complex to allow the ligand to reorient itself in the site in response to specific ligand−protein interactions. Screening of the compound library against the receptor models is then performed with the DOCK4.0 software,[523] employing a special screening and scoring protocol.[522] At the end of the docking process, the compounds are sorted on the basis of their best docking score. The effectiveness of the virtual screening is estimated by the following two measures: "enrichment factor", which is related to the common measure of "hit rate", and "coverage." An enrichment factor measures the success of virtual screening relative to random screening (which is equivalent to experimental high throughput screening) by comparing the virtual screening hit rate with the hit rate of random screening. A hit rate is the fraction of active compounds found in the screening library by whatever method. The other parameter, "coverage", is defined as the number of active compounds M^q that are included within the top q percent of the ranked library vs the total number of active compounds in that library M: $\text{coverage}^q = M^q/M$.

The approach was first tested on the D2, NK1, and NPY Y1 receptors, by enriching the random libraries of known ligands of these receptors, followed by blind screening.[334,522] As for testing, coverage was 85%, 80%, and 50% for D2, NK1, and NPY Y1, respectively. The blind screening targeting the D_2, 5-HT_{1A}, 5-HT_4, and CCR3 receptors always found hits with a binding affinity $< 5\ \mu$M with hit rates of up to 21%. Thus, in general, the assays validated the quality of the hits as lead compounds for drug discovery. We think that these encouraging results of virtual screening would be significantly improved by increasing the resolution of the structural information incorporated while modeling the receptor. So far, the highest resolution information on the members of the rhodopsin family can be inferred from the crystal structure of rhodopsin. Contrarily to what the authors state, we believe that careful comparative modeling aimed at extracting as much stereochemical restraints as possible from the rhodopsin structure, possibly including the extracellular and intracellular domains, would provide more reliable starting models of GPCRs than the most effective ab initio modeling approaches currently available.[334,522] In this respect, the authors' statement that their ab initio models are better than those based upon the rhodopsin structure only because they are in better agreement with site-directed mutagenesis data is not acceptable, since the resolution level of site-directed mutagenesis data is low and cannot be used to evaluate the stereochemical quality of atomistic models. Furthermore, the extracellular loops, completely neglected in the ab initio models by the authors, are expected to play a meaningful role in recognizing a significant number of GPCRs' ligands.

Virtual screening by means of ligand−receptor docking was also used by Gouldson et al. to determine whether the inactive model of the β_2-AR showed preference for antagonists, while the restraint-based active model showed preference for agonists.[139] The models were challenged by 172 GPCR ligands, 50% of which were adrenergic ligands. The binding site was primed for docking by interactively docking norepinephrine or S-propanolol into the active or inactive models, respectively, minimizing, running MD simulations for 100 ps, and then carrying out a final minimization. High performance of the active-state model in retrieving agonists was shown.[139]

Evers and Klebe developed the MOBILE (modeling binding sites including ligand information explicitly) approach to build GPCR models suitable for virtual screening.[411,524] With this approach, the information from a selected ligand binding is added as additional restraints during comparative modeling by MOD-

ELLER.[411,524] A final model is then chosen from the individual models that best explain the observed ligand binding. Combining the information from the 3D receptor model and the analysis of known NK1 receptor antagonists, a pharmacophore model was deduced, which was used for 2D and 3D database searches.[411] The hits from such searches were docked into the modeled binding pocket of the NK1 receptor, leading to the final selection of seven compounds for the biochemical testing, including one in the submicromolar range. A similar approach was also employed to virtual screening of α_{1a}-AR antagonists.[501] In this case, the receptor model was achieved by using a modified version of the MOBILE approach, consisting of the following steps: (a) a ligand is docked into an ensemble of homology models of the target proteins; (b) refinement of the models is done by considering explicitly the interactions with a selected ligand; (c) finally, the best models in terms of intermolecular interactions are selected and assigned side chain conformers from a set of selected models, followed by energy optimization of the entire complex. The finally selected model was used for virtual screening by docking a database of 22 950 compounds resulting from the application of a series of hierarchical filters, which also included fulfillment of the Catalyst-generated pharmacophore requirements.[501,525] The top-scoring 300 compounds were, hence, subjected to cluster analysis, and 80 were sent to in vitro testing of binding affinities. Thirty-seven compounds revealed K_i values better than 10 μM, and the most active compound had a 1.4 nM affinity. We do not think that biasing the results of comparative modeling by low resolution information on the binding properties of a selected compound leads to more reliable receptor models. We rather think that giving the receptor the fingerprints of a lead compound is exclusively instrumental in ligand−receptor docking-based virtual screening that would not be possible otherwise.

A promising computational approach for binding-site identification and in silico screening of GPCRs with a potential to be used to "deorphanize" orphan GPCRs was recently developed and tested on bovine rhodopsin and BRD.[526] The methodology, based on the Internal Coordinate Mechanics (ICM) program, was validated in accurately identifying the ligand-binding pocket of the two TM proteins. Furthermore, ICM flexible docking with and without the loop regions accurately predicted the binding geometry of retinal. The authors also demonstrated that the native ligand can be identified by flexible docking and scoring in 1.5% and 0.2% (for rhodopsin and for bacteriorhodopsin, respectively) of the best scoring compounds from two different types of compound databases. Finally, it was also demonstrated that, even if the side chain positions in the rhodopsin binding pocket are entirely wrong, their correct conformation can be fully restored with high accuracy (0.28 Å) through ICM global optimization with and without the ligand present.[526] These binding site adjustments are critical for flexible docking of new ligands to known structures or for docking to GPCR models based upon the rhodopsin structure. Of course, the effectiveness of the approach in virtual screening of GPCR ligands is expected to strongly depend on the accuracy of the GPCR model.

4.4. Computational Modeling of Ligand-Induced/Stabilized Active States

The few computational experiments reported in the literature, aimed at simulating the active and nonactive states of GPCRs induced/stabilized by agonists or antagonists, respectively, have been essentially based on comparative analyses of the average structures derived from MD simulations of the free and agonist- and antagonist-bound forms of the target receptor.[60,68,71,83,220,267,272,283,286,295,296,301,303,365,480] Ligands were manually or automatically docked either into the input or the average structure of the receptor, driven by the available indications from in vitro experiments.

Early studies include that by Zhang and Weinstein on the TM domains of the 5-HT_{2A} serotonin receptor.[220] MD simulations of the interactions between the 5-HT_{2A} receptor and ligands of different pharmacological efficacies suggested that the binding of a full agonist, but not an antagonist, produces a marked structural change at the cytosolic ends of H5 and H6. MD trajectories showed that the secondary structure of the TM domains of the receptor is well maintained, whereas the conformational changes involve mainly the relative translations and rotations of the helices in the bundle. An algorithm was, hence, developed for improving the analysis of the ligand-induced domain motions over the different MD trajectories obtained for the 5-HT_{2A} receptor.[527]

The β_2-AR was the target of the computational experiments by Gouldson and co-workers, who docked agonists, antagonists, and partial agonists into a receptor model achieved by using Baldwin's α-carbon model as a template and ran MD simulations over a period of 500 ps.[83] The major structural changes were found to occur in the intracellular halves of H5 and H6. The agonist-induced structural changes to H5 and H6 were thought to be large enough to induce a conformational change in I3. An alternative hypothesis inferred by computations was that the change in the tilt of these helices might enhance the formation of a H5,H6-dimer due to the formation of a more optimal helix packing at the dimer interface. Indeed, in the unbound receptor, H5 and H6 were found to be essentially perpendicular to the membrane. However, after the agonist-induced structural changes, the tilt of these helices had changed by approximately 20°.[83]

Sylte and co-workers performed MD simulations on agonist- and antagonist-bound forms of the NPY Y_1 receptor.[272] Packing of the energy-minimized structure of each TM helix was driven by the inferences from the first Baldwin study.[10] All the intracellular and extracellular domains were, then, included in a conformation assigned by MD simulations. The main inferences from this study were that the agonist NPY induces motions of H5 and H6, whereas the antagonist BIBP3226 does not.[272] The same team repeated similar computational experiments on the 5-HT_{1A} receptor, drawing overlapping conclusions

with their previous study on the NPY Y_1 receptor.[272,286] The initial receptor structure was achieved by modeling the helix bundle on Baldwin's Cα-atom model, taking the N-term structure from an earlier model, and using a database loop search to achieve the three extracellular loops as well as I1, I2, and the C-tail.[286] The free as well as full agonist-, partial agonist- and antagonist-bound forms of the receptor were subjected to comparative MD simulations. In the simulations, the agonist induced larger conformational changes in H3 and H6 compared to the other helices, whereas the main conformational differences between the agonist- and antagonist-bound receptor forms essentially resided in H5 and H6. During the simulations, all the ligands constrained helix movements, compared to the case of the empty receptor form.[286] In a more recent study, following docking and MD simulations of a number of ligands characterized by different functionalities at pre- and postsynaptic 5-HT_{1A} receptors into an updated model of the 5-HT_{1A} receptor, the authors speculated that presynaptic antagonistic behavior is connected to large movements of H7, whereas presynaptic agonistic behavior is associated with large displacements of H2 and small displacements of H7. Moreover, postsynaptic partial agonist behavior would be connected to large displacements of H4 and H5, which, instead, would show a small displacement following interaction with postsynaptic antagonists.[480]

Over the last 11 years, we have done extensive studies aimed at investigating the propagation of the structural modifications from the ligand binding site to distal receptor domains, following the docking of selected agonists and antagonists into their cognate receptors.[60,68,71,295,296,301,303,365] Targets of our study have been different members of family A, including the α_{1b}-AR, M_3-muscarinic receptor, OTR, 5-HT_{1A} serotonin receptor, MCHR1, and MCHR2.[60,68,71,295,296,301,303,365]

The computational approach was essentially the same as that employed to infer the structural differences between mutation-induced active and nonactive receptor states, the only difference being the perturbation introduced in the initial model, i.e., ligand docking, in one case, and point mutation, in the other. The computational approach consisted of manual docking of selected agonists and antagonists into the putative binding sites of the average minimized receptor structures, in early studies, or, more recently, in the input receptor structure derived by MODELLER refinements, followed by comparative MD simulations. Manual docking was driven, at least in part, by the available indications of site-directed mutagenesis experiments. The receptor models employed in early studies were achieved by ab initio modeling,[60,296,301,303] whereas those employed in the latest studies were achieved by comparative modeling, using the crystal structure of rhodopsin as a template.[68,71,365]

Early computational experiments focused on the α_{1b}-AR and on the M_3-muscarinic receptors lacking the extracellular and intracellular domains.[295,296] First, the average minimized structures of the free receptor were analyzed to identify the hydrogen bonding interactions involving the most conserved polar amino acids. Successively, analyses of the antagonist- and agonist-bound receptor forms were done, focusing on the ligand-induced perturbation in such a H-bonding network. For both the receptors, antagonists substantially preserved, whereas agonists perturbed, the H-bonding network found in the empty forms. By different binding modes and different dynamics mechanisms, the different agonists induced, as a final result, the movement of R3.50 of the E/DRY motif out from a highly conserved polar pocket, possibly exposing this residue ready for interaction with the G protein. This model was in agreement with the hypotheses by Oliveira and coworkers.[528] MD simulations on more complex ab initio receptor models (i.e. carrying the extracellular and intracellular domains) of the α_{1b}-AR and of OTR indicated as well the fully conserved arginine as the target of the structural modifications induced by agonists.[60,303] These studies suggested also that agonist-triggered helix movements would induce the opening of a solvent accessible crevice involving I2, I3, and the cytosolic extensions of H3 and H6.[60,303]

Further insight was achieved by applying the same computational approach to models of the 5-HT_{1A} serotonin receptor, OTR, MCHR1, and MCHR2 built by comparative modeling, using the crystal structure of rhodopsin as a template.[68,71,365] The models of OTR and MCHRs included all the intracellular and extracellular domains but not the C-tail, whereas the model of the 5-HT_{1A} included all the domains but not the N-terminus and the huge I3. The inferences from the latest studies relied on more extensive MD simulations than the previous ones. Indeed, the approach consisted of comparative analyses of a large number of short (100 ps) and relatively long (up to 2 ns) MD trajectories, probing different starting structures as well as different intrahelical (for the empty receptor form) and intermolecular (for the ligand-bound forms) distance restraints. For the empty receptor form, the criteria for selecting the short trajectory that could be worth prolonging included the stereochemical quality of the average minimized structures and their degree of similarity to the rhodopsin structure. Despite the differences between previous and latest approaches and between the receptor systems under study, hallmarks of the ligand-induced active and nonactive forms of the receptors were found to involve R3.50, the arginine of the E/DRY motif, and the cytosolic extensions of H3 and H6. In fact, for the agonist-bound (i.e. active state) and the antagonist-bound (i.e. nonactive state) forms, the establishment of crucial intermolecular interactions (as suggested by the experimental evidences) was found, respectively, concurrent with destabilization and reinforcement of the intramolecular interactions involving the E/DRY arginine in the empty receptor forms.[68,71,365] Therefore, the choice of the ligand−receptor complexes, which better fulfill the intermolecular interaction requirements, implies selecting divergent interaction patterns of the E/DRY motif in the agonist-bound (active) and antagonist-bound (nonactive) forms.

A common inference from the comparative analysis of the three substantially different receptors was that the agonist-bound forms share the release of all or some of the charge-reinforced H-bonds involving R3.50 in the empty receptor. The aspartate at position 3.49 is one of the arginine partners in the empty and antagonist-bound states (i.e. inactive states) of all the three receptors, whereas E6.30 is the other partner of R3.50 only in the 5-HT_{1A} receptor (Figure 12). In fact, in the empty and antagonist-bound forms of the 5-HT_{1A} receptor, both the D133(3.49)−R134(3.50) and R134(3.50)−E340(6.30) interactions are almost persistent during the 2 ns time of the MD simulation.[68] However, in the empty form, the intrahelical salt bridge is more stable than the interhelical one, as the former is conserved during the whole 2 ns simulation, whereas the latter is alternately lost and gained.[68] The loss of the H-bonding contribution to this interhelical interaction is always concurrent with the establishment of another H-bonding interaction between the E/DRY arginine and T343(6.33). In empty OTR, the lack of a glutamate at 6.30 is compensated by T273(6.33), which is involved in H-bonds with the E/DRY arginine in the second half of the most representative 1 ns trajectories.[71] In contrast, in both the empty forms of the MCHR1 and MCHR2 subtypes, which both lack an anionic amino acid at 6.30, the E/DRY arginine in the inactive forms is engaged in interhelical salt bridge interactions with an aspartate at position 2.40 and, only for MCHR1, in H-bonding interaction with T251(6.30) (Figure 12).[365]

Interestingly, for the 5-HT_{1A} receptor, the analysis of an extensive number of different MD trajectories suggested that the interactions with three relevant amino acids of the putative ligand binding site, in the extracellular half of the helix bundle, are essential for both the agonists (i.e. serotonin and (R)-8-OH-DPAT) to destabilize the salt bridge interactions that, in the empty and antagonist-bound receptor forms, involve the E/DRY arginine. These amino acids are (a) D3.32, which recognizes the protonated nitrogen atom of the ligand, (b) S5.42, which interacts with the indole or tetralin hydroxy groups of the ligands, and (c) F6.51, which interacts with the indole or tetralin ring of the ligands. The occurrence of the latter interaction relies on the establishment of the first two. Interestingly, the establishment of a double H-bonding connection between the agonist serotonin and both S5.42 and T5.43 is concurrent with the highest degree of destabilization of both the salt bridges involving the E/DRY arginine.[68] An interesting suggestion inferred from comparative MD simulations of the 5-HT_{1A} receptor and MCHRs, which differ both in amino acid composition and in the natural agonist (i.e. a small biogenic amine for the 5-HT_{1A} and a huge cyclic peptide for the two MCHRs), is that the essential requirements for the agonist to perturb the peculiar features of the inactive receptor forms are to accomplish charge-reinforced H-bonds with the binding site aspartate and to make van der Waals attractive interactions with one or more members of the aromatic cluster in H6 (Figure 12).[68,365] These interactions involve essentially two out of the 19 amino acids, which constitute MCH. In this scenario, the few critical interactions needed for the huge MCH or for the small serotonin to transfer the chemical information from the extracellular to the intracellular domains of their target receptors are overlapping (Figure 12). This hypothesis has been strengthened by the results of simulations of the OT−OTR complexes, suggesting that the few critical intermolecular interactions include that between Tyr2 of the agonist and F291(6.51) of the aromatic cluster in H6.[71]

Thus, for all the three different GPCRs, the agonist-induced chemical information transfer from the extracellular to the cytosolic domains (i.e. vertical information transfer) appears to be mediated by a cluster of aromatic amino acids in H6, i.e., formed by F6.44, W6.48, and F6.51, following the ligand interaction with selected amino acids in the extracellular half of the receptor.[68,71,365] In detail, the interaction between the aromatic ring of the agonist and F6.51 of the aromatic cluster induces a conformational change of W6.48, which loses its original H-bonding interaction with N7.45 and moves from H7 toward H5, consistent with the results of UV determination on rhodopsin (Figure 12).[350] These changes are concurrent with a significant reduction in the bend at the highly conserved P6.50, as compared to the empty and the antagonist-bound receptor forms, consistent with the results of conformational sampling on the isolated H6 of the β_2-AR and the CB1 receptor and restrained MD simulations on an almost complete model of the β_2-AR.[139,345,349] These results also agree with the computer simulation- and experiment-based hypothesis that GPCR activation would significantly diminish the kink at P6.50.[341] The straightening of H6 is one of the features of the agonist-bound forms correlated with weakening of the interactions made by R3.50, of the E/DRY motif. Another common feature to the agonist-bound forms of the simulated receptors is the release of the interaction found in the inactive forms between Y7.53, of the NPxxY motif, and a conserved phenylalanine in H8. This is particularly true for the computational models of the two MCHRs, whose agonist-bound forms are characterized by the approaching of H8 to the cytosolic extension of H3.[365] Although this structural change may be due to an overestimation of the electrostatic interactions due to the absence of the screening effect of water, it is also true that the approaching of H8 to the cytosolic extension of H3 is a feature of the active forms of MCHRs and not of the nonactive ones. Active and nonactive forms of both receptors have been simulated by using the same computational setup.[365] These structural changes, suggestive of increased flexibility of H7 and H8, are consistent with the experimental evidences that a disulfide bridge is allowed to form between the amino acids at positions 7.63 and 3.55 in the light-activated states of rhodopsin, and not in the dark state.[123] Other structural changes in the cytosolic domains, which are concurrent with the establishment of the few critical interactions between MCH and its cognate receptors, include the opening of a solvent accessible crevice

involving the second intracellular loop and the cytosolic extensions of H2, H3, and H6. This structural change is properly described by the solvent accessible surface area computed over selected amino acids, including the arginine of the E/DRY motif. This index, which was significantly lower than 100 Å^2 in the empty and antagonist-bound receptor forms and significantly higher than 100 Å^2 in the agonist-bound receptor forms, proved effectiveness as a hallmark of the functional receptor state (i.e. active or nonactive) and suitability for virtual functional screening of novel MCH ligands (Figure 13).[365]

Agonist binding induces significant changes in the arrangements of the extracellular ends of the helices and in the conformation/orientation of the N-terminus and E2, the latter being strongly involved in interaction with the peptide agonists OT and MCH.[71,365] The high extent of the structural differences in the putative agonist binding site between free and agonist-bound receptor forms suggests that the role of agonists cannot be limited to a conformation selection but it should also include the triggering of relevant structural changes, unlikely to occur spontaneously in the empty receptor form. The same is true for the antagonist-bound receptor forms. Of course, these inferences should be taken with caution given the indeterminations in the extracellular and intracellular domains of the receptor models and the approximations of the computational approach.

The hypothesis of the intrahelix salt bridge interaction between D/E3.49 and R3.50 of the E/DRY motif as the target of the structural perturbation induced selectively by agonists and not by antagonists was challenged in computational experiments on the free as well as the agonist- and antagonist-bound forms of the δ, κ, and μ subtypes of the mouse opioid receptors.[267] The receptor models were limited to the TM helices, which were packed using the Baldwin's Cα-atom model as a template.[11] MD simulations of the empty forms of all the three opioid receptor subtypes proved the stability of the intrahelix salt bridge between the two charged members of the E/DRY motif. For the δ opioid subtype, selected energy-minimized complexes with an agonist, etorphine, and an antagonist, naltrexone, were prepared in the putative "inactive" and "active" states, differing, respectively, in the presence or absence of the D3.49−R3.50 salt bridge. The four complexes were then subjected to a 1 ns MD simulation, and the outcome of the study was that, starting from the "inactive" form, the charge-reinforced H-bond between D3.39 and R3.50 is maintained over the whole simulation time, independently of the functionality of the bound ligand. In contrast, starting from the "active" form, the antagonist is able to restore the D3.49−R3.50 interaction, whereas the agonist does not.[267] It was also inferred that the agonist produces larger structural effects on H1, H3, and H6, whereas the antagonist induces larger effects on H4.[267]

Disruption of an intramolecular salt bridge between D170(3.32), the key recognition receptor point of amine ligands, and K379(7.36) was suggested to be the primary event leading to the α_{1d}-AR activation.[283] This hypothesis, in agreement with the results of in vitro experiments on the α_{1b}-AR,[529] was inferred from a computational model of the α_{1d}-AR, holding the TM helices, arranged according to Baldwin's model, and the three extracellular loops.[283] According to this model, the inactive state of the receptor would be stabilized by the D170(3.32)−K379(7.36) ionic lock and would move spontaneously into an active state, following the switch of the anionic partner of K379(7.36) from D170(3.32) to E375 (in E3). The active state of the receptor would be, hence, stabilized by the agonist, which would anchor to D170(3.32) of the receptor, through its protonated nitrogen atom. This model would provide an interpretation, with atomic detail, for the hypothesis of agonist selection of pre-existent active receptor states. However, an inconsistency in the mechanistic hypothesis concerns the antagonist-bound state that, similarly to the agonist-bound state, lacks the crucial ionic lock, though being inactive. The inferences of such a computational study could be misled by the low resolution of the computational model. In fact, models of the α_1-AR subtypes, based upon the crystal structure of rhodopsin, suggest that K7.36 would be more likely involved in an interhelical salt bridge with E2.65 instead of D3.32.[407] The structural relationship between E2.65 and K7.36 is also supported by their tendency to mutate in a correlated manner. In fact, in the 5-HT$_{1A}$ receptor, E2.65 and K7.36 are respectively replaced by a glutamine and an alanine.[407] The hypothesis of a potential role of D3.32 in stabilizing an inactive receptor state through an interhelical salt bridge is, hence, not supported by computational models based upon the high resolution structure of rhodopsin and has also been recently challenged by in vitro experiments on the 5-HT$_{2A}$ receptor.[276] Perhaps, upgrading the receptor model according to the advances in the determination of the rhodopsin structure would have led the authors to different conclusions.

The employment of a set of "activating" distance constraints, derived from the experimental studies of different GPCRs in the active conformation, was used to derive active state models of the μ and δ opioid receptors from the crystal structure of inactive rhodopsin.[496,503] These models were used to build the complex with agonists.[496,503] We expect that the reliability of these active state receptor models strongly depends on the specificity, quality, and resolution level of the experimental data employed as additional distance constraints to induce a structural divergence from the rhodopsin template.

5. GPCR Oligomerization

5.1. Insights from in Vitro Experiments

GPCRs have classically been assumed to exist and function as monomeric entities, and the paradigms of ligand binding and signal transduction were based on this hypothesis.

It is often said that, unlike growth factor and cytokine receptors, GPCRs were believed to exist as monomeric proteins in the membrane, and the likelihood that they could form dimers was refuted by the majority of investigators. It is more fair to say that oligomerization among GPCRs was considered a

tedious topic. In fact, data suggesting that GPCRs can oligomerize were presented recurrently over the years, but no evidence was ever found that such a physical interaction could be vital for GPCRs' function. The fundamental questions about GPCRs were focused on the mechanisms of interaction with agonists and G proteins. Whether that occurred between individual or gregarious partners was a matter of secondary importance. What suddenly stirred general interest on oligomerization was evidence showing that direct receptor−receptor interactions can rescue functional activity in complementation experiments. As shown by Maggio and co-workers, pairs of GPCR chimeras or truncated fragments, which were inactive when individually expressed, regained binding and signaling activity if coexpressed in the same membrane.[530,531] This, per se, did not identify which role oligomerization plays in GPCR function, but it definitely demonstrated that receptor−receptor interaction could affect the agonist binding pocket and the G protein interacting interface of GPCRs, as importantly, perhaps, as the intramolecular interactions that directly drive receptor activation.

Following such findings, studies on GPCR dimerization have been appearing at a steadily increasing pace (reviewed in refs 31−40). Although their existence is now largely accepted, their functional importance remains more enigmatic and in some cases even controversial (reviewed in refs 31−41). These conclusions can also be inferred from the recent review article by Terrillon and Bouvier, concerning the state of the art in our understanding of the role of dimerization of GPCRs in the five different steps of their half-life cycle, i.e., ontogeny, ligand-promoted regulation, pharmacological diversity, signal transduction, and internalization.[38]

As for ontogeny, GPCRs of family C were the most generous so far in yielding clues on the possible role of oligomerization. $GABABR_1$ and $GABABR_2$ receptors represent an emblematic case. Each of the two genes makes an incompetent GPCR protein, as the first is not properly glycosylated or inserted in the membrane, and the second has no signaling activity. However, when both are expressed in the same cells, which physiologically occurs in brain neurons,[532,533] they make a functional GABA responsive heterodimer.[532−535] The structural requirement in this case was identified in the C-terminus, where there are recognized sequences prone to make coiled-coil interactions.[536] That is in contrast with another member of family C receptors, the glutamate receptor type 1, where dimerization occurs in the N-terminus.[537] This large "flytrap" domain has structural similarity to the glutamate binding domain of glutamate receptor channels, where they, in fact, form stable oligomeric forms.[4] Although the examples above seem "special" cases, evidence that GPCRs dimerize in order to act as molecular chaperones, i.e., to catalyze their own folding and transport to the cell surface, was also found for family A GPCRs, such as the V_2 vasopressin receptors[538,539] and the chemokine receptor CCR5.[540] See also ref 41 for a review.

A role for dimerization in GPCR ontogeny does not exclude the possibility that, once the receptor has reached the cell surface, its oligomeric state could be dynamically regulated by ligand. Experiments based on FRET and bioluminescence energy transfer (BRET) reveal that many GPCRs exist as oligomers, or at least as closely packed clusters, in the membranes of living cells.[31−38] However, as to the question of functional relevance, they have contributed very little insight, so far. In fact, since the majority of such studies show that agonist binding does not significantly change the optical signal, it appears that GPCRs are constitutive dimers and that this supramolecular organization is not perturbed by the state of activation of the protein. Thus, according to a number of FRET and BRET determinations, neither enhancement nor disruption of the dimeric state seems to be necessary for receptor-mediated activation of G protein. Yet, that may be a glaring case in which nothing tells something, because, by the same token, there is no available evidence to refute the possibility that a dimeric GPCR is an obligatory structure for engaging a productive interaction with the G protein heterotrimer. Indeed, structural and biochemical evidences suggest that the binding surface of the G protein trimer can accommodate the twin footprints of a dimeric GPCR,[541,542] although the same observation can be interpreted as an indication that the cytosolic region of a single GPCR can undergo large conformational changes upon interaction with the G protein.[543] Very recently, elegant experiments consisting of a combination of mass spectrometry after chemical cross-linking and neutron scattering in solution have unambiguously established that the complex formed between the purified, activated leukotriene B_4 receptor BLT1 and $Gi_{\alpha 2\beta 1\gamma 2}$ corresponds to a pentameric assembly of one dimeric receptor and one heterotrimeric G protein.[544]

Oligomerization may be a strategy to diversify and extend the signaling properties that are intrinsic in each individual receptor gene (reviewed also in refs 541 and 545). Opioid receptors have been extensively studied in this regard.[546−548] It was shown that coexpressed δ and κ receptors can generate a different binding pattern and synergistic effects on MAP-kinase.[549] Similarly, mixes of μ and δ receptors gain special signaling and binding properties if jointly expressed.[550] Thus, it is possible that the complex combination of opioid receptor subtypes generated by studies in the precloning era and never matched by the genes that were actually found is explained by the fact that such receptors differ in signaling properties when they exist in homomeric or heteromeric form.[546,547] Similar results have been presented for many other receptor types and suggest a general trend.[31,34,38] Supramolecular arrays containing mixtures of different receptors may constitute specialized signaling patches of the plasma membrane, and their differential distribution in cell regions may have fundamental roles in fine-tuning the complex signaling networks of the central nervous system. This phenomenon, if true, poses a daunting challenge but also new opportunities in the design of new drugs. One signaling pathway, in which oligomerization may be crucial player, is G protein-independent signaling of GPCRs to mitogen-activated protein

kinases. Arrestins, a family of proteins that were previously thought to be exclusively devoted to disconnect receptors from G proteins,[30] are now known to act also as recruiting adaptors that divert GPCR signaling toward intermediate kinases such as ERK1, ERK2, and others.[27,28] Recent AFM experiments on rhodopsin indicate that the photoreceptor forms dimers in the plasma membrane, with cytosolic protrusions that are located 3.7 nm apart.[43] This distance matches that of two evident grooves that are present in the crystal structure of arrestin.[43,90] This would suggest that GPCRs must be in dimeric form to bind arrestin. Very recent in vitro experiments, based on complementation of individually nonfunctional GPCRs, support this hypothesis, providing evidence that binding of β-arrestin-1 to M_3-muscarinic receptors requires paired stimulation of two receptor components within the same receptor dimer.[40,551,552] The consequences of arrestin-mediated signaling in vivo are not entirely understood yet, but studies of opioid effects in mice bearing deletions of arrestin genes suggest that arrestin signaling may be important for the control of tolerance and dependence.[553,554]

Several recent studies have suggested that heterodimerization could affect agonist-promoted GPCR endocytosis, a well-characterized process classically involved in signal attenuation. For many documented heterodimers, stimulation of only one of the protomers was sufficient to promote cointernalization of the two receptors (reviewed in ref 38). By contrast, receptors that do not undergo efficient agonist-promoted endocytosis were found to act as dominant negatives for endocytosis-prone receptors after heterodimerization (reviewed in ref 38). Although of significant potential interest, the physiological consequences of these observations on the regulation of GPCR desensitization/resensitization cycles remain to be determined.

The extremely numerous experimental evidences for GPCR dimerization/oligomerization, however, provided very little insight into the architecture of the supramolecular receptor assemblies (reviewed in refs 31−38 and 541). An elegant study based on a combination of in vitro biochemical and biophysical experiments on the glycoprotein hormone receptors provides evidence that these receptors form homo- and heterodimers via interactions involving primarily the heptahelical domains and that hormone binding occurs with a strong negative cooperativity.[439] For family A GPCRs, recurrent evidence for the involvement of H6 in the intermonomer interface came from early experiments with synthetic peptides holding the sequence of such a helix, which appeared to inhibit homodimerization of either β_2-AR or D_2 receptors.[555,556] For the D_2R, a peptide from the H7 sequence proved to exert an inhibitory effect as well.[556] The involvement of H5 and H6 from D_2 in the heterodimerization with the A_{2A} adenosine receptor ($A_{2A}R$) was indicated by BRET experiments using a $D_{2[1-4,7]}R$ chimera, containing H5, H6, I3, and E3 from the D_1R sequence, a receptor that does not dimerize with $A_{2A}R$.[364] In fact, in contrast to the wild type D_2R, the $D_{2[1-4,7]}R$ chimera was not able to compete for the specific BRET between $A_{2A}R$ and D_2R.[364] Recent experiments on purified leukotriene B_4 receptor provided evidence for the central role of H6 in stabilizing the receptor homodimer.[544] Also, H4 has been suggested to mediate intermonomer contacts on the basis of the results of cysteine cross-linking experiments on D_2R.[557] Interactions between H1 and H4 were suggested to mediate the homodimerization of the CCR5 receptor, on the basis of the experimental observation that the combination of two point mutations, i.e., I52(1.54)V and V150(4.47)A, impaired receptor function and prevented FRET, differently from the case of the wild type.[558] This hypothesis has recently been challenged by Lemay et al.[559] Furthermore, FRET experiments using either various α_{1b}-AR fragments or α_{1b}/β_2 chimeras suggested that, in the whole α_{1b}-AR, H1 and H7, which are adjacent in the helix bundle, act in concert in favoring receptor homo-oligomerization, with H1 being the prominent interface.[560] Curiously, in vitro experiments on the M_3-muscarinic receptor indicated that C140(3.25) and C220 in E2, homologous to the amino acids that in the rhodopsin structure are involved in an intramolecular disulfide bond, can also participate in the formation of intermolecular disulfide bonds in a dimer.[561] This was inferred from the observation that C140(3.25)A and C220A mutations prevented the formation of disulfide-linked receptor aggregates.[561]

5.2. Computational Modeling of GPCR Dimerization/Oligomerization

Gouldson and co-workers have been pioneers in studying GPCR dimerization by using computational modeling and bioinformatics tools (reviewed also in refs 217 and 562).[506,563] They initially proposed a mechanism of receptor activation involving domain swapping, essentially supported by the results by Maggio and co-workers on chimeric M_3-muscarinic and α_2-adrenergic receptors.[530] MD simulations were used to analyze the proposed mechanism of dimer formation.[563] Computations were carried out on a β_2-AR model constructed on the basis of the 3D electron density map of rhodopsin.[12] Three possible dimer arrangements were investigated: a H1−H2 dimer, a H1−H7 dimer, and a H5−H6 domain-swapped dimer. A single ligand was docked into half of the receptor dimer, and the complexes underwent energy minimization and a MD simulation of up to 450 ps. The potential energy of these complexes, plotted against simulation time, revealed that both the apo H1−H2 dimer and a H1−H7 dimer were significantly lower in energy, which was also the case when an antagonist was present. However, when an agonist was docked, the energy of the H5−H6 dimer was significantly lowered relative to those of the other structures. This was thought to be consistent with the idea that agonist-induced activation is caused by a shift in the equilibrium toward the H5−H6 dimer. Since simulations were done on a dimer model in the absence of loops, the H5−H6 contact dimer and the H5−H6 domain-swapped dimer were identical (reviewed also in ref 217).[506,563]

To provide support to the inferences of their simplified molecular models, the authors integrated

the results of MD with those of sequence analyses, using both the correlated mutation (CM)[564] and evolutionary trace (ET)[565,566] methods.[541,563] In particular, they examined the occurrence of both correlated mutations and class-conserved residues. The CM method has been shown to provide information about interdomain contacts.[564] The correlation is interpreted as a result of the tendency of positions in proteins to mutate in a coordinated manner if the interface has to be preserved for structural or functional reasons. Thus, sequence changes occurring during evolution at the interface of dimerization of a given monomer A would be compensated by changes in the interacting monomer B, to preserve the interaction interface. The ET method is another approach to determining functional sites for a protein, given its 3D structure and a multiple sequence alignment.[565,566] It bears some similarities to correlated mutation analysis, as the evolutionary trace residues may also be correlated, but it has the advantage that conserved residues are also included in the analysis.[565,566] The basic assumptions of the ET method are as follows: (a) that within a multiple sequence alignment the protein family retains its fold, on the basis of the idea that proteins that have evolved from a common ancestor will show similar backbone structure;[367] (b) that the location of the functional sites is conserved; (c) that these sites have distinctly lower mutation rates because of the evolutionary pressures of residues defining the functional sites; and (d) that this lower mutation rate is punctuated by mutation events that cause divergence.[565]

For the adrenergic receptors, correlated mutation analysis on 50 aligned sequences showed that the correlated mutations do accumulate at the H5−H6 interface.[563] The remaining external correlated residues on H1, H2, and H7 were thought to be involved in the formation of a H1−H7 dimer intermediate or in the formation of higher order oligomers.[563] The ET method was applied to over 700 aligned GPCR sequences.[541] The method predicted the occurrence of functionally important clusters of residues on the external faces of H5 and H6 for each family or subfamily of receptors; similar clusters were observed on H2 and H3. The probability that these clusters are not random was determined using MC techniques. The cluster on H5 and H6 is consistent with both H5−H6 contact- and H5−H6 domain-swapped dimer formation. The observation of functionally important clusters of residues on H2 and H3 raised some possible interpretations, including heterodimerization and oligomerization.[541] In a more recent report, being aware of some experimental evidence against domain swapping as a general mechanism for receptor dimerization,[227,567] and conscious that their computational approaches were unable to distinguish between contact- and domain-swapped dimers, the authors proposed that domain-swapped and contact dimers are essentially equivalent in their ability to signal, and this could underlie any failure to observe domain swapping. The role of domain swapping in GPCR dimerization still remains an open question, as it is apparently supported also by recent in vitro evidences,[568] although it remains inconsistent with oligomerization. The same team, very recently, combined the ET method with the entropy method to improve predictions.[569] They, indeed, recognized that the CM analysis, while being mathematically well-defined, had the limit to predict relatively few amino acids, inconsistent with the expected extensions of protein−protein interfaces.[569] On the other hand, the ET method, although able to predict protein−protein contact areas, was, however, subjective. While earlier work tended to favor H6 as the most likely dimerization interface, the novel analyses favored H4.[569]

A modification of the CM approach, the so-called "subtractive correlated mutation method", has been used to predict homo- and heterodimer interfaces in the opioid subfamily of GPCRs.[372,570] Application of the method to δ and μ opioid receptors showed that most of the correlated residues of the δ opioid receptor are located on the outer (lipid-facing) surface of H4, H5, and H6, whereas, in the μ opioid receptor, H1 is the helix that is likely to be involved in the heterodimerization with δ. Since H4 and H6 cannot participate simultaneously in the intermonomer interface, these results were thought consistent with higher order oligomers. The same study applied to homodimerization of the δ, κ, and μ opioid subtypes suggested that (a) H4 and/or H5 would participate in the formation of the δ-homodimers; that (b) H5 would participate in the formation of κ-homodimers; and that (c) H1 would participate in homodimerization of the μ-subtype.[372] The consistency with the results of in vitro experiments, which implicate disulfide bonds in κ homodimerization and in $\delta-\mu$ heterodimerization[549,571] as well as the need of the intact C-tail for δ homodimerization,[572] will be the challenge for further computational investigations.

The combination of CMA and automatic methods, such as the level entropy and the sequence space automatization methods, aimed at detecting amino acid positions that could have some functional significance for the whole family and, at the same time, are specific for each subfamily (i.e. tree-determinant positions),[573] was used to predict the interface in CCR5 homodimers.[558] With this approach, positions 1.54 and 4.47 were predicted to participate in the interface.[558]

A hidden-site class model of evolution, which employs different substitution matrixes to represent substitutions in different parts of the protein, was used to predict possible dimerization interfaces in aminergic GPCRs.[574] The approach predicted the involvement of H5 and H6 in most aminergic subfamilies and H4 and H5 in the muscarinic and opsin sub-families.[574]

Predictions of oligomerization interfaces in a number of family A GPCRs belonging to the opsin, dopamine, adrenergic, and muscarinic acetylcholine families were also done by a multistep method made of two convergent sets of in silico experiments, one sequence-based and the other structure-based.[575] The sequence-based set of experiments consists of multiple sequence alignment and search for conservation patterns. The structure-based set consists of reducing the 3D coordinates of the rhodopsin structure to a

bidimensional plane and individuation of the residues on the molecular surface of the rhodopsin monomer, in particular, the lipid-facing ones. The integration of the two sets of experiments leads to 2D maps of conserved lipid-exposed residue clusters, which are interpreted as potentially involved in the dimerization interface. The predicted interfaces differed among subtypes. The following domains were predicted as implicated in homodimerization of selected receptor subtypes: (a) H4, H5, and I2, for rhodopsin, (b) H4, for the D_2R, (c) H6, for the β_2-AR, and (d) amino acids from the cytosolic extension of H3 as well as from I2, H4, and I3, for the M_3-muscarinic receptor.[575]

Collectively, the different sequence-based methods found consensus in predicting H4, H5, and H6 as the domains most likely involved in GPCR oligomerization (reviewed also in ref 576). Despite this general consensus, it is clear that different sequence analysis methods identify different key positions in different receptor subfamilies; this is partly because each subfamily is different, but the differences may also arise because of insufficient sequence data.[569,576] Furthermore, predictions by sequence-based methods have low resolution, as interatomic interactions are not explicitly considered.[569]

About seven years ago, we employed a computational approach based upon rigid-body docking to simulate the homodimerization of mutation- and ligand-induced active states of the α_{1b}-AR.[303] Simulations were carried out on ab initio models of the receptor. The best resulting dimer structures for the D142(3.49)A active mutant and the epinephrine-bound receptor were found to involve H5, H6, and H7.[303]

Very recently, we have developed a computational procedure for predicting the supramolecular organization of TM α-helical proteins.[577] The approach consists of rigid-body docking by means of the ZDOCK program.[578] The best solutions selected by the docking program(s) as the best in terms of shape complementarity are then subjected to an "in-house" made filter, i.e., the "membrane topology" filter, which discharges all the solutions that violate the membrane topology requirements. In detail, the filter discards all the solutions characterized by a deviation angle from the original *z*-axis, i.e., tilt angle, and a displacement of the geometrical center along the *z*-axis, i.e., *z*-offset, above defined threshold values. For the tilt angle and the *z*-offset, thresholds of 0.4 rad and 6.0 Å were, respectively, employed. Discarded solutions generally constitute more than 94% of the solutions selected by the docking program, thus improving the effectiveness of the following cluster analysis. The strength of the approach stands in its independence from the size of the system, symmetry information, and extension of the water-soluble domains.[577] Benchmarks of the approach, done on the tetrameric potassium channel (Kch, 384 amino acids),[579] on the pentameric MscL (540 amino acids)[580] and eptameric MscS (1771 amino acids)[581] mechanosensitive channels, and on trimeric bacteriorhodopsin (698 amino acids),[582] in all the test cases led to nativelike quaternary structures, i.e., with C_α-RMSDs lower than 2.5 Å from the native oligomer.[577]

The effectiveness of the prediction protocol makes it suitable for ab initio quaternary structure predictions of other integral membrane proteins, including GPCRs. An attempt in this respect has already been reported, though based on an early and different version of the computational protocol.[364] In detail, integrating the rigid-body docking approach with the results of protein engineering and FRET and BRET experiments provided insights into the putative interaction interface of $D_2R-A_{2A}R$ heterodimers.[364] The initial models of the two receptors were achieved by comparative modeling, using modified rhodopsin structures as templates. The whole sequences of both receptors were modeled, since dimerization and/or oligomerization might involve also the cytosolic and/or the extracellular domains, as recently suggested for rhodopsin.[90] Since most of the structural errors are expected to reside in the intracellular and extracellular domains, nine average minimized structures of the $A_{2A}R$, differing in the conformations of the intracellular and extracellular loops as well as in the topology of the huge C-tail, were used to probe the effect of such structural differences on the results of docking simulations. Each of these structures was docked with the selected average minimized structure of D_2R. The different $A_{2A}R$ structures share preferential docking modes at the D_2R. These docking modes were broadly grouped into two clusters, CLUSTER 1 and CLUSTER 2 (Figure 14). In particular, in the most populated cluster, CLUSTER 1, H5 and/or H6 and the N-terminal portion of I3, from D_2R, respectively, approach H4 and the C-terminal portion of the C-tail, from the $A_{2A}R$ (Figure 14, left). H7 of D_2R may also participate, together with H6, in the contacts with H4 of $A_{2A}R$. A very short but significant portion of the huge I3 of D_2R, i.e., the N-terminal 217−220 amino acid stretch that is made of four consecutive arginines (^{217}RRRR^{220}), is frequently involved in the heterodimer interface (Figure 14, left).[364] Some of the four cationic amino acids are frequently found interacting with D401 and/or D402 in the C-terminal portion of the $A_{2A}R$ C-tail. A few more amino acids from the $A_{2A}R$ C-tail are suggested to participate in the heterodimer interface. Thus, very limited and almost invariant portions of the D_2R I3 and of the $A_{2A}R$ C-tail are likely to mediate $D_2R-A_{2A}R$ contacts. The heterodimer architecture shared by the members of CLUSTER 1 found consistency with the results of BRET experiments using a D_2R/D_1R chimera, which implicated the H5−I3−H6 portion of D_2R in the interaction with $A_{2A}R$.[364] The predicted interface according to the members of CLUSTER 1 was also consistent with the results of pull-down and mass spectrometry experiments, which suggested that $A_{2A}R-D_2R$ heteromerization depends on an electrostatic interaction between an arginine-rich epitope from the I3 of the D_2R (^{217}RRRRKR^{222}) and two adjacent aspartates (D401 and D402) or a phosphorylated Ser (S374) residue in the C-tail of the $A_{2A}R$.[583]

The second cluster of docking solutions, CLUSTER 2, less populated than CLUSTER 1 but characterized by high docking scores, resembled the intradimer contact model proposed for rhodopsin (Figure 14, right).[90] In this cluster, the heterodimer interface is

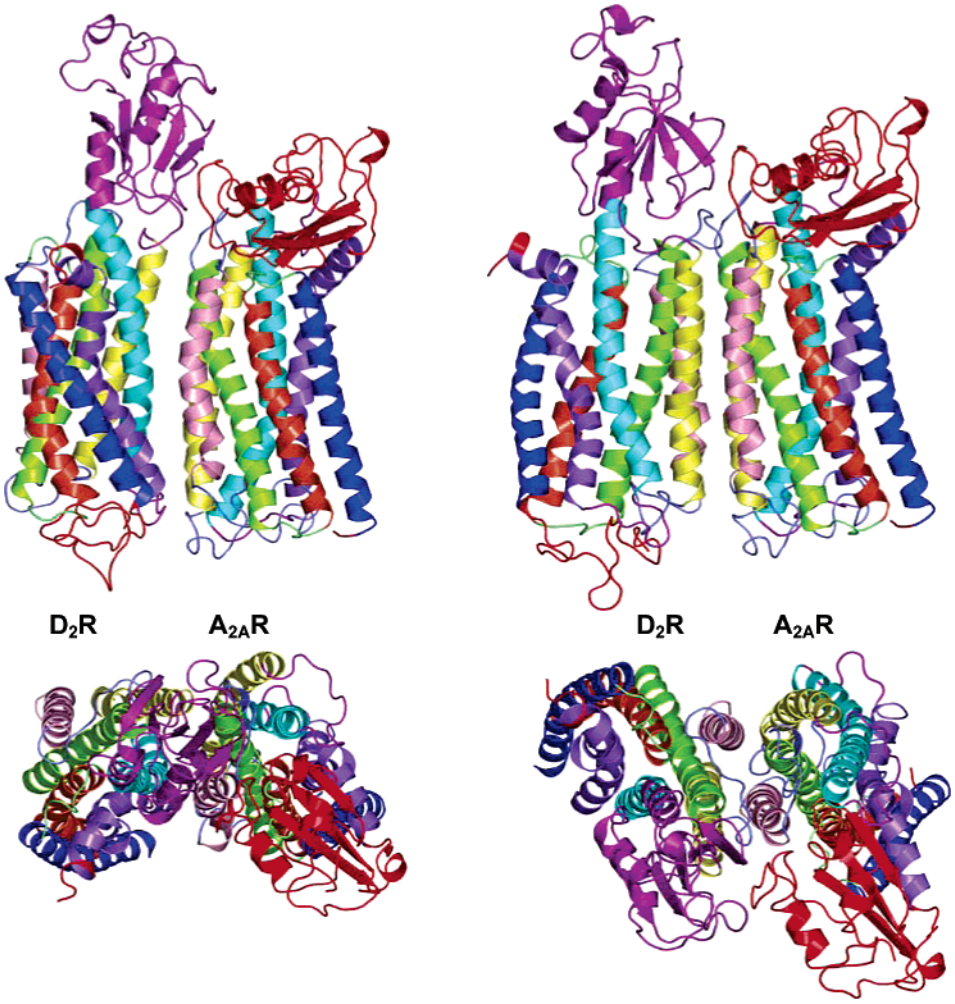


Figure 14. Examples of the $D_2R-A_{2A}R$ heterodimers belonging to CLUSTER 1 (top and bottom left) and to CLUSTER 2 (top and bottom right).[364] In the top views, the helix bundles are seen in a direction parallel to the membrane surface, with the intracellular side being at the top. In the bottom views, the helix bundles are seen from the intracellular side in a direction almost perpendicular to the membrane surface. H1, H2, H3, H4, H5, H6, and H7 are respectively colored in blue, orange, yellow-green, pink, yellow, cyan, and violet. The amino acid stretch corresponding to rhodopsin's H8 is colored in violet as well. The N- and C-termini are in red, I1 and E1 are in lime, I2 and E2 are in slate, and I3 and E3 are in purple. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

mainly formed by I2, H4, H3, and H5, from D_2R, and I2, H5, H3, and H4, from $A_{2A}R$. In these dimers, the extracellular end of the interface is made of contacts between aromatic amino acids from E2 and H5. These features could have functional implications, as the extracellular end of H5 is involved in agonist binding. The role of agonist-induced activation on the homo- and heterodimerization of these receptors is still obscure. Recent experimental evidences suggest that it may have a role in the formation of higher order oligomers, rather than in the formation of dimers, which should be constitutive features of the receptors.[364,584]

Collectively, the results of simulations of $D_2R-A_{2A}R$ heterodimerization showed a very limited involvement of the intracellular and extracellular domains in the intermonomer interfaces, thus overemphasizing the role of the TM helices. These results support the reliability of predictions, as the models of the TM helix bundles are expected to be acceptably accurate.

We are extensively challenging the computational approach in quaternary structure predictions of a number of GPCRs, including members of the amine, peptide, and hormone[585] subfamilies.

Although sequence-analysis-based approaches can be sources of invaluable information, the potential of molecular simulation methods is significantly higher and is going to increase with improvements in the quality of the GPCR models. Given the approximations in the current protein−protein docking algorithms, an integration between sequence- and structure-based approaches and in vitro experiments is, however, expected to improve the accuracy of quaternary structure predictions.

While protein−protein docking simulations are essential to gain insight, at the atomic level, into the architecture of the intermonomer interfaces, a mesoscopic MC simulation approach, such as that proposed by Woolf and Linderman, might create a more direct link with in vitro experiments on GPCR oligomerization and localization in selected membrane compartments.[586,587] By incorporating information from in vitro experiments, the computational approach by Woolf and Linderman has begun to address the problem of how the kinetics of protein−

protein binding affects the overall organization of the membrane and how dimerization affects the global protein organization.[586,587] Preliminary results by this approach suggest that changes in dimerization interactions affect cellular physiology. The possibility of predicting the ligand's ability to induce homo- or heterodimerization is another intriguing potential of this approach in drug discovery.

The integration between atomistic and mesoscopic simulations is expected to be a promising tool to unravel functioning mechanisms that involve intricate protein networks.

6. Receptor–G Protein Interaction

6.1. Insights from in Vitro Experiments

G proteins consist of three subunits α, β, and γ, forming one of nature's most important miniature (nano-) machines (reviewed in refs 21 and 588). The α-subunits are enzymes of the Ras superfamily, which hydrolyze GTP to regulate events within cells and to transduce external signals. In the inactive state, G proteins form membrane-associated $\alpha\beta\gamma$ heterotrimers, with GDP tightly bound to the α-subunit. Upon activation by extracellular signals, receptors catalyze the exchange of bound GDP for GTP. The GTP-bound form of the heterotrimer is unstable and heterolytically dissociates to form active GTP-α and $\beta\gamma$ complexes (reviewed in refs 589 and 590). Acting either coordinately or independently, these two species bind and modulate the activities of downstream effector molecules. G proteins are released from effectors upon the breakdown of GTP that results from the slow GTPase activity of the α-subunit. The inactive α-subunit can recombine with $\beta\gamma$, reforming the heterotrimer, which can then reassociate with its receptor and undergo a new cycle of signal transduction. Recent experimental evidences, however, indicate that G protein activation may not be concurrent with dissociation of α from $\beta\gamma$.[591,592] The family of heterotrimeric G proteins includes over 20 isoforms from four classes of α (Gs, Gi, Gq, and G12), five of β, and at least six of γ (reviewed in refs 21, 589, and 590).

Crystallographic studies of G protein α-subunits and heterotrimers provided significant insight into our understanding of how these extraordinary nanomachines might work (reviewed in refs 543 and 589). Structural studies of α-subunits have focused on Gt_{α}, transducin involved in vertebrate vision (reviewed in refs 589 and 593), and $Gi_{\alpha 1}$ (reviewed in refs 589 and 593) and Gs_{α},[594] respectively, involved in hormone-regulated inhibition and activation of adenylate cyclase. The structures of the GDP-bound forms of the $Gi_{\alpha 1\beta 1\gamma 2}$ and of the $Gt_{\alpha\beta 1\gamma 1}$ heterotrimers provided the first view of the β-subunit in complex with γ.[595,596] The α-subunit consists of two domains, the GTPase domain, which contains a six-stranded β-sheet surrounded by six α-helices, and the helical domain, constituted by a long central helix surrounded by five shorter helices. GDP is bound into a cleft between the GTPase and helical domains (Figure 15). Both domains have almost identical structures in the GTP- and GDP-bound states. Significant

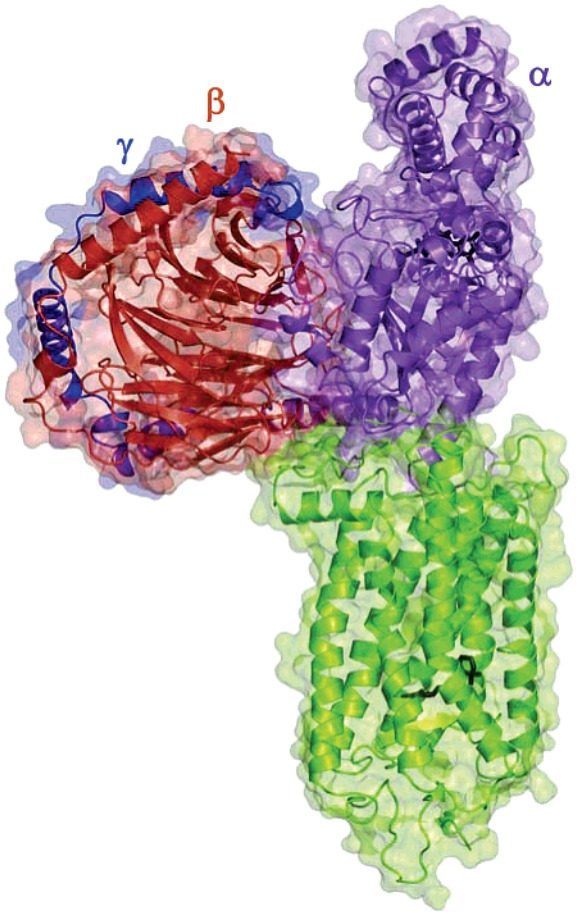


Figure 15. Computational model of the complex between monomeric dark rhodopsin, 1U19, and heterotrimeric Gt.[641] The rhodopsin molecule is colored in yellow-green, whereas the α-, β-, and γ-subunits of the Gt are colored in violet, orange, and blue, respectively. The retinal and GDP molecules are represented by black sticks. The molecular surface is also shown colored according to the protein chains. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

changes are observed within the GTPase domain contacting $Gt_{\beta\gamma}$; in fact, these regions are disordered in the inactive heterotrimeric forms, whereas they are ordered in the Mg^{2+}-GTPγS-activated structures of Gt_{α} and $Gi_{\alpha 1}$ (reviewed in ref 589). The N-terminal region of the α-subunit consists of a long α-helix pointing out from the rest of the subunit (Figures 15 and 16). This structural feature was revealed by the structures of heterotrimeric Gt and Gi, as the α-helical conformation of the N-term is stabilized by the $\beta\gamma$ complex, whereas such a domain is disordered in the isolated α-subunits.[589,597] The last 10 amino acids of the Gt_{α} are predicted to hold an α-helical conformation in the MII-bound forms, whereas this short segment appears to be disordered in the receptor-dissociated forms of the α-subunit.[589,598,599] The β-subunit, a member of the WD family, has a long N-terminal helix followed by a repeating module of seven β-sheets, each with four antiparallel strands, forming a β-propeller structure (Figures 15 and 16).[589,593,595,596] The γ-subunit contains two helices: the N-terminal helix interacts with the N-terminal helix of β, whereas the remaining polypeptide chain of γ interacts with the β-propeller structure of β.[589,593,595,596] Similarly to the C-tail of the α-subunit,

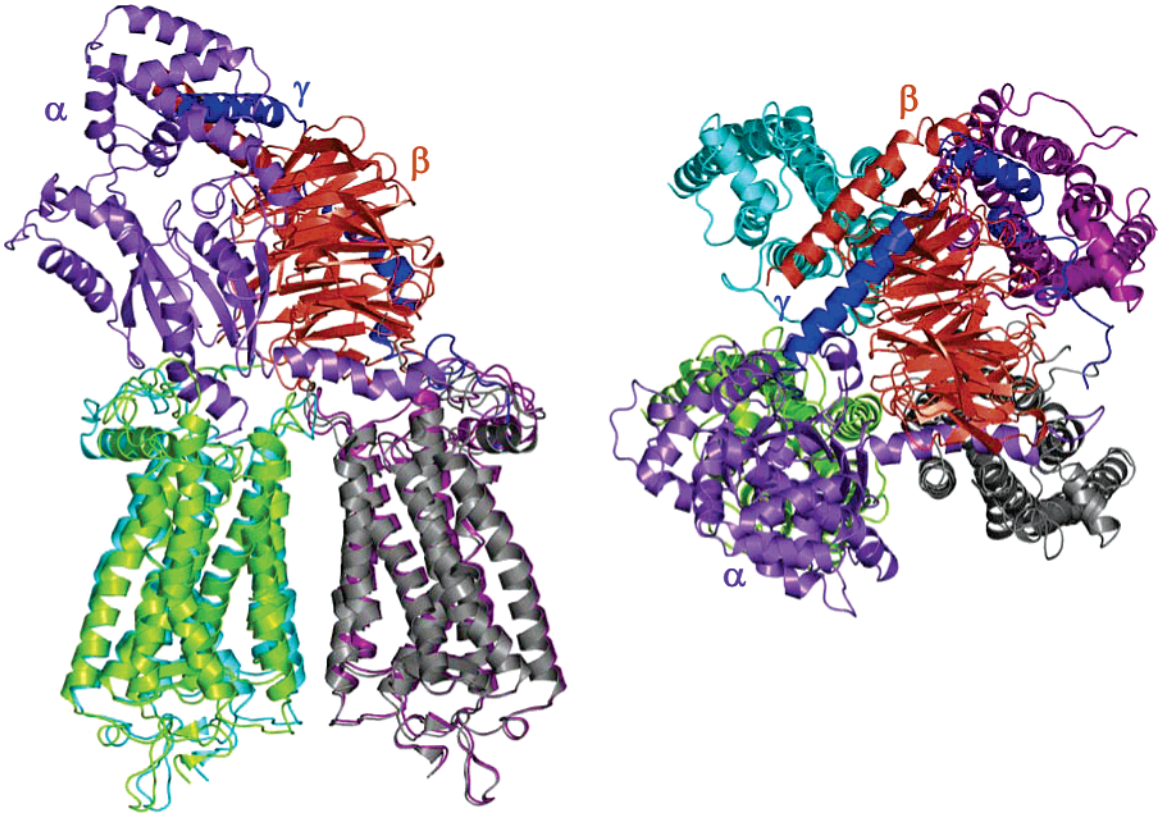


Figure 16. Two side views of the computational model of the complex between a rhodopsin tetramer and heterotrimeric Gt.[542] The four rhodopsin monomers are colored in yellow-green, gray, purple, and cyan, whereas the α-, β-, and γ-subunits of the Gt are colored in violet, orange, and blue, respectively. Drawings were done by means of the software PYMOL 0.97 (http://pymol.sourceforge.net/).

the C-tail of the γ chain (i.e. the (60−71)farnesyl peptide) holds a regular α-helical structure when bound to MII, whereas its conformation is disordered in the receptor-dissociated forms of the βγ complex.[600] Thus, in vitro experiments suggest that activated rhodopsin controls the conformation of the C-tails of the G protein α- and γ-subunits.[598−600]

The gross experimental information on receptor-G protein interaction concern the rhodopsin−transducin system (reviewed in refs 543 and 601).

Early experimental results on rhodopsin suggested that I2 and I3 of rhodopsin are involved in binding and/or activation of transducin, whereas the peptide corresponding to I1 of rhodopsin does not compete with MII for binding to Gt.[602,603] The important role of I2 and I3 in G protein recognition has also been demonstrated by several experiments on different GPCRs.[604−606] Experimental evidences suggest that the amino acids C140(3.55), K141, R147, and F148, at the end of H3 as well as at the N- and C-termini of I2, as well as residues T229, V230, S240, T242-(6.25), T243(6.26), and Q244(6.27), in I3 and the cytosolic extension of H6, participate in the interaction with Gt.[607] The loop regions proximal to the TM helices appear to be more important for the MII-Gt_α coupling than regions in the middle. The cytosolic extension of H3, including the conserved E/DRY motif, and the N-terminal end of I2 have been widely implicated in Gt_α activation.[608−610] The cytosolic extension of H3 and the N-terminus of I2, in concert with the C-terminus of I3 and the cytosolic extension of H6, are proposed to be involved in the recognition of the Gt_α C-terminus.[609] In a recent study, the I3 of rhodopsin was demonstrated as one of the regions responsible for the specific coupling with the Gt. It has also been inferred that the 6-amino acid sequence adjacent to the C-terminal 5-amino acids of Gt_α interacts with I3 of MII.[611] Recent experiments demonstrated that the role of H6 movements during MII formation is to provide a binding site on the cytoplasmic face of rhodopsin for the Gt_α C-terminus. This movement appears to open a cleft and expose a hydrophobic patch, which directly interacts with the Gt_α C-terminus and increases the affinity for transducin. These results also suggest that the Gt_α C-terminus binding is specific for the MII state and that the presence of the retinal Schiff base linkage is required to maintain the exposure of the cleft required for interaction with such domains of transducin. The amino acids L226(5.61), T229, and V230, at the boundaries of H5 and I3, which lie in the solvent-exposed cleft, are suggested to play a key role by imparting high affinity binding for the transducin Gt_α C-terminus.[127]

Experimental evidences suggest that different G protein portions participate in the receptor−G protein interface. It has been, in fact, demonstrated that, to stabilize the high affinity state of the A1 adenosine receptor, a multiple interaction between the receptor and the G_α or even the heterotrimeric G protein is required.[612] The crucial role of the C-terminal amino acids in coupling to the receptor was widely demonstrated for different α-subunits.[590,598,599,613−619] A C-terminal 11-amino acid peptide from transducin, Gt_α-

(340−350), has been shown to both bind and stabilize the MII conformation, mimicking heterotrimeric Gt. Indeed, using a combinatorial library, analogues of Gt_α-(340−350) were identified that bound light-activated rhodopsin with high affinity.[543,620,621] The same authors made peptides with key substitutions either on the background of the native Gt_α-(340−350) sequence or on the high affinity sequences and used the stabilization of MII as a tool to determine which amino acids are critical in the G protein−rhodopsin interaction.[621] Removal of the positive charge at the N-termini by acylation, or delocalization of the charge by K to R substitution enhanced the affinity of the Gt_α-(340−350) peptides for MII, whereas a decrease in affinity was observed following C-terminal amidation.[621] C347, a residue conserved in pertussis toxin-sensitive G proteins, was shown to interact with a hydrophobic site in MII.[621]

The role of the $\alpha4/\beta6$ loop of the α-subunit in receptor−G protein interaction is supported by the results of proteolytic digestion experiments and alanine scanning mutagenesis on transducin as well as by biochemical studies on $Gt_\alpha/Gi_{\alpha1}$ chimerae.[622−625]

More resolved information came from mutational and cross-linking studies aimed at identifying residues in Gt_α and rhodopsin that are in close proximity. Acharya et al. identified Y136(3.51)−V139(3.54) in H3 of rhodopsin as interaction sites with Gt_α.[609] Experimental evidences from Khorana's group seem to indicate that L19−R28, R310−K313, and E342−K345 of Gt_α are cross-linked to S240C in I3 of MII.[626,627] Evidence for interactions between the N310−Q312 region (H8) of rhodopsin and residues 340−350 of Gt_α has been also reported.[628,629]

Using stabilization and photoregeneration of the receptor's signaling state and Gt activation assays, Herrmann and co-workers provided evidence for a two-site sequential fit mechanism of Gt activation. According to this model, receptor−G protein recognition is suggested to be initiated by an encounter of Gt_γ(50−71)farnesyl with MII. This would make the C-terminal tail of the Gt_α available for binding with MII, triggering the GDP release and the formation of a stable empty site complex that is ready to receive the activating cofactor, GTP.[630]

In summary, the patchwork of the most relevant information from in vitro experiments on rhodopsin−transducin recognition suggests that the $\alpha4/\beta6$ loop and the C-terminus of Gt_α recognize a solvent-accessible cleft on MII, formed by amino acids from the cytosolic extensions of H3, H5, and H6, from the N-terminus of I2, and from the N-terminus of H8.

Very recent in vitro experiments provide evidence for possible interactions between inactivated rhodopsin and Gt_α, thus suggesting that the cytosolic domains in the crystal structures of inactive rhodopsin can recognize transducin and that the accomplishment of the MII state would require a precoupling between dark rhodopsin and transducin.[24,631] However, experimental evidence that dark rhodopsin and heterotrimeric transducin may exist as a preformed complex has appeared early in the literature, athough it was not pursued any further.[632] This hypothesis is also supported by the experimental evidences that proton uptake from the cytosol, which accompanies MII formation, would require the presence of transducin to occur.[106] The hypothesis of receptor−G protein precoupling may be a common feature to all GPCRs, as also demonstrated by recent PWR spectroscopy observations of a binding between a ligand-free δ opioid receptor ant the G protein.[23]

6.2. Bioinformatics and Computational Modeling Approaches to Predictions of the Receptor−G Protein Interface

Most of the computational approaches aimed at predicting the receptor−G protein interface in the years that preceded the release of the first crystal structure of rhodopsin were essentially based on sequence analyses approaches such as the ET method.[565,633] Mapping the results of the ET method into the van der Waals surface of the G_α led to identifying residues in $\beta4$, the $\alpha4/\beta6$ loop, $\beta6$, α5, and the C-terminus as potential recognition points for the receptor. Many of the residues identified have also been implicated by in vitro experiments.[565,624,633]

A successive reinvestigation, through the ET method, of the potential receptor contact sites on the G protein, in light of emerging evidence for GPCR dimerization, predicted an ET functional site approximately twice as big as was originally reported by Lichtarge et al.[541,633] The ET functional sites predicted by Dean et al. involved both the RAS and the α-helical domain of the α-subunit and were estimated to be large enough to interact with a GPCR dimer.[541] The simultaneous interactions between the N-terminus of the α-subunit and phospholipids and between the α-helical domain of the α-subunit and the receptor were thought possible only if the reciprocal orientation of these two portions of the α-subunit is similar to that in the crystal structure of the RGS-bound form of AlF4-activated $Gi_{\alpha1}$.[634]

Other sequence analyses on the G_α chains identified 12 residues, which are fully conserved.[635] However, the majority of these were found to be involved in the GDP/GTP binding site. A correlated mutational analysis of these sequences identified residue groups that had remained conserved or mutated as a group. These groups included residues in the $\beta2/\beta3$ loop as well as in the N-terminal and α5 helices. Mapping these residues in the crystal structure showed that they clustered around a conserved negatively charged aspartate at position 337 and suitable for interacting with the receptor. It was, therefore, suggested that this conserved aspartate and the surrounding residues form the binding site for R3.50 of the E/DRY motif in the receptor, consistent with a number of in vitro deletion experiments.[635] The authors also assumed that the C-tail of the receptor interacts with the N-terminal helix and possibly with the last 10 amino acids of the α-subunit.[635]

Early computational modeling of receptor−G protein interaction was done by Mahmoudian.[636] The target of the study was the complex of the human Gs_α with the β_3-AR.[636] The crystal structure of the *E. coli* EF-tu nucleotide binding domain was used as the backbone for the Gs_α model, with its amino acid

sequence being mutated to the Gs_α sequence. This procedure produced the core of the Gs_α protein, to which a number of loops were added using the COMPOSER program.[637,638] The β_3-AR model was achieved by comparative modeling using the BRD structure as a template.[636] The receptor−G protein complex was built using interactive molecular graphics based on low resolution experimental constraints. Interface domains from the receptor involved I3 and the C-terminus, whereas the Gs_α side of the interface was made by the N-terminus and the C-terminus.[636]

We employed computational modeling to investigate the early steps of receptor−G protein recognition.[302,303] On the basis of the inference that the mutation-induced active states of the α_{1b}-AR share the opening of a solvent accessible cleft between I2 and I3, electrostatic analysis and rigid body docking simulations were carried out to identify whether the solvent exposure of peculiar receptor portions is important in receptor−G protein recognition.[302,303] The ab initio models of the free and agonist-bound forms of wild type α_{1b}-AR and of the free forms of the D142(3.49)A and A293(6.34)E constitutively active mutants of the receptor were docked with their cognate heterotrimeric $Gq_{\alpha\beta1\gamma2}$ heterotrimer using the ESCHER program.[639] Attempts to address the issue of G protein coupling selectivity were also made by docking the same inactive and active receptor forms with heterotrimeric Gs and Gi and Gt.[302] The models of the Gq_α and Gs_α were achieved by comparative modeling using the structures of heterotrimeric Gi_α (PDB entry 1GP2)[595] as a template.[302] All the incomplete termini of the $\alpha\beta\gamma$-subunits were modeled. Consistent with the experimental findings, it was found that the cytosolic crevice formed by I2, I3, and the cytosolic extensions of H5 and H6 and characterized by a large solvent accessible surface and a positive electrostatic potential participates in the receptor−G protein interface. A suggestion from this computational study was that while the cationic nature of the cytosolic portions of the receptor seems to complement an anionic surface of the Gq_α protein, only a selected number of cationic residues could be contact sites on the receptor for the G protein.[302] Thus, although the majority of cationic residues on the cytosolic surface of the α_{1b}-AR contribute to reciprocal electrostatic properties between the receptor and the Gq_α, they appear to not be directly involved in receptor−G protein interaction and/or receptor-induced G protein activation.[302] These results are in agreement with the experimental findings that mutating the majority of basic residues in the cytosolic loops and extensions of H4, H5, and H6 of the α_{1b}-AR did not impair the receptor-mediated IP response, thus demonstrating that most basic amino acids play no direct role in receptor−G protein coupling.[62] The results of computational modeling overemphasized the role of R254 and K258, in I3, as potential contact sites on the receptor for the Gq_α.[302] These predictions found validation in in vitro findings that, among all the basic residues of the cytosolic surface of the receptor, only the combined mutations of R254 and K258 totally impaired the IP response of the agonist-stimulated receptor as well as of the constitutively active mutants D142(3.49)A and A293(6.34)E.[62]

Consistent with the results of many in vitro experiments, computations suggested that the G protein solvent-exposed portions that recognize the intracellular loops of activated receptors are the N-terminal portion of α3, αG, the αG/α4 loop, α4, the α4/β6 loop, α5, and the C-terminus.[302] Docking simulations suggested also that the two constitutively active mutants D142(3.49)A and A293(6.34)E recognize different G proteins with similar selectivity orders, i.e., $Gq_\alpha \cong Gs_\alpha > Gi_\alpha \gg Gt_\alpha$.[302]

Yeagle and co-workers proposed an interaction model between their computational model of MII and heterotrimeric Gt.[138] The construction of the complex was driven by the experimental information on rhodopsin−transducin interaction, taking also into account the electrostatic complementarity between interacting domains. In the model, the N-terminal helix of the Gt_α binds to the groove that becomes more exposed on the cytosol when MII is formed. Furthermore, the C-terminus of Gt_α binds to the crevice at the bottom of the groove in the cytoplasmic surface of MII, between I2 and I3. This mode of binding places residues K340 and D341 of Gt_α close to E134(3.49) and R135(3.50) of MII, which are hypothesized to be separated from E247(6.30) upon MII formation. The authors hypothesized that the C-terminus of Gt_α may substitute, in part, for the interactions of E134(3.49) and R135(3.50) with E247(6.30).[138]

An invaluable framework to interpret the results of biophysical and biochemical experiments on rhodopsin−transducin interaction is represented by the crystal structures of rhodopsin[13,51−54] and the rhodopsin semiempirical oligomeric model recently built, based upon geometrical constraints from AFM experiments.[43,90] An interaction model between the $Gt_{\alpha\beta1\gamma1}$ heterotrimer and a tetrameric model of rhodopsin has very recently been produced by computational modeling (Figure 16).[542] In detail, after modeling of all the missing termini in heterotrimeric Gt,[596] transducin was docked to the rhodopsin dimer, where one molecule was activated by movement and rotation of H6, followed by slight movements of neighboring helices to accommodate changes in the structure.[542] The least amount of tension was achieved during docking of the C-terminal region of Gt_α along the longest primary axis of activated rhodopsin. Docking was continued until the N-terminal helix of Gt_α (parallel to the cytoplasmic surface of the rhodopsin dimer) interacted firmly with the cytoplasmic cavities of adjacent inactive rhodopsin in the dimer. Next, β- and γ-subunits were added, obtained from the crystal structure of Gt,[596] and the whole complex was optimized by energy minimization. Short molecular dynamics runs of 5−10 ps were applied to different parts of the model to ensure proper interactions between molecules in the complex. A longer molecular dynamics run (100 ps) validated that the complex is stable.[542] Despite the probable artifacts in the MII model and the extremely short time length of MD simulations, such a computational experiment resulted in an intriguing supramolecular model

comprising tetrameric rhodopsin and heterotrimeric Gt (Figure 16).[542] According to this model, the activated rhodopsin mainly interacts with the RAS-like domain of the Gt_α, with all the cytosolic domains plus the cytosolic ends of H3 and H6 being involved in this interaction. Interactions of the N-terminal region of Gt_α with the second rhodopsin molecule in the dimer mainly involve I1, I2, H6, and the C-tail of the photoreceptor (Figure 16). Gt_β interacts with all four rhodopsin monomers in the tetramer, whereas Gt_γ forms hydrogen bonds with two rhodopsin monomers (Figure 16).[542] These interactions are suggested to provide a temporary anchor and can be replaced with water when $Gt_{\beta\gamma}$ dissociates.[542] The oligomeric model of rhodopsin−transducin interaction by Filipek and co-workers, thus, suggests that the Gt heterotrimer covers two rhodopsin dimers, and when $\beta\gamma$-subunits dissociate away, another Gt heterotrimer can bind to an adjacent tetramer. After dissociation of $\beta\gamma$-subunits from the second Gt, the next Gt can bind to a neighboring rhodopsin tetramer, and so forth. Both α-subunits not only bind to adjacent rhodopsin dimers but also interact with each other when located on adjacent rhodopsin dimers, suggesting that one Gt_α can facilitate binding of a second G protein on an adjacent rhodopsin tetramer.[542] An alternative oligomeric model of activated rhodopsin (i.e. with H6/I3 moved 8 Å away from the helix bundle)[640] in complex with heterotrimeric Gt has been proposed.[94] In this model, E134(3.49) and R135(3.50) of the rhodopsin's E/DRY motif interact respectively with K345 and D346 from the C-tail of Gt_α.[94]

In a very recent study, we have challenged the excellent protein−protein docking program ZDOCK[578] in predicting the architecture of possible complexes between dark rhodopsin, in its monomeric, dimeric, and tetrameric forms, and heterotrimeric Gt.[641] In this respect, a number of structural models of dark rhodopsin, differing in the cytosolic domains,[52−54,90] and a number of Gt heterotrimers, essentially differing in the conformation of the C-tails of the α- and γ-subunits, have been probed.[578] The most convincing results have been achieved by using the most highly resolved rhodopsin structure, 1U19 .[54] The results of these computational experiments, based upon the accomplishment of shape and electrostatic complementarities, desolvation effects, and consistency with in vitro experiments, suggest that dark rhodopsin has the potential to recognize GDP-bound heterotrimeric Gt and that the conformation of the C-tail of Gt_α affects the orientation of the Gt, used as a probe versus the target rhodopsin.[641] One of the most reliable rhodopsin−Gt complexes, obtained using monomeric rhodopsin as a target, is shown in Figure 15. Since the arrangements of the cytosolic domains of dark rhodopsin and of MI seem to be almost identical, on the basis of recent evidences from electron microscopy experiments,[104] simulations of rhodopsin−transducin recognition using dark rhodopsin or MI would be equivalent. Collectively, the results of our experiments suggest that MII formation follows the precoupling between dark rhodopsin and heterotrimeric transducin.[641]Our conclusions find support in recent evidences from PWR spectroscopy[24] and in previous evidences that proton uptake from the cytoplasm, which accompanies the transition from MI to MII, would require the presence of transducin to occur.[106] A worthy result is that the most reliable complex between monomeric rhodopsin and Gt is found also when dimeric or tetrameric rhodopsin is used as a target. This suggests that the receptor monomer holds the structural determinants for G protein activation, consistent with the results of in vitro experiments.[89]

7. Conclusions and Perspectives

GPCRs constitute the largest superfamily of membrane proteins known to date and regulate any aspect of cell activity by transmitting extracellular signals inside the cell (reviewed in refs 1 and 4−7). They have an enormous physiological and biomedical relevance, being the primary site of action of many of today's life-saving drugs and the most promising targets for those to be developed in the future (reviewed in refs 1 and 4−7). Therefore, it is not surprising that there are a tremendous number of reports published so far concerning in vitro, in vivo, and computational experiments on these systems, which make it very difficult to produce a comprehensive review.

The main drawback of computational modeling of GPCRs is the lack of high resolution information, as the unique atomistic models resolved so far, the crystal structures of rhodopsin, became available only five years ago.[13] Until that event, extremely variable computational approaches, more or less integrated with the available information from in vitro experiments, have been challenged in building 3D models of these receptors. These models proved to have high interpretative and predictive potential toward low resolution in vitro experiments. A significant example of this is a constitutively active receptor mutant designed on the basis of predictions of a computational model.[45,46,57] Along the same line, computational models of GPCRs have been successful in interpreting the results of SAR (QSAR) analyses and eventually in aiding lead optimization in those cases in which highly informative and robust ligand-based SAR could be available. Such models indeed had low potential in de novo drug design, which would require a highly resolved knowledge of the target binding site.

The need for GPCR structure prediction spurred the development of different, more or less complicated ab initio approaches, which proved successful in predicting the architecture of α-helical membrane proteins in those cases in which only low resolution 2D electron density maps are available.[10,310]

The crystal structure of rhodopsin offers the opportunity to improve significantly the reliability and quality of the computational models. We believe that the ad hoc transfer of the stereochemical restraints from the rhodopsin template to the target receptor is currently the most successful approach. We stress the point that this approach has to be customized ad hoc for each single GPCR, as the sequence identity between rhodopsin and the homologous GPCRs is low and, in most cases, extra restraints need to be

introduced and/or ab initio modeling has to be integrated with comparative modeling. The best structure achieved by comparative modeling cannot be used as it is, but rather it must be considered as the starting point of further calculations aimed at improving the reliability of the computational model. These conclusions, which are based upon critical comparisons between rhodopsin and each member of family A, discourage automatic high throughput modeling of GPCRs based upon the rhodopsin structure and discourage the use of high throughput modeled GPCRs as targets of high throughput screening of libraries of compounds. In fact, in the majority of the virtual screening experiments based on GPCR models, the receptor model was first primed with a representative ligand[139,334,385,522] or built by adding external restraints concerning information on ligand binding.[411,501]

GPCRs are allosteric proteins which exist as complex statistical conformation ensembles.[444,457,642] They hold regions at high stability (i.e. low flexibility) and regions at low stability (i.e. high flexibility) that communicate with each other, even if distal. The functional properties of a GPCR are related to the distribution of states within the native ensemble.[444,457] Such a distribution is differently affected by ligands and/or interacting proteins and/or amino acid mutations.[444,457] Of course, the different oligomeric states of a GPCR contribute to differentiating the distribution of the receptor states.

The challenge of future computational modeling approaches is to provide a stochastic description of the GPCR systems and to predict, with atomic detail, the mechanisms by which structural information is transferred from the extracellular site to the intracellular site of the same receptor molecule or within a molecular network of receptors or within a multicomponent signaling unit. This would require an extensive integration between different molecular simulation methods and a careful weighing of approximations. In this respect, we think that continuous electrostatic models of the protein environment should be supported over explicit representation of lipids and water, to allow effective simulations of multiprotein assemblies. Explicit membrane/water representation does not necessarily improve the reliability of the system, as the lipid composition of a biological membrane is quite variable and the lipid−receptor stoichiometry is ill-defined, due also to indeterminations in the oligomeric states of the receptor.

GPCRs are very complex integrated chemical information systems that can be described and represented by chemical−structural formalisms. Quantitative molecular descriptors can be computed on average structural representatives of relevant conformational ensembles rather than on a single structure. This approach is expected to increase the information content of the selected molecular descriptors and overcome, at least in part, the inadequacy of the static single structure description.

The future of computational modeling of GPCRs relies in protein−protein and ligand−protein docking as well as in effective and exhaustive molecular simulation approaches aimed at predicting the probability distribution of the conformational ensembles of supramolecular assemblies. Moreover, the collective organizing principles of GPCRs appropriate to the mesoscopic scale domain will probably be the next formidable challenge for theoretical and computational modeling approaches aimed at describing the mesoscopic principles of GPCR-mediated signal transduction. These mesoscopic principles formally derive from atomistic/molecular rules but are, at the same time, independent of them.[643]

8. Acknowledgment

We thank Susanna Cotecchia and Tommaso Costa for valuable discussion. We are very grateful to Ongun Onaran for the valuable hints about the thermodynamic models and for providing us with effective schematizations of such models (i.e. shown in Figure 10). We thank Chris Reynolds for giving us the coordinates of their computational models of inactive and active rhodopsin, shown in Figure 7. We are also grateful to Slawomir Filipek for sending us the coordinates of the complex between tetrameric rhodopsin and heterotrimeric transducin, shown in Figure 16. We apologize to the investigators whose contributions have been unintentionally neglected in this review. This study was supported by Telethon-Italy Grant No. TCP00068 (to F.F.).

9. Note Added after ASAP Publication

This manuscript was originally published August 25, 2005 with a typographical error in the penultimate paragraph of section 5.2, and an incorrect reference number in the caption to Figure 15. The correct version of this manuscript was published September 1, 2005.

10. References

(1) Pierce, K. L.; Premont, R. T.; Lefkowitz, R. J. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 639.
(2) Brady, A. E.; Limbird, L. E. *Cell Signalling* **2002**, *14*, 297.
(3) Zhang, M.; Wang, W. *Acc. Chem. Res.* **2003**, *36*, 530.
(4) Bockaert, J.; Pin, J. P. *EMBO J.* **1999**, *18*, 1723.
(5) Gether, U. *Endocr. Rev.* **2000**, *21*, 90.
(6) Lefkowitz, R. J. *Nat. Cell Biol.* **2000**, *2*, E133.
(7) Kristiansen, K. *Pharmacol. Ther.* **2004**, *103*, 21.
(8) Tyndall, J. D.; Pfeiffer, B.; Abbenante, G.; Fairlie, D. P. *Chem. Rev.* **2005**, *105*, 793.
(9) Schertler, G. F.; Villa, C.; Henderson, R. *Nature* **1993**, *362*, 770.
(10) Baldwin, J. M. *EMBO J.* **1993**, *12*, 1693.
(11) Baldwin, J. M.; Schertler, G. F.; Unger, V. M. *J. Mol. Biol.* **1997**, *272*, 144.
(12) Unger, V. M.; Hargrave, P. A.; Baldwin, J. M.; Schertler, G. F. *Nature* **1997**, *389*, 203.
(13) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739.
(14) Kunishima, N.; Shimada, Y.; Tsuji, Y.; Sato, T.; Yamamoto, M.; Kumasaka, T.; Nakanishi, S.; Jingami, H.; Morikawa, K. *Nature* **2000**, *407*, 971−977.
(15) Fan, Q. R.; Hendrickson, W. A. *Nature* **2005**, *433*, 269.
(16) Foord, S. M. *Curr. Opin. Pharmacol.* **2002**, *2*, 561.
(17) Fredriksson, R.; Lagerstrom, M. C.; Lundin, L. G.; Schioth, H. B. *Mol. Pharmacol.* **2003**, *63*, 1256.
(18) Fredriksson, R.; Schioth, H. B. *Mol. Pharmacol.* **2005**, *67*, 1414.
(19) Kolakowski, L. F., Jr. *Receptors Channels* **1994**, *2*, 1.
(20) Eilers, M.; Hornak, V.; Smith, S. O.; Konopka, J. B. *Biochemistry* **2005**, *44*, 8959.
(21) Gilman, A. G. *Annu. Rev. Biochem.* **1987**, *56*, 615.

(22) Bourne, H. R.; Sanders, D. A.; McCormick, F. *Nature* **1991**, *349*, 117.
(23) Alves, I. D.; Salamon, Z.; Varga, E.; Yamamura, H. I.; Tollin, G.; Hruby, V. J. *J. Biol. Chem.* **2003**, *278*, 48890.
(24) Alves, I. D.; Salgado, G. F.; Salamon, Z.; Brown, M. F.; Tollin, G.; Hruby, V. J. *Biophys. J.* **2005**, *88*, 198.
(25) Marinissen, M. J.; Gutkind, J. S. *Trends Pharmacol. Sci.* **2001**, *22*, 368.
(26) Brzostowski, J. A.; Kimmel, A. R. *Trends Biochem. Sci.* **2001**, *26*, 291.
(27) Luttrell, L. M.; Roudabush, F. L.; Choy, E. W.; Miller, W. E.; Field, M. E.; Pierce, K. L.; Lefkowitz, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2449.
(28) Hall, R. A.; Lefkowitz, R. J. *Circ. Res.* **2002**, *91*, 672.
(29) Arvanitakis, L.; Geras-Raaka, E.; Gershengorn, M. C. *Trends Endocrinol. Metab.* **1998**, *9*, 27.
(30) Krupnick, J. G.; Benovic, J. L. *Annu. Rev. Pharmacol. Toxicol.* **1998**, *38*, 289.
(31) Bouvier, M. *Nat. Rev. Neurosci.* **2001**, *2*, 274.
(32) Milligan, G. *J. Cell Sci.* **2001**, *114*, 1265.
(33) Rios, C. D.; Jordan, B. A.; Gomes, I.; Devi, L. A. *Pharmacol. Ther.* **2001**, *92*, 71.
(34) George, S. R.; O'Dowd, B. F.; Lee, S. P. *Nat. Rev. Drug Discovery* **2002**, *1*, 808.
(35) Agnati, L. F.; Ferre, S.; Lluis, C.; Franco, R.; Fuxe, K. *Pharmacol. Rev.* **2003**, *55*, 509.
(36) Franco, R.; Canals, M.; Marcellino, D.; Ferre, S.; Agnati, L.; Mallol, J.; Casado, V.; Ciruela, F.; Fuxe, K.; Lluis, C.; Canela, E. I. *Trends Biochem. Sci.* **2003**, *28*, 238.
(37) Kroeger, K. M.; Pfleger, K. D.; Eidne, K. A. *Front Neuroendocrinol.* **2003**, *24*, 254.
(38) Terrillon, S.; Bouvier, M. *EMBO Rep.* **2004**, *5*, 30.
(39) Park, P. S.; Filipek, S.; Wells, J. W.; Palczewski, K. *Biochemistry* **2004**, *43*, 15643.
(40) Maggio, R.; Novi, F.; Scarselli, M.; Corsini, G. U. *FEBS J.* **2005**, *272*, 2939.
(41) Bulenger, S.; Marullo, S.; Bouvier, M. *Trends Pharmacol. Sci.* **2005**, *26*, 131.
(42) Reddy, P. S.; Corley, R. B. *Bioessays* **1998**, *20*, 546.
(43) Fotiadis, D.; Liang, Y.; Filipek, S.; Saperstein, D. A.; Engel, A.; Palczewski, K. *Nature* **2003**, *421*, 127.
(44) Lefkowitz, R. J.; Cotecchia, S.; Samama, P.; Costa, T. *Trends Pharmacol. Sci.* **1993**, *14*, 303.
(45) Cotecchia, S.; Fanelli, F.; Scheer, A.; De Benedetti, P. G. In *Structure/function analysis of GPCRs*; Wess, J., Ed.; Wiley-Liss: New York, 1999; Vol. 3.
(46) Cotecchia, S.; Fanelli, F.; Scheer, A.; Costa, T. In *G Protein-Coupled Receptors as Drug Targets. Analysis of Activation and Constitutive Activity*; Seifert, R., Wieland, T., Eds.; 2005.
(47) Chalmers, D. T.; Behan, D. P. *Nat. Rev. Drug Discovery* **2002**, *1*, 599.
(48) Shenker, A. *Baillieres Clin. Endocrinol. Metab.* **1995**, *9*, 427.
(49) Themmen, A. P.; Martens, J. W.; Brunner, H. G. *Mol. Cell Endocrinol.* **1998**, *145*, 137.
(50) Mendes, H. F.; van der Spuy, J.; Chapple, J. P.; Cheetham, M. E. *Trends Mol. Med.* **2005**, *11*, 177.
(51) Teller, D. C.; Okada, T.; Behnke, C. A.; Palczewski, K.; Stenkamp, R. E. *Biochemistry* **2001**, *40*, 7761.
(52) Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5982.
(53) Li, J.; Edwards, P. C.; Burghammer, M.; Villa, C.; Schertler, G. F. *J. Mol. Biol.* **2004**, *343*, 1409.
(54) Okada, T.; Sugihara, M.; Bondar, A. N.; Elstner, M.; Entel, P.; Buss, V. *J. Mol. Biol.* **2004**, *342*, 571.
(55) Ballesteros, J. A.; Weinstein, H. *Methods Neurosci.* **1995**, *25*, 366.
(56) Cohen, G. B.; Yang, T.; Robinson, P. R.; Oprian, D. D. *Biochemistry* **1993**, *32*, 6111.
(57) Scheer, A.; Fanelli, F.; Costa, T.; De Benedetti, P. G.; Cotecchia, S. *EMBO J.* **1996**, *15*, 3566.
(58) Scheer, A.; Fanelli, F.; Costa, T.; De Benedetti, P. G.; Cotecchia, S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 808.
(59) Ballesteros, J.; Kitanovic, S.; Guarnieri, F.; Davies, P.; Fromme, B. J.; Konvicka, K.; Chi, L.; Millar, R. P.; Davidson, J. S.; Weinstein, H.; Sealfon, S. C. *J. Biol. Chem.* **1998**, *273*, 10445.
(60) Fanelli, F.; Barbier, P.; Zanchetta, D.; de Benedetti, P. G.; Chini, B. *Mol. Pharmacol.* **1999**, *56*, 214.
(61) Ballesteros, J. A.; Jensen, A. D.; Liapakis, G.; Rasmussen, S. G.; Shi, L.; Gether, U.; Javitch, J. A. *J. Biol. Chem.* **2001**, *276*, 29171.
(62) Greasley, P. J.; Fanelli, F.; Scheer, A.; Abuin, L.; Nenniger-Tosato, M.; DeBenedetti, P. G.; Cotecchia, S. *J. Biol. Chem.* **2001**, *276*, 46485.
(63) Li, J.; Huang, P.; Chen, C.; de Riel, J. K.; Weinstein, H.; Liu-Chen, L. Y. *Biochemistry* **2001**, *40*, 12039.
(64) Greasley, P. J.; Fanelli, F.; Rossier, O.; Abuin, L.; Cotecchia, S. *Mol. Pharmacol.* **2002**, *61*, 1025.
(65) Angelova, K.; Fanelli, F.; Puett, D. *J. Biol. Chem.* **2002**, *277*, 32202.
(66) Visiers, I.; Ebersole, B.; Dracheva, S.; Ballesteros, J.; Sealfon, S. C.; Weinstein, H. *Int. J. Quantum Chem.* **2002**, *88*, 65.
(67) Shapiro, D. A.; Kristiansen, K.; Weiner, D. M.; Kroeze, W. K.; Roth, B. L. *J. Biol. Chem.* **2002**, *277*, 11441.
(68) Seeber, M.; De Benedetti, P. G.; Fanelli, F. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1520.
(69) Fanelli, F.; Verhoef-Post, M.; Timmerman, M.; Zeilemaker, A.; Martens, J. W.; Themmen, A. P. *Mol. Endocrinol.* **2004**, *18*, 1499.
(70) Vassart, G.; Pardo, L.; Costagliola, S. *Trends Biochem. Sci.* **2004**, *29*, 119.
(71) Favre, N.; Fanelli, F.; Missotten, M.; Nichols, A.; Wilson, J.; Di Tiani, M.; Rommel, C.; Scheer, A. *Biochemistry* **2005**, *44*, 9990.
(72) Zhang, M.; Mizrachi, D.; Fanelli, F.; Segaloff, D. L. *J. Biol. Chem.* **2005**, *280*, 26169.
(73) Flanagan, C. A. *Mol. Pharmacol.* **2005**, *68*, 1.
(74) Rosenkilde, M. M.; Kledal, T. N.; Schwartz, T. W. *Mol. Pharmacol.* **2005**, *68*, 11.
(75) Mirzadegan, T.; Benko, G.; Filipek, S.; Palczewski, K. *Biochemistry* **2003**, *42*, 2759.
(76) Fritze, O.; Filipek, S.; Kuksa, V.; Palczewski, K.; Hofmann, K. P.; Ernst, O. P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2290.
(77) Jacobs, D. J.; Rader, A. J.; Kuhn, L. A.; Thorpe, M. F. *Proteins* **2001**, *44*, 150.
(78) Bahar, I.; Atilgan, A. R.; Erman, B. *Folding Des.* **1997**, *2*, 173.
(79) Rader, A. J.; Anderson, G.; Isin, B.; Khorana, H. G.; Bahar, I.; Klein-Seetharaman, J. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7246.
(80) Klco, J. M.; Wiegand, C. B.; Narzinski, K.; Baranski, T. J. *Nat. Struct. Mol. Biol.* **2005**, *12*, 320.
(81) Massotte, D.; Kieffer, B. L. *Nat. Struct. Mol. Biol.* **2005**, *12*, 287.
(82) Findlay, J. B.; Pappin, D. J. *Biochem. J.* **1986**, *238*, 625.
(83) Gouldson, P. R.; Snell, C. R.; Reynolds, C. A. *J. Med. Chem.* **1997**, *40*, 3871.
(84) Konvicka, K.; Guarnieri, F.; Ballesteros, J. A.; Weinstein, H. *Biophys. J.* **1998**, *75*, 601.
(85) Chabre, M.; Cone, R.; Saibil, H. *Nature* **2003**, *426*, 30.
(86) Fotiadis, D.; Liang, Y.; Filipek, S.; Saperstein, D. A.; Engel, A.; Palczewski, K. *FEBS Lett.* **2004**, *564*, 281.
(87) Jastrzebska, B.; Maeda, T.; Zhu, L.; Fotiadis, D.; Filipek, S.; Engel, A.; Stenkamp, R. E.; Palczewski, K. *J. Biol. Chem.* **2004**, *279*, 54663.
(88) Suda, K.; Filipek, S.; Palczewski, K.; Engel, A.; Fotiadis, D. *Mol. Membr. Biol.* **2004**, *21*, 435.
(89) Chabre, M.; le Maire, M. *Biochemistry* **2005**, *44*, 9395.
(90) Liang, Y.; Fotiadis, D.; Filipek, S.; Saperstein, D. A.; Palczewski, K.; Engel, A. *J. Biol. Chem.* **2003**, *278*, 21655.
(91) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
(92) Okada, T. *Biochem. Soc. Trans.* **2004**, *32*, 738.
(93) Filipek, S. *J. Mol. Model*, published online 2005, 10.1007/s00894-005-0268-3.
(94) Ciarkowski, J.; Witt, M.; Slusarz, R. *J. Mol. Model*, published online 2005, 10.1007/s00894-005-0270-9.
(95) Davies, A.; Gowen, B. E.; Krebs, A. M.; Schertler, G. F.; Saibil, H. R. *J. Mol. Biol.* **2001**, *314*, 455.
(96) Meng, E. C.; Bourne, H. R. *Trends Pharmacol. Sci.* **2001**, *22*, 587.
(97) McBee, J. K.; Palczewski, K.; Baehr, W.; Pepperberg, D. R. *Prog. Retinal Eye Res.* **2001**, *20*, 469.
(98) Okada, T.; Ernst, O. P.; Palczewski, K.; Hofmann, K. P. *Trends Biochem. Sci.* **2001**, *26*, 318.
(99) Kliger, D. S.; Lewis, J. W. *Isr. J. Chem.* **1995**, *35*, 289.
(100) Yan, E. C.; Kazmi, M. A.; Ganim, Z.; Hou, J. M.; Pan, D.; Chang, B. S.; Sakmar, T. P.; Mathies, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9262.
(101) Bartl, F. J.; Ritter, E.; Hofmann, K. P. *J. Biol. Chem.* **2001**, *276*, 30161.
(102) Heck, M.; Schadel, S. A.; Maretzki, D.; Bartl, F. J.; Ritter, E.; Palczewski, K.; Hofmann, K. P. *J. Biol. Chem.* **2003**, *278*, 3162.
(103) Fahmy, K.; Jager, F.; Beck, M.; Zvyaga, T. A.; Sakmar, T. P.; Siebert, F. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10206.
(104) Ruprecht, J. J.; Mielke, T.; Vogel, R.; Villa, C.; Schertler, G. F. *EMBO J.* **2004**, *23*, 3609.
(105) Arnis, S.; Hofmann, K. P. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7849.
(106) Fahmy, K.; Sakmar, T. P.; Siebert, F. *Biochemistry* **2000**, *39*, 10607.
(107) Kim, J. M.; Altenbach, C.; Thurmond, R. L.; Khorana, H. G.; Hubbell, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14273.
(108) Arnis, S.; Fahmy, K.; Hofmann, K. P.; Sakmar, T. P. *J. Biol. Chem.* **1994**, *269*, 23879.
(109) Meyer, C. K.; Bohme, M.; Ockenfels, A.; Gartner, W.; Hofmann, K. P.; Ernst, O. P. *J. Biol. Chem.* **2000**, *275*, 19713.
(110) Palczewski, K. *Eur. J. Biochem.* **1997**, *248*, 261.
(111) Pulvermuller, A.; Palczewski, K.; Hofmann, K. P. *Biochemistry* **1993**, *32*, 14082.
(112) Farrens, D. L.; Khorana, H. G. *J. Biol. Chem.* **1995**, *270*, 5073.
(113) Heck, M.; Schadel, S. A.; Maretzki, D.; Hofmann, K. P. *Vision Res.* **2003**, *43*, 3003.

(114) Ritter, E.; Zimmermann, K.; Heck, M.; Hofmann, K. P.; Bartl, F. J. *J. Biol. Chem.* **2004**, *279*, 48102.
(115) Schadel, S. A.; Heck, M.; Maretzki, D.; Filipek, S.; Teller, D. C.; Palczewski, K.; Hofmann, K. P. *J. Biol. Chem.* **2003**, *278*, 24896.
(116) Melia, T. J., Jr.; Cowan, C. W.; Angleson, J. K.; Wensel, T. G. *Biophys. J.* **1997**, *73*, 3182.
(117) Borhan, B.; Souto, M. L.; Imai, H.; Shichida, Y.; Nakanishi, K. *Science* **2000**, *288*, 2209.
(118) Patel, A. B.; Crocker, E.; Eilers, M.; Hirshfeld, A.; Sheves, M.; Smith, S. O. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10048.
(119) Patel, A. B.; Crocker, E.; Reeves, P. J.; Getmanova, E. V.; Eilers, M.; Khorana, H. G.; Smith, S. O. *J. Mol. Biol.* **2005**, *347*, 803.
(120) Spooner, P. J.; Sharples, J. M.; Goodall, S. C.; Bovee-Geurts, P. H.; Verhoeven, M. A.; Lugtenburg, J.; Pistorius, A. M.; Degrip, W. J.; Watts, A. *J. Mol. Biol.* **2004**, *343*, 719.
(121) Klein-Seetharaman, J.; Yanamala, N. V.; Javeed, F.; Reeves, P. J.; Getmanova, E. V.; Loewen, M. C.; Schwalbe, H.; Khorana, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3409.
(122) Getmanova, E.; Patel, A. B.; Klein-Seetharaman, J.; Loewen, M. C.; Reeves, P. J.; Friedman, N.; Sheves, M.; Smith, S. O.; Khorana, H. G. *Biochemistry* **2004**, *43*, 1126.
(123) Yu, H.; Kono, M.; Oprian, D. D. *Biochemistry* **1999**, *38*, 12028.
(124) Farahbakhsh, Z. T.; Ridge, K. D.; Khorana, H. G.; Hubbell, W. L. *Biochemistry* **1995**, *34*, 8812.
(125) Altenbach, C.; Yang, K.; Farrens, D. L.; Farahbakhsh, Z. T.; Khorana, H. G.; Hubbell, W. L. *Biochemistry* **1996**, *35*, 12470.
(126) Altenbach, C.; Cai, K.; Khorana, H. G.; Hubbell, W. L. *Biochemistry* **1999**, *38*, 7931.
(127) Janz, J. M.; Farrens, D. L. *J. Biol. Chem.* **2004**, *279*, 29767.
(128) Altenbach, C.; Klein-Seetharaman, J.; Hwa, J.; Khorana, H. G.; Hubbell, W. L. *Biochemistry* **1999**, *38*, 7945.
(129) Kim, J. M.; Altenbach, C.; Kono, M.; Oprian, D. D.; Hubbell, W. L.; Khorana, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12508.
(130) Farrens, D. L.; Altenbach, C.; Yang, K.; Hubbell, W. L.; Khorana, H. G. *Science* **1996**, *274*, 768.
(131) Huber, T.; Botelho, A. V.; Beyer, K.; Brown, M. F. *Biophys. J.* **2004**, *86*, 2078.
(132) Crozier, P. S.; Stevens, M. J.; Forrest, L. R.; Woolf, T. B. *J. Mol. Biol.* **2003**, *333*, 493.
(133) Pitman, M. C.; Grossfield, A.; Suits, F.; Feller, S. E. *J. Am. Chem. Soc.* **2005**, *127*, 4576.
(134) Rohrig, U. F.; Guidoni, L.; Rothlisberger, U. *Biochemistry* **2002**, *41*, 10799.
(135) Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K. *Biophys. J.* **2002**, *83*, 3097.
(136) Yeagle, P. L.; Choi, G.; Albert, A. D. *Biochemistry* **2001**, *40*, 11932.
(137) Choi, G.; Landin, J.; Galan, J. F.; Birge, R. R.; Albert, A. D.; Yeagle, P. L. *Biochemistry* **2002**, *41*, 7318.
(138) Yeagle, P. L.; Albert, A. D. *Biochemistry* **2003**, *42*, 1365.
(139) Gouldson, P. R.; Kidley, N. J.; Bywater, R. P.; Psaroudakis, G.; Brooks, H. D.; Diaz, C.; Shire, D.; Reynolds, C. A. *Proteins* **2004**, *56*, 67.
(140) Nikiforovich, G. V.; Marshall, G. R. *Biochemistry* **2003**, *42*, 9110.
(141) Chambers, J. J.; Nichols, D. E. *J. Comput-Aided Mol. Des.* **2002**, *16*, 511.
(142) Ishiguro, M.; Hirano, T.; Oyama, Y. *ChemBioChem* **2003**, *4*, 228.
(143) Ishiguro, M.; Oyama, Y.; Hirano, T. *ChemBioChem* **2004**, *5*, 298.
(144) Ishiguro, M. *ChemBioChem* **2004**, *5*, 1210.
(145) Horn, F.; Weare, J.; Beukers, M. W.; Horsch, S.; Bairoch, A.; Chen, W.; Edvardsen, O.; Campagne, F.; Vriend, G. *Nucleic Acids Res.* **1998**, *26*, 275.
(146) Horn, F.; Vriend, G.; Cohen, F. E. *Nucleic Acids Res* **2001**, *29*, 346.
(147) Horn, F.; Bettler, E.; Oliveira, L.; Campagne, F.; Cohen, F. E.; Vriend, G. *Nucleic Acids Res.* **2003**, *31*, 294.
(148) Horn, F.; Lau, A. L.; Cohen, F. E. *Bioinformatics* **2004**, *20*, 557.
(149) Crasto, C.; Marenco, L.; Miller, P.; Shepherd, G. *Nucleic Acids Res.* **2002**, *30*, 354.
(150) Kristiansen, K.; Dahl, S. G.; Edvardsen, O. *Proteins* **1996**, *26*, 81.
(151) Edvardsen, O.; Kristiansen, K. *Trends Biochem. Sci.* **1997**, *22*, 226.
(152) Beukers, M. W.; Kristiansen, I.; AP, I. J.; Edvardsen, I. *Trends Pharmacol. Sci.* **1999**, *20*, 475.
(153) Edvardsen, O.; Reiersen, A. L.; Beukers, M. W.; Kristiansen, K. *Nucleic Acids Res.* **2002**, *30*, 361.
(154) Campagne, F.; Jestin, R.; Reversat, J. L.; Bernassau, J. M.; Maigret, B. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 625.
(155) Peitsch, M. C. *Biochem. Soc. Trans.* **1996**, *24*, 274.
(156) Bhasin, M.; Raghava, G. P. *Nucleic Acids Res.* **2004**, *32*, W383.
(157) Gaulton, A.; Attwood, T. K. *Curr. Opin. Pharmacol.* **2003**, *3*, 114.
(158) Elefsinioti, A. L.; Bagos, P. G.; Spyropoulos, I. C.; Hamodrakas, S. J. *BMC Bioinformatics* **2004**, *5*, 208.
(159) Henderson, R.; Schertler, G. F. *Philos. Trans. R. Soc. London, B: Biol. Sci.* **1990**, *326*, 379.
(160) Hibert, M. F.; Trumpp-Kallmeyer, S.; Bruinvels, A.; Hoflack, J. *Mol. Pharmacol.* **1991**, *40*, 8.
(161) Ijzerman, A. P.; Van Galen, P. J.; Jacobson, K. A. *Drug Des. Discovery* **1992**, *9*, 49.
(162) Livingstone, C. D.; Strange, P. G.; Naylor, L. H. *Biochem. J.* **1992**, *287 (Pt 1)*, 277.
(163) Trumpp-Kallmeyer, S.; Hoflack, J.; Bruinvels, A.; Hibert, M. *J. Med. Chem.* **1992**, *35*, 3448.
(164) Cronet, P.; Sander, C.; Vriend, G. *Protein Eng.* **1993**, *6*, 59.
(165) De Benedetti, P. G.; Menziani, M. C.; Fanelli, F.; Cocchi, M. *THEOCHEM* **1993**, *285*, 147.
(166) Hoflack, J.; Hibert, M. F.; Trumpp-Kallmeyer, S.; Bidart, J. M. *Drug Des. Discovery* **1993**, *10*, 157.
(167) Huggins, J. P.; Trumpp-Kallmeyer, S.; Hibert, M. F.; Hoflack, J. M.; Fanger, B. O.; Jones, C. R. *Eur. J. Pharmacol.* **1993**, *245*, 203.
(168) Nordvall, G.; Hacksell, U. *J. Med. Chem.* **1993**, *36*, 967.
(169) Oliveira, L.; Paiva, A. C. M.; Vriend, G. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 649.
(170) Yamamoto, Y.; Kamiya, K.; Terao, S. *J. Med. Chem.* **1993**, *36*, 820.
(171) Ijzerman, A. P.; van der Wenden, E. M.; van Galen, P. J.; Jacobson, K. A. *Eur. J. Pharmacol.* **1994**, *268*, 95.
(172) Fanelli, F.; Menziani, M. C.; Carotti, A.; De Benedetti, P. G. *Bioorg. Med. Chem.* **1994**, *2*, 195.
(173) Fanelli, F.; Menziani, M. C.; Cocchi, M.; Leonardi, A.; De Benedetti, P. G. *THEOCHEM* **1994**, *314*, 265.
(174) Kuipers, W.; Van Wijngaarden, I.; Ijzerman, A. P. *Drug Des. Discovery* **1994**, *11*, 231.
(175) Kyle, D. J.; Chakravarty, S.; Sinsko, J. A.; Stormann, T. M. *J. Med. Chem.* **1994**, *37*, 1347.
(176) Malmberg, A.; Nordvall, G.; Johansson, A. M.; Mohell, N.; Hacksell, U. *Mol. Pharmacol.* **1994**, *46*, 299.
(177) Teeter, M. M.; Froimowitz, M.; Stec, B.; DuRand, C. J. *J. Med. Chem.* **1994**, *37*, 2874.
(178) van Galen, P. J.; van Bergen, A. H.; Gallo-Rodriguez, C.; Melman, N.; Olah, M. E.; AP, I. J.; Stiles, G. L.; Jacobson, K. A. *Mol. Pharmacol.* **1994**, *45*, 1101.
(179) Groblewski, T.; Maigret, B.; Nouet, S.; Larguier, R.; Lombard, C.; Bonnafous, J. C.; Marie, J. *Biochem. Biophys. Res. Commun.* **1995**, *209*, 153.
(180) Hedberg, M. H.; Johansson, A. M.; Nordvall, G.; Yliniemela, A.; Li, H. B.; Martin, A. R.; Hjorth, S.; Unelius, L.; Sundell, S.; Hacksell, U. *J. Med. Chem.* **1995**, *38*, 647.
(181) Jagerschmidt, A.; Guillaume, N.; Goudreau, N.; Maigret, B.; Roques, B. P. *Mol. Pharmacol.* **1995**, *48*, 783.
(182) Joseph, M. P.; Maigret, B.; Scheraga, H. A. *Int. J. Peptide Protein Res.* **1995**, *46*, 514.
(183) Joseph, M. P.; Maigret, B.; Bonnafous, J. C.; Marie, J.; Scheraga, H. A. *J. Protein Chem.* **1995**, *14*, 381.
(184) Kaupmann, K.; Bruns, C.; Raulf, F.; Weber, H. P.; Mattes, H.; Lubbert, H. *EMBO J.* **1995**, *14*, 727.
(185) ter Laak, A. M.; Timmerman, H.; Leurs, R.; Nederkoorn, P. H.; Smit, M. J.; Donne-Op den Kelder, G. M. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 319.
(186) Yamano, Y.; Ohyama, K.; Kikyo, M.; Sano, T.; Nakagomi, Y.; Inoue, Y.; Nakamura, N.; Morishima, I.; Guo, D. F.; Hamakubo, T.; et al. *J. Biol. Chem.* **1995**, *270*, 14024.
(187) Befort, K.; Tabbara, L.; Kling, D.; Maigret, B.; Kieffer, B. L. *J. Biol. Chem.* **1996**, *271*, 10161.
(188) Forbes, I. T.; Dabbs, S.; Duckworth, D. M.; Ham, P.; Jones, G. E.; King, F. D.; Saunders: D. V.; Blaney, F. E.; Naylor, C. B.; Baxter, G. S.; Blackburn, T. P.; Kennett, G. A.; Wood, M. D. *J. Med. Chem.* **1996**, *39*, 4966.
(189) Jagerschmidt, A.; Guillaume-Rousselet, N.; Vikland, M. L.; Goudreau, N.; Maigret, B.; Roques, B. P. *Eur. J. Pharmacol.* **1996**, *296*, 97.
(190) Wieland, K.; Zuurmond, H. M.; Krasel, C.; Ijzerman, A. P.; Lohse, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9276.
(191) Bromidge, S. M.; Duckworth, M.; Forbes, I. T.; Ham, P.; King, F. D.; Thewlis, K. M.; Blaney, F. E.; Naylor, C. B.; Blackburn, T. P.; Kennett, G. A.; Wood, M. D.; Clarke, S. E. *J. Med. Chem.* **1997**, *40*, 3494.
(192) Mahmoudian, M. *J. Mol. Graphics Modell.* **1997**, *15*, 149.
(193) Kuipers, W.; Link, R.; Standaar, P. J.; Stoit, A. R.; Van Wijngaarden, I.; Leurs, R.; Ijzerman, A. P. *Mol. Pharmacol.* **1997**, *51*, 889.
(194) Bromidge, S. M.; Dabbs, S.; Davies, D. T.; Duckworth, D. M.; Forbes, I. T.; Ham, P.; Jones, G. E.; King, F. D.; Saunders: D. V.; Starr, S.; Thewlis, K. M.; Wyman, P. A.; Blaney, F. E.; Naylor, C. B.; Bailey, F.; Blackburn, T. P.; Holland, V.; Kennett, G. A.; Riley, G. J.; Wood, M. D. *J. Med. Chem.* **1998**, *41*, 1598.
(195) Spedding, M.; Newman-Tancredi, A.; Millan, M. J.; Dacquet, C.; Michel, A. N.; Jacoby, E.; Vickery, B.; Tallentire, D. *Neuropharmacology* **1998**, *37*, 769.
(196) Jacoby, E.; Fauchere, J.-L.; Raimbaud, E.; Ollivier, S.; Michel, A.; Spedding, M. *Quant. Struct.-Act. Relat.* **1999**, *18*, 561.
(197) Zuurmond, H. M.; Hessling, J.; Bluml, K.; Lohse, M.; Ijzerman, A. P. *Mol. Pharmacol.* **1999**, *56*, 909.
(198) Bromidge, S. M.; Dabbs, S.; Davies, D. T.; Davies, S.; Duckworth, D. M.; Forbes, I. T.; Gaster, L. M.; Ham, P.; Jones, G. E.; King,

F. D.; Mulholland, K. R.; Saunders: D. V.; Wyman, P. A.; Blaney, F. E.; Clarke, S. E.; Blackburn, T. P.; Holland, V.; Kennett, G. A.; Lightowler, S.; Middlemiss, D. N.; Trail, B.; Riley, G. J.; Wood, M. D. *J. Med. Chem.* **2000**, *43*, 1123.
(199) Lavecchia, A.; Greco, G.; Novellino, E.; Vittorio, F.; Ronsisvalle, G. *J. Med. Chem.* **2000**, *43*, 2124.
(200) Macdonald, D.; Murgolo, N.; Zhang, R.; Durkin, J. P.; Yao, X.; Strader, C. D.; Graziano, M. P. *Mol. Pharmacol.* **2000**, *58*, 217.
(201) Zhou, N.; Luo, Z.; Hall, J. W.; Luo, J.; Han, X.; Huang, Z. *Eur. J. Immunol.* **2000**, *30*, 164.
(202) Zhou, N.; Luo, Z.; Luo, J.; Liu, D.; Hall, J. W.; Pomerantz, R. J.; Huang, Z. *J. Biol. Chem.* **2001**, *276*, 42826.
(203) Iadanza, M.; Holtje, M.; Ronsisvalle, G.; Holtje, H. D. *J. Med. Chem.* **2002**, *45*, 4838.
(204) Manivet, P.; Schneider, B.; Smith, J. C.; Choi, D. S.; Maroteaux, L.; Kellermann, O.; Launay, J. M. *J. Biol. Chem.* **2002**, *277*, 17170.
(205) Marie, J.; Richard, E.; Pruneau, D.; Paquet, J. L.; Siatka, C.; Larguier, R.; Ponce, C.; Vassault, P.; Groblewski, T.; Maigret, B.; Bonnafous, J. C. *J. Biol. Chem.* **2001**, *276*, 41100.
(206) Pardo, L.; Ballesteros, J. A.; Osman, R.; Weinstein, H. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4009.
(207) Taylor, E. W.; Agarwal, A. *FEBS Lett.* **1993**, *325*, 161.
(208) Metzger, T. G.; Paterlini, M. G.; Portoghese, P. S.; Ferguson, D. M. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 857.
(209) Metzger, T. G.; Paterlini, M. G.; Portoghese, P. S.; Ferguson, D. M. *Neurochem. Res.* **1996**, *21*, 1287.
(210) Paterlini, G.; Portoghese, P. S.; Ferguson, D. M. *J. Med. Chem.* **1997**, *40*, 3254.
(211) Soppa, J. *FEBS Lett.* **1994**, *342*, 7.
(212) Jones, D.; Thornton, J. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 439.
(213) Wodak, S.; Rooman, M. *Curr. Opin. Struct. Biol.* **1993**, *3*, 247.
(214) Escrieut, C.; Gigoux, V.; Archer, E.; Verrier, S.; Maigret, B.; Behrendt, R.; Moroder, L.; Bignon, E.; Silvente-Poirot, S.; Pradayrol, L.; Fourmy, D. *J. Biol. Chem.* **2002**, *277*, 7546.
(215) Fourmy, D.; Escrieut, C.; Archer, E.; Gales, C.; Gigoux, V.; Maigret, B.; Moroder, L.; Silvente-Poirot, S.; Martinez, J.; Fehrentz, J. A.; Pradayrol, L. *Pharmacol. Toxicol.* **2002**, *91*, 313.
(216) Gales, C.; Poirot, M.; Taillefer, J.; Maigret, B.; Martinez, J.; Moroder, L.; Escrieut, C.; Pradayrol, L.; Fourmy, D.; Silvente-Poirot, S. *Mol. Pharmacol.* **2003**, *63*, 973.
(217) Higgs, C.; Reynolds, C. A. In *Theoretical and Computational Chemistry*; Eriksson, L. A., Ed.; Elsevier: Amsterdam, 2001; Vol. 9.
(218) Dahl, S. G.; Edvardsen, O.; Sylte, I. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 8111.
(219) MaloneyHuss, K.; Lybrand, T. P. *J. Mol. Biol.* **1992**, *225*, 859.
(220) Zhang, D.; Weinstein, H. *J. Med. Chem.* **1993**, *36*, 934.
(221) Zhang, D.; Weinstein, H. *FEBS Lett.* **1994**, *337*, 207.
(222) Almaula, N.; Ebersole, B. J.; Zhang, D.; Weinstein, H.; Sealfon, S. C. *J. Biol. Chem.* **1996**, *271*, 14672.
(223) Kontoyianni, M.; DeWeese, C.; Penzotti, J. E.; Lybrand, T. P. *J. Med. Chem.* **1996**, *39*, 4406.
(224) Ji, Z.; Hadac, E. M.; Henne, R. M.; Patel, S. A.; Lybrand, T. P.; Miller, L. J. *J. Biol. Chem.* **1997**, *272*, 24393.
(225) Hadac, E. M.; Pinon, D. I.; Ji, Z.; Holicky, E. L.; Henne, R. M.; Lybrand, T. P.; Miller, L. J. *J. Biol. Chem.* **1998**, *273*, 12988.
(226) Hamaguchi, N.; True, T. A.; Goetz, A. S.; Stouffer, M. J.; Lybrand, T. P.; Jeffs, P. W. *Biochemistry* **1998**, *37*, 5730.
(227) Hadac, E. M.; Ji, Z.; Pinon, D. I.; Henne, R. M.; Lybrand, T. P.; Miller, L. J. *J. Med. Chem.* **1999**, *42*, 2105.
(228) Ding, X. Q.; Dolu, V.; Hadac, E. M.; Holicky, E. L.; Pinon, D. I.; Lybrand, T. P.; Miller, L. J. *J. Biol. Chem.* **2001**, *276*, 4236.
(229) Elling, C. E.; Nielsen, S. M.; Schwartz, T. W. *Nature* **1995**, *374*, 74.
(230) Liu, J.; Schoneberg, T.; van Rhee, M.; Wess, J. *J. Biol. Chem.* **1995**, *270*, 19532.
(231) Mizobe, T.; Maze, M.; Lam, V.; Suryanarayana, S.; Kobilka, B. K. *J. Biol. Chem.* **1996**, *271*, 2387.
(232) Elling, C. E.; Schwartz, T. W. *EMBO J.* **1996**, *15*, 6213.
(233) Ding, X. Q.; Pinon, D. I.; Furse, K. E.; Lybrand, T. P.; Miller, L. J. *Mol. Pharmacol.* **2002**, *61*, 1041.
(234) Miller, L. J.; Lybrand, T. P. *Pharmacol. Toxicol.* **2002**, *91*, 282.
(235) Furse, K. E.; Lybrand, T. P. *J. Med. Chem.* **2003**, *46*, 4450.
(236) Harikumar, K. G.; Pinon, D. I.; Wessels, W. S.; Dawson, E. S.; Lybrand, T. P.; Prendergast, F. G.; Miller, L. J. *Mol. Pharmacol.* **2004**, *65*, 28.
(237) Perlman, J. H.; Laakkonen, L.; Osman, R.; Gershengorn, M. C. *J. Biol. Chem.* **1994**, *269*, 23383.
(238) Chini, B.; Mouillac, B.; Ala, Y.; Balestre, M. N.; Trumpp-Kallmeyer, S.; Hoflack, J.; Elands, J.; Hibert, M.; Manning, M.; Jard, S. *EMBO J.* **1995**, *14*, 2176.
(239) Kim, J.; Wess, J.; van Rhee, A. M.; Schoneberg, T.; Jacobson, K. A. *J. Biol. Chem.* **1995**, *270*, 13987.
(240) Mouillac, B.; Chini, B.; Balestre, M. N.; Elands, J.; Trumpp-Kallmeyer, S.; Hoflack, J.; Hibert, M.; Jard, S.; Barberis, C. *J. Biol. Chem.* **1995**, *270*, 25771.
(241) Van Rhee, A. M.; Fischer, B.; Van Galen, P. J.; Jacobson, K. A. *Drug Des. Discovery* **1995**, *13*, 133.
(242) Cappelli, A.; Anzini, M.; Vomero, S.; Menziani, M. C.; De Benedetti, P. G.; Sbacchi, M.; Clarke, G. D.; Mennuni, L. *J. Med. Chem.* **1996**, *39*, 860.
(243) Kristiansen, K.; Dahl, S. G. *Eur. J. Pharmacol.* **1996**, *306*, 195.
(244) Laakkonen, L. J.; Guarnieri, F.; Perlman, J. H.; Gershengorn, M. C.; Osman, R. *Biochemistry* **1996**, *35*, 7651.
(245) Perlman, J. H.; Laakkonen, L.; Guarnieri, F.; Osman, R.; Gershengorn, M. C. *Biochemistry* **1996**, *35*, 7643.
(246) Sylte, I.; Edvardsen, O.; Dahl, S. G. *Protein Eng.* **1996**, *9*, 149.
(247) Wetzel, J. M.; Salon, J. A.; Tamm, J. A.; Forray, C.; Craig, D.; Nakanishi, H.; Cui, W.; Vaysse, P. J.; Chiu, G.; Weinshank, R. L.; Hartig, P. R.; Branchek, T. A.; Gluchowski, C. *Recept. Channels* **1996**, *4*, 165.
(248) Bourdon, H.; Trumpp-Kallmeyer, S.; Schreuder, H.; Hoflack, J.; Hibert, M.; Wermuth, C. G. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 317.
(249) Gouldson, P. R.; Reynolds, C. A. *Biochem. Soc. Trans.* **1997**, *25*, 1066.
(250) Jiang, Q.; Guo, D.; Lee, B. X.; Van Rhee, A. M.; Kim, Y. C.; Nicholas, R. A.; Schachter, J. B.; Harden, T. K.; Jacobson, K. A. *Mol. Pharmacol.* **1997**, *52*, 499.
(251) Lin, Z.; Shenker, A.; Pearlstein, R. *Protein Eng.* **1997**, *10*, 501.
(252) Perlman, J. H.; Colson, A. O.; Jain, R.; Czyzewski, B.; Cohen, L. A.; Osman, R.; Gershengorn, M. C. *Biochemistry* **1997**, *36*, 15670.
(253) Perlman, J. H.; Colson, A. O.; Wang, W.; Bence, K.; Osman, R.; Gershengorn, M. C. *J. Biol. Chem.* **1997**, *272*, 11937.
(254) Strahs, D.; Weinstein, H. *Protein Eng.* **1997**, *10*, 1019.
(255) Sylte, I.; Chilmonczyk, Z.; Dahl, S. G.; Cybulski, J.; Edvardsen, O. *J. Pharm. Pharmacol.* **1997**, *49*, 698.
(256) Colson, A. O.; Perlman, J. H.; Smolyar, A.; Gershengorn, M. C.; Osman, R. *Biophys. J.* **1998**, *74*, 1087.
(257) Colson, A. O.; Perlman, J. H.; Jinsi-Parimoo, A.; Nussenzveig, D. R.; Osman, R.; Gershengorn, M. C. *Mol. Pharmacol.* **1998**, *54*, 968.
(258) Cotte, N.; Balestre, M. N.; Phalipou, S.; Hibert, M.; Manning, M.; Barberis, C.; Mouillac, B. *J. Biol. Chem.* **1998**, *273*, 29462.
(259) Kim, Y. C.; de Zwart, M.; Chang, L.; Moro, S.; von Frijtag Drabbe Kunzel, J. K.; Melman, N.; AP, I. J.; Jacobson, K. A. *J. Med. Chem.* **1998**, *41*, 2835.
(260) Labbe-Jullie, C.; Barroso, S.; Nicolas-Eteve, D.; Reversat, J. L.; Botto, J. M.; Mazella, J.; Bernassau, J. M.; Kitabgi, P. *J. Biol. Chem.* **1998**, *273*, 16351.
(261) Moro, S.; Li, A. H.; Jacobson, K. A. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1239.
(262) Moro, S.; Guo, D.; Camaioni, E.; Boyer, J. L.; Harden, T. K.; Jacobson, K. A. *J. Med. Chem.* **1998**, *41*, 1456.
(263) Prado, G. N.; Mierke, D. F.; Pellegrini, M.; Taylor, L.; Polgar, P. *J. Biol. Chem.* **1998**, *273*, 33548.
(264) Robin-Jagerschmidt, C.; Sylte, I.; Bihoreau, C.; Hendricksen, L.; Calvet, A.; Dahl, S. G.; Benicourt, C. *Mol. Cell Endocrinol.* **1998**, *139*, 187.
(265) Sylte, I.; Robin-Jagerschmidt, C.; Bihoreau, C.; Hendricksen, L.; Calvet, A.; Benicourt, C.; Dahl, S. G. *J. Mol. Model.* **1998**, *4*, 221.
(266) Donnelly, D.; Maudsley, S.; Gent, J. P.; Moser, R. N.; Hurrell, C. R.; Findlay, J. B. *Biochem. J.* **1999**, *339*, 55.
(267) Filizola, M.; Laakkonen, L.; Loew, G. H. *Protein Eng.* **1999**, *12*, 927.
(268) Li, A. H.; Moro, S.; Forsyth, N.; Melman, N.; Ji, X. D.; Jacobson, K. A. *J. Med. Chem.* **1999**, *42*, 706.
(269) Phalipou, S.; Seyer, R.; Cotte, N.; Breton, C.; Barberis, C.; Hibert, M.; Mouillac, B. *J. Biol. Chem.* **1999**, *274*, 23316.
(270) Rivkees, S. A.; Barbhaiya, H.; AP, I. J. *J. Biol. Chem.* **1999**, *274*, 3617.
(271) Song, Z. H.; Slowey, C. A.; Hurst, D. P.; Reggio, P. H. *Mol. Pharmacol.* **1999**, *56*, 834.
(272) Sylte, I.; Andrianjara, C. R.; Calvet, A.; Pascal, Y.; Dahl, S. G. *Bioorg. Med. Chem.* **1999**, *7*, 2737.
(273) Tao, Q.; McAllister, S. D.; Andreassi, J.; Nowell, K. W.; Cabral, G. A.; Hurst, D. P.; Bachtel, K.; Ekman, M. C.; Reggio, P. H.; Abood, M. E. *Mol. Pharmacol.* **1999**, *55*, 605.
(274) Moro, S.; Hoffmann, C.; Jacobson, K. A. *Biochemistry* **1999**, *38*, 3498.
(275) Salminen, T.; Varis, M.; Nyronen, T.; Pihlavisto, M.; Hoffren, A. M.; Lonnberg, T.; Marjamaki, A.; Frang, H.; Savola, J. M.; Scheinin, M.; Johnson, M. S. *J. Biol. Chem.* **1999**, *274*, 23405.
(276) Kristiansen, K.; Kroeze, W. K.; Willins, D. L.; Gelber, E. I.; Savage, J. E.; Glennon, R. A.; Roth, B. L. *J. Pharmacol. Exp. Ther.* **2000**, *293*, 735.
(277) Nandanan, E.; Jang, S. Y.; Moro, S.; Kim, H. O.; Siddiqui, M. A.; Russ, P.; Marquez, V. E.; Busson, R.; Herdewijn, P.; Harden, T. K.; Boyer, J. L.; Jacobson, K. A. *J. Med. Chem.* **2000**, *43*, 829.
(278) Shapiro, D. A.; Kristiansen, K.; Kroeze, W. K.; Roth, B. L. *Mol. Pharmacol.* **2000**, *58*, 877.
(279) Nyronen, T.; Pihlavisto, M.; Peltonen, J. M.; Hoffren, A. M.; Varis, M.; Salminen, T.; Wurster, S.; Marjamaki, A.; Kanerva, L.; Katainen, E.; Laaksonen, L.; Savola, J. M.; Scheinin, M.; Johnson, M. S. *Mol. Pharmacol.* **2001**, *59*, 1343.

(280) Blaney, F. E.; Raveglia, L. F.; Artico, M.; Cavagnera, S.; Dartois, C.; Farina, C.; Grugni, M.; Gagliardi, S.; Luttmann, M. A.; Martinelli, M.; Nadler, G. M.; Parini, C.; Petrillo, P.; Sarau, H. M.; Scheideler, M. A.; Hay, D. W.; Giardina, G. A. *J. Med. Chem.* **2001**, *44*, 1675.
(281) Bronowska, A.; Chilmonczyk, Z.; Les, A.; Edvardsen, O.; Ostensen, R.; Sylte, I. *J. Comput-Aided Mol. Des.* **2001**, *15*, 1005.
(282) Bronowska, A.; Les, A.; Chilmonczyk, Z.; Filipek, S.; Edvardsen, O.; Ostensen, R.; Sylte, I. *Bioorg. Med. Chem.* **2001**, *9*, 881.
(283) Carrieri, A.; Centeno, N. B.; Rodrigo, J.; Sanz, F.; Carotti, A. *Proteins* **2001**, *43*, 382.
(284) Da Settimo, F.; Primofiore, G.; Taliani, S.; Marini, A. M.; La Motta, C.; Novellino, E.; Greco, G.; Lavecchia, A.; Trincavelli, L.; Martini, C. *J. Med. Chem.* **2001**, *44*, 316.
(285) Labrou, N. E.; Bhogal, N.; Hurrell, C. R.; Findlay, J. B. *J. Biol. Chem.* **2001**, *276*, 37944.
(286) Sylte, I.; Bronowska, A.; Dahl, S. G. *Eur. J. Pharmacol.* **2001**, *416*, 33.
(287) Brea, J.; Rodrigo, J.; Carrieri, A.; Sanz, F.; Cadavid, M. I.; Enguix, M. J.; Villazon, M.; Mengod, G.; Caro, Y.; Masaguer, C. F.; Ravina, E.; Centeno, N. B.; Carotti, A.; Loza, M. I. *J. Med. Chem.* **2002**, *45*, 54.
(288) Novellino, E.; Abignente, E.; Cosimelli, B.; Greco, G.; Iadanza, M.; Laneri, S.; Lavecchia, A.; Rimoli, M. G.; Settimo, F. D.; Primofiore, G.; Tuscano, D.; Trincavelli, L.; Martini, C. *J. Med. Chem.* **2002**, *45*, 5030.
(289) Wilkes, B. C.; Masaro, L.; Schiller, P. W.; Carpenter, K. A. *J. Med. Chem.* **2002**, *45*, 4410.
(290) Lu, X.; Huang, W.; Worthington, S.; Drabik, P.; Osman, R.; Gershengorn, M. C. *Mol. Pharmacol.* **2004**, *66*, 1192.
(291) Vermeulen, E. S.; Schmidt, A. W.; Sprouse, J. S.; Wikstrom, H. V.; Grol, C. J. *J. Med. Chem.* **2003**, *46*, 5365.
(292) Vermeulen, E. S.; van Smeden, M.; Schmidt, A. W.; Sprouse, J. S.; Wikstrom, H. V.; Grol, C. J. *J. Med. Chem.* **2004**, *47*, 5451.
(293) Fanelli, F.; Menziani, M. C.; Cocchi, M.; De Benedetti, P. G. *THEOCHEM* **1995**, *333*, 49.
(294) Fanelli, F.; Menziani, M. C.; Scheer, A.; Cotecchia, S.; De Benedetti, P. G. *Methods* **1998**, *14*, 302.
(295) Fanelli, F.; Menziani, M. C.; De Benedetti, P. G. *Bioorg. Med. Chem.* **1995**, *3*, 1465.
(296) Fanelli, F.; Menziani, M. C.; De Benedetti, P. G. *Protein Eng.* **1995**, *8*, 557.
(297) Barlocco, D.; Fanelli, F.; Cignarella, G.; Villa, S.; Cattabeni, F.; Balduini, W.; Cimino, M.; De Benedetti, P. G. *Drug Des. Discovery* **1996**, *14*, 129.
(298) Cavalli, A.; Fanelli, F.; Taddei, C.; De Benedetti, P. G.; Cotecchia, S. *FEBS Lett.* **1996**, *399*, 9.
(299) Barlocco, D.; Cignarella, G.; Fanelli, F.; Vitalis, B.; Matyus, P.; De Benedetti, P. G. *Drug Des. Discovery* **1997**, *14*, 273.
(300) De Benedetti, P. G.; Fanelli, F.; Menziani, M. C.; Cocchi, M.; Testa, R.; Leonardi, A. *Bioorg. Med. Chem.* **1997**, *5*, 809.
(301) Reaper, C. M.; Fanelli, F.; Buckingham, S. D.; Millar, N. S.; Sattelle, D. B. *Recept. Channels* **1998**, *5*, 331.
(302) Fanelli, F.; Menziani, C.; Scheer, A.; Cotecchia, S.; De Benedetti, P. G. *Proteins* **1999**, *37*, 145.
(303) Fanelli, F.; Menziani, M. C.; Scheer, A.; Cotecchia, S.; De Benedetti, P. *Int. J. Quantum Chem.* **1999**, *73*, 71.
(304) Fanelli, F. *J. Mol. Biol.* **2000**, *296*, 1333.
(305) Scheer, A.; Costa, T.; Fanelli, F.; De Benedetti, P. G.; Mhaouty-Kodja, S.; Abuin, L.; Nenniger-Tosato, M.; Cotecchia, S. *Mol. Pharmacol.* **2000**, *57*, 219.
(306) Alkorta, I.; Du, P. *Protein Eng.* **1994**, *7*, 1231.
(307) Alkorta, I.; Loew, G. H. *Protein Eng.* **1996**, *9*, 573.
(308) Donnelly, D.; Findlay, J. B. *Curr. Opin. Struct. Biol.* **1994**, *4*, 582.
(309) Donnelly, D.; Findlay, J. B.; Blundell, T. L. *Recept. Channels* **1994**, *2*, 61.
(310) Herzyk, P.; Hubbard, R. E. *Biophys. J.* **1995**, *69*, 2419.
(311) Du, P.; Salon, J. A.; Tamm, J. A.; Hou, C.; Cui, W.; Walker, M. W.; Adham, N.; Dhanoa, D. S.; Islam, I.; Vaysse, P. J.; Dowling, B.; Shifman, Y.; Boyle, N.; Rueger, H.; Schmidlin, T.; Yamaguchi, Y.; Branchek, T. A.; Weinshank, R. L.; Gluchowski, C. *Protein Eng.* **1997**, *10*, 109.
(312) Prusis, P.; Schioth, H. B.; Muceniece, R.; Herzyk, P.; Afshar, M.; Hubbard, R. E.; Wikberg, J. E. *J. Mol. Graphics Modell.* **1997**, *15*, 307.
(313) Homan, E. J.; Wikstrom, H. V.; Grol, C. J. *Bioorg. Med. Chem.* **1999**, *7*, 1805.
(314) Peitsch, M. C.; Herzyk, P.; Wells, T. N.; Hubbard, R. E. *Recept. Channels* **1996**, *4*, 161.
(315) Czaplewski, C.; Kazmierkiewicz, R.; Ciarkowski, J. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 275.
(316) Czaplewski, C.; Pasenkiewicz-Gierula, M.; Ciarkowski, J. *J. Recept. Signal Transduction Res.* **1999**, *19*, 355.
(317) Politowska, E.; Czaplewski, C.; Ciarkowski, J. *Acta Biochim. Pol.* **1999**, *46*, 581.
(318) Rehwald, M.; Neuschafer-Rube, F.; de Vries, C.; Puschel, G. P. *FEBS Lett.* **1999**, *443*, 357.
(319) Ciarkowski, J.; Drabik, P.; Gieldon, A.; Kazmierkiewicz, R.; Slusarz, R. *Acta Biochim. Pol.* **2001**, *48*, 1203.
(320) Pogozheva, I. D.; Lomize, A. L.; Mosberg, H. I. *Biophys. J.* **1997**, *72*, 1963.
(321) Pogozheva, I. D.; Lomize, A. L.; Mosberg, H. I. *Biophys. J.* **1998**, *75*, 612.
(322) Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 325.
(323) Gieldon, A.; Kazmierkiewicz, R.; Slusarz, R.; Ciarkowski, J. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 1085.
(324) Zhang, L.; DeHaven, R. N.; Goodman, M. *Biochemistry* **2002**, *41*, 61.
(325) Orry, A. J.; Wallace, B. A. *Biophys. J.* **2000**, *79*, 3083.
(326) Filizola, M.; Perez, J. J.; Carteni-Farina, M. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 111.
(327) Filizola, M.; Carteni-Farina, M.; Perez, J. J. *J. Phys. Chem. B* **1999**, *103*, 2520.
(328) Fleishman, S. J.; Harrington, S.; Friesner, R. A.; Honig, B.; Ben-Tal, N. *Biophys. J.* **2004**, *87*, 3448.
(329) Floriano, W. B.; Vaidehi, N.; Goddard, W. A., 3rd; Singer, M. S.; Shepherd, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10712.
(330) Vaidehi, N.; Floriano, W. B.; Trabanino, R.; Hall, S. E.; Freddolino, P.; Choi, E. J.; Zamanakos, G.; Goddard, W. A., 3rd. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12622.
(331) Trabanino, R. J.; Hall, S. E.; Vaidehi, N.; Floriano, W. B.; Kam, V. W.; Goddard, W. A., 3rd. *Biophys. J.* **2004**, *86*, 1904.
(332) Schertler, G. F. *Eye* **1998**, *12 (Pt 3b)*, 504.
(333) Vriend, G. *J. Mol. Graph.* **1990**, *8*, 52.
(334) Shacham, S.; Marantz, Y.; Bar-Haim, S.; Kalid, O.; Warshaviak, D.; Avisar, N.; Inbal, B.; Heifetz, A.; Fichman, M.; Topf, M.; Naor, Z.; Noiman, S.; Becker, O. M. *Proteins* **2004**, *57*, 51.
(335) Becker, O. M.; Shacham, S.; Marantz, Y.; Noiman, S. *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 353.
(336) Sale, K.; Faulon, J. L.; Gray, G. A.; Schoeniger, J. S.; Young, M. M. *Protein Sci.* **2004**, *13*, 2613.
(337) Gether, U.; Ballesteros, J. A.; Seifert, R.; Sanders-Bush, E.; Weinstein, H.; Kobilka, B. K. *J. Biol. Chem.* **1997**, *272*, 2587.
(338) Gether, U.; Lin, S.; Ghanouni, P.; Ballesteros, J. A.; Weinstein, H.; Kobilka, B. K. *EMBO J.* **1997**, *16*, 6737.
(339) Javitch, J. A.; Ballesteros, J. A.; Weinstein, H.; Chen, J. *Biochemistry* **1998**, *37*, 998.
(340) Javitch, J. A.; Shi, L.; Simpson, M. M.; Chen, J.; Chiappa, V.; Visiers, I.; Weinstein, H.; Ballesteros, J. A. *Biochemistry* **2000**, *39*, 12190.
(341) Sansom, M. S.; Weinstein, H. *Trends Pharmacol. Sci.* **2000**, *21*, 445.
(342) Ballesteros, J. A.; Shi, L.; Javitch, J. A. *Mol. Pharmacol.* **2001**, *60*, 1.
(343) Govaerts, C.; Blanpain, C.; Deupi, X.; Ballet, S.; Ballesteros, J. A.; Wodak, S. J.; Vassart, G.; Pardo, L.; Parmentier, M. *J. Biol. Chem.* **2001**, *276*, 13217.
(344) Shi, L.; Simpson, M. M.; Ballesteros, J. A.; Javitch, J. A. *Biochemistry* **2001**, *40*, 12339.
(345) Shi, L.; Liapakis, G.; Xu, R.; Guarnieri, F.; Ballesteros, J. A.; Javitch, J. A. *J. Biol. Chem.* **2002**, *277*, 40989.
(346) Visiers, I.; Hassan, S. A.; Weinstein, H. *Protein Eng.* **2001**, *14*, 409.
(347) Barnett-Norris, J.; Hurst, D. P.; Lynch, D. L.; Guarnieri, F.; Makriyannis, A.; Reggio, P. H. *J. Med. Chem.* **2002**, *45*, 3649.
(348) Lopez-Rodriguez, M. L.; Vicente, B.; Deupi, X.; Barrondo, S.; Olivella, M.; Morcillo, M. J.; Behamu, B.; Ballesteros, J. A.; Salles, J.; Pardo, L. *Mol. Pharmacol.* **2002**, *62*, 15.
(349) Singh, R.; Hurst, D. P.; Barnett-Norris, J.; Lynch, D. L.; Reggio, P. H.; Guarnieri, F. *J. Peptide Res.* **2002**, *60*, 357.
(350) Lin, S. W.; Sakmar, T. P. *Biochemistry* **1996**, *35*, 11149.
(351) Zhang, R.; Hurst, D. P.; Barnett-Norris, J.; Reggio, P. H.; Song, Z. H. *Mol. Pharmacol.* **2005**, *68*, 69.
(352) Ballesteros, J.; Palczewski, K. *Curr. Opin. Drug Discovery Dev.* **2001**, *4*, 561.
(353) Sakmar, T. P. *Curr. Opin. Cell Biol.* **2002**, *14*, 189.
(354) Archer, E.; Maigret, B.; Escrieut, C.; Pradayrol, L.; Fourmy, D. *Trends Pharmacol. Sci.* **2003**, *24*, 36.
(355) Oliveira, L.; Hulsen, T.; Lutje Hulsik, D.; Paiva, A. C.; Vriend, G. *FEBS Lett.* **2004**, *564*, 269.
(356) Bissantz, C.; Logean, A.; Rognan, D. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1162.
(357) Moyle, W. R.; Campbell, R. K.; Rao, S. N.; Ayad, N. G.; Bernard, M. P.; Han, Y.; Wang, Y. *J. Biol. Chem.* **1995**, *270*, 20020.
(358) Bhowmick, N.; Huang, J.; Puett, D.; Isaacs, N. W.; Lapthorn, A. J. *Mol. Endocrinol.* **1996**, *10*, 1147.
(359) Smits, G.; Campillo, M.; Govaerts, C.; Janssens, V.; Richter, C.; Vassart, G.; Pardo, L.; Costagliola, S. *EMBO J.* **2003**, *22*, 2692.
(360) Moyle, W. R.; Xing, Y.; Lin, W.; Cao, D.; Myers, R. V.; Kerrigan, J. E.; Bernard, M. P. *J. Biol. Chem.* **2004**, *279*, 44442.
(361) Nunez Miguel, R.; Sanders, J.; Jeffreys, J.; Depraetere, H.; Evans, M.; Richards, T.; Blundell, T. L.; Rees Smith, B.; Furmaniak, J. *Thyroid* **2004**, *14*, 991.
(362) Fiser, A.; Do, R. K.; Sali, A. *Protein Sci.* **2000**, *9*, 1753.

(363) Mehler, E. L.; Periole, X.; Hassan, S. A.; Weinstein, H. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 841.
(364) Canals, M.; Marcellino, D.; Fanelli, F.; Ciruela, F.; de Benedetti, P.; Goldberg, S. R.; Neve, K.; Fuxe, K.; Agnati, L. F.; Woods, A. S.; Ferre, S.; Lluis, C.; Bouvier, M.; Franco, R. *J. Biol. Chem.* **2003**, *278*, 46741.
(365) Vitale, R. M.; Pedone, C.; De Benedetti, P. G.; Fanelli, F. *Proteins* **2004**, *56*, 430.
(366) Fanelli, F.; Puett, D. *Endocrine* **2002**, *18*, 285.
(367) Chothia, C.; Lesk, A. M. *EMBO J.* **1986**, *5*, 823.
(368) Jacobson, K. A.; Gao, Z. G.; Chen, A.; Barak, D.; Kim, S. A.; Lee, K.; Link, A.; Rompaey, P. V.; van Calenbergh, S.; Liang, B. T. *J. Med. Chem.* **2001**, *44*, 4125.
(369) Lopez-Rodriguez, M. L.; Morcillo, M. J.; Fernandez, E.; Rosado, M. L.; Pardo, L.; Schaper, K. *J. Med. Chem.* **2001**, *44*, 198.
(370) Lopez-Rodriguez, M. L.; Murcia, M.; Benhamu, B.; Olivella, M.; Campillo, M.; Pardo, L. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 1025.
(371) Baraldi, P. G.; Cacciari, B.; Moro, S.; Spalluto, G.; Pastorin, G.; Da Ros, T.; Klotz, K. N.; Varani, K.; Gessi, S.; Borea, P. A. *J. Med. Chem.* **2002**, *45*, 770.
(372) Filizola, M.; Weinstein, H. *Biopolymers* **2002**, *66*, 317.
(373) Furukawa, H.; Hamada, T.; Hayashi, M. K.; Haga, T.; Muto, Y.; Hirota, H.; Yokoyama, S.; Nagasawa, K.; Ishiguro, M. *Mol. Pharmacol.* **2002**, *62*, 778.
(374) Gao, Z. G.; Chen, A.; Barak, D.; Kim, S. K.; Muller, C. E.; Jacobson, K. A. *J. Biol. Chem.* **2002**, *277*, 19056.
(375) Gao, Z. G.; Kim, S. K.; Biadatti, T.; Chen, W.; Lee, K.; Barak, D.; Kim, S. G.; Johnson, C. R.; Jacobson, K. A. *J. Med. Chem.* **2002**, *45*, 4471.
(376) Lequin, O.; Bolbach, G.; Frank, F.; Convert, O.; Girault-Lagrange, S.; Chassaing, G.; Lavielle, S.; Sagan, S. *J. Biol. Chem.* **2002**, *277*, 22386.
(377) Maconi, A.; Pastorin, G.; Da Ros, T.; Spalluto, G.; Gao, Z. G.; Jacobson, K. A.; Baraldi, P. G.; Cacciari, B.; Varani, K.; Moro, S.; Borea, P. A. *J. Med. Chem.* **2002**, *45*, 3579.
(378) Moro, S.; Jacobson, K. A. *Curr. Pharm. Des.* **2002**, *8*, 2401.
(379) Paterlini, M. G. *Biophys. J.* **2002**, *83*, 3012.
(380) Shin, N.; Coates, E.; Murgolo, N. J.; Morse, K. L.; Bayne, M.; Strader, C. D.; Monsma, F. J., Jr. *Mol. Pharmacol.* **2002**, *62*, 38.
(381) Ulfers, A. L.; Piserchio, A.; Mierke, D. F. *Biopolymers* **2002**, *66*, 339.
(382) Johren, K.; Holtje, H. D. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 795.
(383) Anzini, M.; Canullo, L.; Braile, C.; Cappelli, A.; Gallelli, A.; Vomero, S.; Menziani, M. C.; De Benedetti, P. G.; Rizzo, M.; Collina, S.; Azzolina, O.; Sbacchi, M.; Ghelardini, C.; Galeotti, N. *J. Med. Chem.* **2003**, *46*, 3853.
(384) Berkhout, T. A.; Blaney, F. E.; Bridges, A. M.; Cooper, D. G.; Forbes, I. T.; Gribble, A. D.; Groot, P. H.; Hardy, A.; Ife, R. J.; Kaur, R.; Moores, K. E.; Shillito, H.; Willetts, J.; Witherington, J. *J. Med. Chem.* **2003**, *46*, 4070.
(385) Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. *Proteins* **2003**, *50*, 5.
(386) Broer, B. M.; Gurrath, M.; Holtje, H. D. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 739.
(387) Ebersole, B. J.; Visiers, I.; Weinstein, H.; Sealfon, S. C. *Mol. Pharmacol.* **2003**, *63*, 36.
(388) Gao, Z. G.; Kim, S. K.; Gross, A. S.; Chen, A.; Blaustein, J. B.; Jacobson, K. A. *Mol. Pharmacol.* **2003**, *63*, 1021.
(389) Kaltenbock, A.; Hibert, M.; Langer, T. *Recept. Channels* **2003**, *9*, 93.
(390) Kim, S. K.; Gao, Z. G.; Van Rompaey, P.; Gross, A. S.; Chen, A.; Van Calenbergh, S.; Jacobson, K. A. *J. Med. Chem.* **2003**, *46*, 4847.
(391) Tahtaoui, C.; Balestre, M. N.; Klotz, P.; Rognan, D.; Barberis, C.; Mouillac, B.; Hibert, M. *J. Biol. Chem.* **2003**, *278*, 40010.
(392) Trent, J. O.; Wang, Z. X.; Murray, J. L.; Shao, W.; Tamamura, H.; Fujii, N.; Peiper, S. C. *J. Biol. Chem.* **2003**, *278*, 47136.
(393) Xie, X. Q.; Chen, J. Z.; Billings, E. M. *Proteins* **2003**, *53*, 307.
(394) Gieldon, A.; Kazmierkiewicz, R.; Slusarz, R.; Pasenkiewicz-Gierula, M.; Ciarkowski, J. *J. Mol. Model. (Online)* **2003**, *9*, 372.
(395) Munshi, U. M.; Pogozheva, I. D.; Menon, K. M. *Biochemistry* **2003**, *42*, 3708.
(396) Shim, J. Y.; Welsh, W. J.; Howlett, A. C. *Biopolymers* **2003**, *71*, 169.
(397) Lopez-Rodriguez, M. L.; Porras, E.; Morcillo, M. J.; Benhamu, B.; Soto, L. J.; Lavandera, J. L.; Ramos, J. A.; Olivella, M.; Campillo, M.; Pardo, L. *J. Med. Chem.* **2003**, *46*, 5638.
(398) Aburi, M.; Smith, P. E. *Protein Sci.* **2004**, *13*, 1997.
(399) Bhogal, N.; Blaney, F. E.; Ingley, P. M.; Rees, J.; Findlay, J. B. *Biochemistry* **2004**, *43*, 3027.
(400) Bondensgaard, K.; Ankersen, M.; Thogersen, H.; Hansen, B. S.; Wulff, B. S.; Bywater, R. P. *J. Med. Chem.* **2004**, *47*, 888.
(401) Cappelli, A.; Pericot Mohr Gl, G.; Gallelli, A.; Rizzo, M.; Anzini, M.; Vomero, S.; Mennuni, L.; Ferrari, F.; Makovec, F.; Menziani, M. C.; De Benedetti, P. G.; Giorgi, G. *J. Med. Chem.* **2004**, *47*, 2574.
(402) Colotta, V.; Catarzi, D.; Varano, F.; Calabri, F. R.; Lenzi, O.; Filacchioni, G.; Martini, C.; Trincavelli, L.; Deflorian, F.; Moro, S. *J. Med. Chem.* **2004**, *47*, 3580.
(403) Costanzi, S.; Mamedova, L.; Gao, Z. G.; Jacobson, K. A. *J. Med. Chem.* **2004**, *47*, 5393.
(404) Giragossian, C.; Schaschke, N.; Moroder, L.; Mierke, D. F. *Biochemistry* **2004**, *43*, 2724.
(405) Gutierrez-de-Teran, H.; Centeno, N. B.; Pastor, M.; Sanz, F. *Proteins* **2004**, *54*, 705.
(406) Gutierrez-de-Teran, H.; Pastor, M.; Centeno, N. B.; Aqvist, J.; Sanz, F. *ChemBioChem* **2004**, *5*, 841.
(407) Leonardi, A.; Barlocco, D.; Montesano, F.; Cignarella, G.; Motta, G.; Testa, R.; Poggesi, E.; Seeber, M.; De Benedetti, P. G.; Fanelli, F. *J. Med. Chem.* **2004**, *47*, 1900.
(408) Rosenkilde, M. M.; Gerlach, L. O.; Jakobsen, J. S.; Skerlj, R. T.; Bridger, G. J.; Schwartz, T. W. *J. Biol. Chem.* **2004**, *279*, 3033.
(409) Wilczynski, A.; Wang, X. S.; Bauzo, R. M.; Xiang, Z.; Shaw, A. M.; Millard, W. J.; Richards, N. G.; Edison, A. S.; Haskell-Luevano, C. *J. Med. Chem.* **2004**, *47*, 5662.
(410) Wilczynski, A.; Wang, X. S.; Joseph, C. G.; Xiang, Z.; Bauzo, R. M.; Scott, J. W.; Sorensen, N. B.; Shaw, A. M.; Millard, W. J.; Richards, N. G.; Haskell-Luevano, C. *J. Med. Chem.* **2004**, *47*, 2194.
(411) Evers, A.; Klebe, G. *J. Med. Chem.* **2004**, *47*, 5381.
(412) Fowler, C. B.; Pogozheva, I. D.; LeVine, H., 3rd; Mosberg, H. I. *Biochemistry* **2004**, *43*, 8700.
(413) Han, S. J.; Hamdan, F. F.; Kim, S. K.; Jacobson, K. A.; Brichta, L.; Bloodworth, L. M.; Li, J. H.; Wess, J. *J. Biol. Chem.* **2005**, *280*, 24870.
(414) Kinsella, G. K.; Rozas, I.; Watson, G. W. *Biochem. Biophys. Res. Commun.* **2004**, *324*, 916.
(415) Kiss, R.; Kovari, Z.; Keseru, G. M. *Eur. J. Med. Chem.* **2004**, *39*, 959.
(416) Major, D. T.; Fischer, B. *J. Med. Chem.* **2004**, *47*, 4391.
(417) Mazna, P.; Obsilova, V.; Jelinkova, I.; Balik, A.; Berka, K.; Sovova, Z.; Ettrich, R.; Svoboda, P.; Obsil, T.; Teisinger, J. *J. Neurochem.* **2004**, *91*, 836.
(418) Salo, O. M.; Lahtela-Kakkonen, M.; Gynther, J.; Jarvinen, T.; Poso, A. *J. Med. Chem.* **2004**, *47*, 3048.
(419) Santos, E. L.; Pesquero, J. B.; Oliveira, L.; Paiva, A. C.; Costa-Neto, C. M. *Regul. Pept.* **2004**, *119*, 183.
(420) Xu, Y.; Liu, H.; Niu, C.; Luo, C.; Luo, X.; Shen, J.; Chen, K.; Jiang, H. *Bioorg. Med. Chem.* **2004**, *12*, 6193.
(421) Zhang, Y.; Sham, Y. Y.; Rajamani, R.; Gao, J.; Portoghese, P. S. *ChemBioChem* **2005**, *6*, 853.
(422) Allegretti, M.; Bertini, R.; Cesta, M. C.; Bizzarri, C.; Di Bitondo, R.; Di Cioccio, V.; Galliera, E.; Berdini, V.; Topai, A.; Zampella, G.; Russo, V.; Di Bello, N.; Nano, G.; Nicolini, L.; Locati, M.; Fantucci, P.; Florio, S.; Colotta, F. *J. Med. Chem.* **2005**, *48*, 4312.
(423) Ambrosio, C.; Molinari, P.; Fanelli, F.; Chuman, Y.; Sbraccia, M.; Ugur, O.; Costa, T. *J. Biol. Chem.* **2005**, *280*, 23464.
(424) Boeckler, F.; Lanig, H.; Gmeiner, P. *J. Med. Chem.* **2005**, *48*, 694.
(425) Cappellacci, L.; Franchetti, P.; Pasqualini, M.; Petrelli, R.; Vita, P.; Lavecchia, A.; Novellino, E.; Costa, B.; Martini, C.; Klotz, K. N.; Grifantini, M. *J. Med. Chem.* **2005**, *48*, 1550.
(426) Catarzi, D.; Colotta, V.; Varano, F.; Calabri, F. R.; Lenzi, O.; Filacchioni, G.; Trincavelli, L.; Martini, C.; Tralli, A.; Montopoli, C.; Moro, S. *Bioorg. Med. Chem.* **2005**, *13*, 705.
(427) Chai, B. X.; Pogozheva, I. D.; Lai, Y. M.; Li, J. Y.; Neubig, R. R.; Mosberg, H. I.; Gantz, I. *Biochemistry* **2005**, *44*, 3418.
(428) Ha, S. N.; Hey, P. J.; Ransom, R. W.; Harrell, C. M., Jr.; Murphy, K. L.; Chang, R.; Chen, T. B.; Su, D. S.; Markowitz, M. K.; Bock, M. G.; Freidinger, R. M.; Hess, F. J. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 159.
(429) Hjerde, E.; Dahl, S. G.; Sylte, I. *Eur J. Med. Chem.* **2005**, *40*, 185.
(430) Johren, K.; Holtje, H. D. *Arch. Pharm. (Weinheim)* **2005**, *338*, 260.
(431) Jongejan, A.; Bruysters, M.; Ballesteros, J.; Haaksma, E.; Bakker, R. A.; Pardo, L.; Leurs, R. *Nature Chem. Biol.* **2005**, *1*, 98.
(432) Jongejan, A.; Leurs, R. *Arch. Pharm. (Weinheim)* **2005**, *338*, 248.
(433) Lavecchia, A.; Cosconati, S.; Novellino, E. *J. Med. Chem.* **2005**, *48*, 2480.
(434) Lopez-Rodriguez, M. L.; Benhamu, B.; de la Fuente, T.; Sanz, A.; Pardo, L.; Campillo, M. *J. Med. Chem.* **2005**, *48*, 4216.
(435) Lopez-Rodriguez, M. L.; Morcillo, M. J.; Fernandez, E.; Benhamu, B.; Tejada, I.; Ayala, D.; Viso, A.; Campillo, M.; Pardo, L.; Delgado, M.; Manzanares, J.; Fuentes, J. A. *J. Med. Chem.* **2005**, *48*, 2548.
(436) Moro, S.; Braiuca, P.; Deflorian, F.; Ferrari, C.; Pastorin, G.; Cacciari, B.; Baraldi, P. G.; Varani, K.; Borea, P. A.; Spalluto, G. *J. Med. Chem.* **2005**, *48*, 152.
(437) Rivara, S.; Lorenzi, S.; Mor, M.; Plazzi, P. V.; Spadoni, G.; Bedini, A.; Tarzia, G. *J. Med. Chem.* **2005**, *48*, 4049.
(438) Soderhall, J. A.; Polymeropoulos, E. E.; Paulini, K.; Gunther, E.; Kuhne, R. *Biochem. Biophys. Res. Commun.* **2005**, *333*, 568.

(439) Urizar, E.; Montanelli, L.; Loy, T.; Bonomi, M.; Swillens, S.; Gales, C.; Bouvier, M.; Smits, G.; Vassart, G.; Costagliola, S. *EMBO J.* **2005**, *24*, 1954.
(440) Xhaard, H.; Nyronen, T.; Rantanen, V. V.; Ruuskanen, J. O.; Laurila, J.; Salminen, T.; Scheinin, M.; Johnson, M. S. *J. Struct. Biol.* **2005**, *150*, 126.
(441) Montero, C.; Campillo, N. E.; Goya, P.; Paez, J. A. *Eur J. Med. Chem.* **2005**, *40*, 75.
(442) Nemoto, W.; Imai, T.; Takahashi, T.; Kikuchi, T.; Fujita, N. *Protein J.* **2004**, *23*, 427.
(443) Kenakin, T. *FASEB J.* **2001**, *15*, 598.
(444) Kenakin, T. *Nat. Rev. Drug Discovery* **2002**, *1*, 103.
(445) De Lean, A.; Stadel, J. M.; Lefkowitz, R. J. *J. Biol. Chem.* **1980**, *255*, 7108.
(446) Costa, T.; Ogino, Y.; Munson, P. J.; Onaran, H. O.; Rodbard, D. *Mol. Pharmacol.* **1992**, *41*, 549.
(447) Samama, P.; Cotecchia, S.; Costa, T.; Lefkowitz, R. J. *J. Biol. Chem.* **1993**, *268*, 4625.
(448) Weiss, J. M.; Morgan, P. H.; Lutz, M. W.; Kenakin, T. *J. Theor. Biol.* **1996**, *178*, 151.
(449) Weiss, J. M.; Morgan, P. H.; Lutz, M. W.; Kenakin, T. *J. Theor. Biol.* **1996**, *178*, 169.
(450) Weiss, J. M.; Morgan, P. H.; Lutz, M. W.; Kenakin, T. *J. Theor. Biol.* **1996**, *181*, 381.
(451) Costa, T.; Herz, A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7321.
(452) Tian, W. N.; Duzic, E.; Lanier, S. M.; Deth, R. C. *Mol. Pharmacol.* **1994**, *45*, 524.
(453) Agnati, L. F.; Fuxe, K.; Ferre, S. *Trends Biochem. Sci.* **2005**, *30*, 188.
(454) Durroux, T. *Trends Pharmacol. Sci.* **2005**, *26*, 376.
(455) Perez, D. M.; Karnik, S. S. *Pharmacol. Rev.* **2005**, *57*, 147.
(456) Onaran, H. O.; Costa, T. *Ann. N.Y. Acad. Sci.* **1997**, *812*, 98.
(457) Onaran, H. O.; Scheer, A.; Cotecchia, S.; Costa, T. In *Handbook of Experimental Pharmacology*; Kenakin, T., Angus, J., Eds.; Springer: Heidelberg, 2000; Vol. 148.
(458) Gether, U.; Lin, S.; Kobilka, B. K. *J. Biol. Chem.* **1995**, *270*, 28268.
(459) Ghanouni, P.; Gryczynski, Z.; Steenhuis, J. J.; Lee, T. W.; Farrens, D. L.; Lakowicz, J. R.; Kobilka, B. K. *J. Biol. Chem.* **2001**, *276*, 24433.
(460) Ghanouni, P.; Steenhuis, J. J.; Farrens, D. L.; Kobilka, B. K. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5997.
(461) Baneres, J. L.; Martin, A.; Hullot, P.; Girard, J. P.; Rossi, J. C.; Parello, J. *J. Mol. Biol.* **2003**, *329*, 801.
(462) Kobilka, B. *Mol. Pharmacol.* **2004**, *65*, 1060.
(463) Swaminath, G.; Xiang, Y.; Lee, T. W.; Steenhuis, J.; Parnot, C.; Kobilka, B. K. *J. Biol. Chem.* **2004**, *279*, 686.
(464) Fanelli, F.; Themmen, A. P.; Puett, D. *IUBMB Life* **2001**, *51*, 149.
(465) Shinozaki, H.; Fanelli, F.; Liu, X.; Jaquette, J.; Nakamura, K.; Segaloff, D. L. *Mol. Endocrinol.* **2001**, *15*, 972.
(466) Ascoli, M.; Fanelli, F.; Segaloff, D. L. *Endocr. Rev.* **2002**, *23*, 141.
(467) Kjelsberg, M. A.; Cotecchia, S.; Ostrowski, J.; Caron, M. G.; Lefkowitz, R. J. *J. Biol. Chem.* **1992**, *267*, 1430.
(468) Periole, X.; Ceruso, M. A.; Mehler, E. L. *Biochemistry* **2004**, *43*, 6858.
(469) Huang, P.; Li, J.; Chen, C.; Visiers, I.; Weinstein, H.; Liu-Chen, L. Y. *Biochemistry* **2001**, *40*, 13501.
(470) Capra, V.; Veltri, A.; Foglia, C.; Crimaldi, L.; Habib, A.; Parenti, M.; Rovati, G. E. *Mol. Pharmacol.* **2004**, *66*, 880.
(471) Jones, D. T. *Curr. Opin. Struct. Biol.* **1997**, *7*, 377.
(472) Lazaridis, T. *Proteins* **2003**, *52*, 176.
(473) Bikker, J. A.; Trumpp-Kallmeyer, S.; Humblet, C. *J. Med. Chem.* **1998**, *41*, 2911.
(474) Gershengorn, M. C.; Osman, R. *Endocrinology* **2001**, *142*, 2.
(475) Klabunde, T.; Hessler, G. *ChemBioChem* **2002**, *3*, 928.
(476) Moro, S.; Deflorian, F.; Spalluto, G.; Pastorin, G.; Cacciari, B.; Kim, S. K.; Jacobson, K. A. *Chem. Commun. (Cambridge)* **2003**, *24*, 2949.
(477) Bywater, R. P. *J. Mol. Recognit.* **2005**, *18*, 60.
(478) Dahl, S. G.; Sylte, I. *Basic Clin. Pharmacol. Toxicol.* **2005**, *96*, 151.
(479) Moro, S.; Spalluto, G.; Jacobson, K. A. *Trends Pharmacol. Sci.* **2005**, *26*, 44.
(480) Strzelczyk, A. A.; Jaronczyk, M.; Chilmonczyk, Z.; Mazurek, A. P.; Chojnacka-Wojcik, E.; Sylte, I. *Biochem. Pharmacol.* **2004**, *67*, 2219.
(481) Kamiya, Y.; Reynolds, C. A. *THEOCHEM* **1999**, *469*, 229.
(482) Filizola, M.; Carteni-Farina, M.; Perez, J. J. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 397.
(483) Freddolino, P. L.; Kalani, M. Y.; Vaidehi, N.; Floriano, W. B.; Hall, S. E.; Trabanino, R. J.; Kam, V. W.; Goddard, W. A., 3rd. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2736.
(484) Kalani, M. Y.; Vaidehi, N.; Hall, S. E.; Trabanino, R. J.; Freddolino, P. L.; Kalani, M. A.; Floriano, W. B.; Kam, V. W.; Goddard, W. A., 3rd. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3815.
(485) Hedberg, M. H.; Jansen, J. M.; Nordvall, G.; Hjorth, S.; Unelius, L.; Johansson, A. M. *J. Med. Chem.* **1996**, *39*, 3491.
(486) Hedberg, M. H.; Linnanen, T.; Jansen, J. M.; Nordvall, G.; Hjorth, S.; Unelius, L.; Johansson, A. M. *J. Med. Chem.* **1996**, *39*, 3503.
(487) Kask, K.; Berthold, M.; Kahl, U.; Nordvall, G.; Bartfai, T. *EMBO J.* **1996**, *15*, 236.
(488) Nordvall, G.; Nilsson, B. M.; Sundquist, S.; Johansson, G.; Glas, G.; Nilvebrant, L.; Hacksell, U. *Prog. Brain Res.* **1996**, *109*, 141.
(489) Nordvall, G.; Sundquist, S.; Johansson, G.; Glas, G.; Nilvebrant, L.; Hacksell, U. *J. Med. Chem.* **1996**, *39*, 3269.
(490) Archer-Lahlou, E.; Escrieut, C.; Clerc, P.; Martinez, J.; Moroder, L.; Logsdon, C.; Kopin, A.; Seva, C.; Dufresne, M.; Pradayrol, L.; Maigret, B.; Fourmy, D. *J. Biol. Chem.* **2005**, *280*, 10664.
(491) Archer-Lahlou, E.; Tikhonova, I.; Escrieut, C.; Dufresne, M.; Seva, C.; Pradayrol, L.; Moroder, L.; Maigret, B.; Fourmy, D. *J. Med. Chem.* **2005**, *48*, 180.
(492) Lopez-Rodriguez, M. L.; Murcia, M.; Benhamu, B.; Viso, A.; Campillo, M.; Pardo, L. *J. Med. Chem.* **2002**, *45*, 4806.
(493) McAllister, S. D.; Rizvi, G.; Anavi-Goffer, S.; Hurst, D. P.; Barnett-Norris, J.; Lynch, D. L.; Reggio, P. H.; Abood, M. E. *J. Med. Chem.* **2003**, *46*, 5139.
(494) Lopez-Rodriguez, M. L.; Benhamu, B.; Murcia, M.; Alvaro, E.; Campillo, M.; Pardo, L. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 515.
(495) Lopez-Rodriguez, M. L.; Morcillo, M. J.; Fernandez, E.; Benhamu, B.; Tejada, I.; Ayala, D.; Viso, A.; Olivella, M.; Pardo, L.; Delgado, M.; Manzanares, J.; Fuentes, J. A. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1429.
(496) Fowler, C. B.; Pogozheva, I. D.; Lomize, A. L.; LeVine, H., 3rd; Mosberg, H. I. *Biochemistry* **2004**, *43*, 15796.
(497) Hiramoto, T.; Nonaka, Y.; Inoue, K.; Yamamoto, T.; Omatsu-Kanbe, M.; Matsuura, H.; Gohda, K.; Fujita, N. *J. Pharmacol. Sci.* **2004**, *95*, 81.
(498) Jung, K. Y.; Kim, S. K.; Gao, Z. G.; Gross, A. S.; Melman, N.; Jacobson, K. A.; Kim, Y. C. *Bioorg. Med. Chem.* **2004**, *12*, 613.
(499) Major, D. T.; Nahum, V.; Wang, Y.; Reiser, G.; Fischer, B. *J. Med. Chem.* **2004**, *47*, 4405.
(500) Tchilibon, S.; Kim, S. K.; Gao, Z. G.; Harris, B. A.; Blaustein, J. B.; Gross, A. S.; Duong, H. T.; Melman, N.; Jacobson, K. A. *Bioorg. Med. Chem.* **2004**, *12*, 2021.
(501) Evers, A.; Klabunde, T. *J. Med. Chem.* **2005**, *48*, 1088.
(502) Boeckler, F.; Ohnmacht, U.; Lehmann, T.; Utz, W.; Hubner, H.; Gmeiner, P. *J. Med. Chem.* **2005**, *48*, 2493.
(503) Przydzial, M. J.; Pogozheva, I. D.; Bosse, K. E.; Andrews, S. M.; Tharp, T. A.; Traynor, J. R.; Mosberg, H. I. *J. Pept. Res.* **2005**, *65*, 333.
(504) Gigoux, V.; Escrieut, C.; Fehrentz, J. A.; Poirot, S.; Maigret, B.; Moroder, L.; Gully, D.; Martinez, J.; Vaysse, N.; Fourmy, D. *J. Biol. Chem.* **1999**, *274*, 20457.
(505) Gigoux, V.; Maigret, B.; Escrieut, C.; Silvente-Poirot, S.; Bouisson, M.; Fehrentz, J. A.; Moroder, L.; Gully, D.; Martinez, J.; Vaysse, N.; Fourmy, A. D. *Protein Sci.* **1999**, *8*, 2347.
(506) Gouldson, P. R.; Bywater, R. P.; Reynolds, C. A. *Biochem. Soc. Trans.* **1997**, *25*, 529S.
(507) Govaerts, C.; Lefort, A.; Costagliola, S.; Wodak, S. J.; Ballesteros, J. A.; Van Sande, J.; Pardo, L.; Vassart, G. *J. Biol. Chem.* **2001**, *276*, 22991.
(508) Urizar, E.; Claeysen, S.; Deupi, X.; Govaerts, C.; Costagliola, S.; Vassart, G.; Pardo, L. *J. Biol. Chem.* **2005**, *280*, 17135.
(509) Chollet, A.; Turcatti, G. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 209.
(510) Breton, C.; Chellil, H.; Kabbaj-Benmansour, M.; Carnazzi, E.; Seyer, R.; Phalipou, S.; Morin, D.; Durroux, T.; Zingg, H.; Barberis, C.; Mouillac, B. *J. Biol. Chem.* **2001**, *276*, 26931.
(511) Yeagle, P. L.; Alderfer, J. L.; Albert, A. D. *Biochemistry* **1995**, *34*, 14621.
(512) Yeagle, P. L.; Alderfer, J. L.; Albert, A. D. *Nat. Struct. Biol.* **1995**, *2*, 832.
(513) Yeagle, P. L.; Alderfer, J. L.; Salloum, A. C.; Ali, L.; Albert, A. D. *Biochemistry* **1997**, *36*, 3864.
(514) Yeagle, P. L.; Alderfer, J. L.; Albert, A. D. *Biochemistry* **1997**, *36*, 9649.
(515) Jimonet, P.; Jager, R. *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 325.
(516) Fanelli, F.; Menziani, M. C.; Carotti, A.; De Benedetti, P. G. *THEOCHEM* **1993**, *283*, 63.
(517) Montorsi, M.; Menziani, M. C.; Cocchi, M.; Fanelli, F.; De Benedetti, P. G. *Methods* **1998**, *14*, 239.
(518) De Benedetti, P. G.; Fanelli, F.; Menziani, M. C.; Cocchi, M. *THEOCHEM* **2000**, *503*, 1.
(519) Barlocco, D.; Cignarella, G.; Piaz, V. D.; Giovannoni, M. P.; De Benedetti, P. G.; Fanelli, F.; Montesano, F.; Poggesi, E.; Leonardi, A. *J. Med. Chem.* **2001**, *44*, 2403.
(520) De Benedetti, P. G.; Menziani, M. C.; Cocchi, M.; Fanelli, F. *THEOCHEM* **1995**, *333*, 1.
(521) Hiramoto, T.; Nemoto, W.; Kikuchi, T.; Fujita, N. *J. Protein Chem.* **2002**, *21*, 537.
(522) Becker, O. M.; Marantz, Y.; Shacham, S.; Inbal, B.; Heifetz, A.; Kalid, O.; Bar-Haim, S.; Warshaviak, D.; Fichman, M.; Noiman, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11304.

(523) Ewing, T. J.; Makino, S.; Skillman, A. G.; Kuntz, I. D. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411.
(524) Evers, A.; Klebe, G. *Angew. Chem., Int. Ed.* **2004**, *43*, 248.
(525) Klabunde, T.; Evers, A. *ChemBioChem* **2005**, *6*, 876.
(526) Cavasotto, C. N.; Orry, A. J.; Abagyan, R. A. *Proteins* **2003**, *51*, 423.
(527) Luo, X.; Zhang, D.; Weinstein, H. *Protein Eng.* **1994**, *7*, 1441.
(528) Oliveira, L.; Paiva, A. C.; Sander, C.; Vriend, G. *Trends Pharmacol. Sci.* **1994**, *15*, 170.
(529) Porter, J. E.; Hwa, J.; Perez, D. M. *J. Biol. Chem.* **1996**, *271*, 28318.
(530) Maggio, R.; Vogel, Z.; Wess, J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3103.
(531) Maggio, R.; Barbier, P.; Fornai, F.; Corsini, G. U. *J. Biol. Chem.* **1996**, *271*, 31055.
(532) White, J. H.; Wise, A.; Main, M. J.; Green, A.; Fraser, N. J.; Disney, G. H.; Barnes, A. A.; Emson, P.; Foord, S. M.; Marshall, F. H. *Nature* **1998**, *396*, 679.
(533) Kaupmann, K.; Malitschek, B.; Schuler, V.; Heid, J.; Froestl, W.; Beck, P.; Mosbacher, J.; Bischoff, S.; Kulik, A.; Shigemoto, R.; Karschin, A.; Bettler, B. *Nature* **1998**, *396*, 683.
(534) Marshall, F. H.; Jones, K. A.; Kaupmann, K.; Bettler, B. *Trends Pharmacol. Sci.* **1999**, *20*, 396.
(535) Mohler, H.; Fritschy, J. M. *Trends Pharmacol. Sci.* **1999**, *20*, 87.
(536) Margeta-Mitrovic, M.; Jan, Y. N.; Jan, L. Y. *Neuron* **2000**, *27*, 97.
(537) Kunishima, N.; Shimada, Y.; Tsuji, Y.; Sato, T.; Yamamoto, M.; Kumasaka, T.; Nakanishi, S.; Jingami, H.; Morikawa, K. *Nature* **2000**, *407*, 971.
(538) Zhu, X.; Wess, J. *Biochemistry* **1998**, *37*, 15773.
(539) Morello, J. P.; Salahpour, A.; Laperriere, A.; Bernier, V.; Arthus, M. F.; Lonergan, M.; Petaja-Repo, U.; Angers, S.; Morin, D.; Bichet, D. G.; Bouvier, M. *J. Clin. Invest.* **2000**, *105*, 887.
(540) Benkirane, M.; Jin, D. Y.; Chun, R. F.; Koup, R. A.; Jeang, K. T. *J. Biol. Chem.* **1997**, *272*, 30603.
(541) Dean, M. K.; Higgs, C.; Smith, R. E.; Bywater, R. P.; Snell, C. R.; Scott, P. D.; Upton, G. J.; Howe, T. J.; Reynolds, C. A. *J. Med. Chem.* **2001**, *44*, 4595.
(542) Filipek, S.; Krzysko, K. A.; Fotiadis, D.; Liang, Y.; Saperstein, D. A.; Engel, A.; Palczewski, K. *Photochem. Photobiol. Sci.* **2004**, *3*, 628.
(543) Hamm, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4819.
(544) Baneres, J. L.; Parello, J. *J. Mol. Biol.* **2003**, *329*, 815.
(545) Park, P. S.; Palczewski, K. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8793.
(546) Jordan, B. A.; Cvejic, S.; Devi, L. A. *Neuropsychopharmacology* **2000**, *23*, S5.
(547) Levac, B. A.; O'Dowd, B. F.; George, S. R. *Curr. Opin. Pharmacol.* **2002**, *2*, 76.
(548) Waldhoer, M.; Fong, J.; Jones, R. M.; Lunzer, M. M.; Sharma, S. K.; Kostenis, E.; Portoghese, P. S.; Whistler, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9050.
(549) Jordan, B. A.; Devi, L. A. *Nature* **1999**, *399*, 697.
(550) George, S. R.; Fan, T.; Xie, Z.; Tse, R.; Tam, V.; Varghese, G.; O'Dowd, B. F. *J. Biol. Chem.* **2000**, *275*, 26128.
(551) Novi, F.; Scarselli, M.; Corsini, G. U.; Maggio, R. *J. Biol. Chem.* **2004**, *279*, 7476.
(552) Novi, F.; Stanasila, L.; Giorgi, F.; Corsini, G. U.; Cotecchia, S.; Maggio, R. *J. Biol. Chem.* **2005**, *280*, 19768.
(553) Bohn, L. M.; Lefkowitz, R. J.; Gainetdinov, R. R.; Peppel, K.; Caron, M. G.; Lin, F. T. *Science* **1999**, *286*, 2495.
(554) Bohn, L. M.; Gainetdinov, R. R.; Lin, F. T.; Lefkowitz, R. J.; Caron, M. G. *Nature* **2000**, *408*, 720.
(555) Hebert, T. E.; Moffett, S.; Morello, J. P.; Loisel, T. P.; Bichet, D. G.; Barret, C.; Bouvier, M. *J. Biol. Chem.* **1996**, *271*, 16384.
(556) Ng, G. Y.; O'Dowd, B. F.; Lee, S. P.; Chung, H. T.; Brann, M. R.; Seeman, P.; George, S. R. *Biochem. Biophys. Res. Commun.* **1996**, *227*, 200.
(557) Guo, W.; Shi, L.; Javitch, J. A. *J. Biol. Chem.* **2003**, *278*, 4385.
(558) Hernanz-Falcon, P.; Rodriguez-Frade, J. M.; Serrano, A.; Juan, D.; del Sol, A.; Soriano, S. F.; Roncal, F.; Gomez, L.; Valencia, A.; Martinez, A. C.; Mellado, M. *Nat. Immunol.* **2004**, *5*, 216.
(559) Lemay, J.; Marullo, S.; Jockers, R.; Alizon, M.; Brelot, A. *Nat. Immunol.* **2005**, *6*, 535; author reply 535.
(560) Stanasila, L.; Perez, J. B.; Vogel, H.; Cotecchia, S. *J. Biol. Chem.* **2003**, *278*, 40239.
(561) Zeng, F.; Wess, J. *Neuropsychopharmacology* **2000**, *23*, S19.
(562) Gouldson, P. R.; Higgs, C.; Smith, R. E.; Dean, M. K.; Gkoutos, G. V.; Reynolds, C. A. *Neuropsychopharmacology* **2000**, *23*, S60.
(563) Gouldson, P. R.; Snell, C. R.; Bywater, R. P.; Higgs, C.; Reynolds, C. A. *Protein Eng.* **1998**, *11*, 1181.
(564) Pazos, F.; Helmer-Citterich, M.; Ausiello, G.; Valencia, A. *J. Mol. Biol.* **1997**, *271*, 511.
(565) Lichtarge, O.; Bourne, H. R.; Cohen, F. E. *J. Mol. Biol.* **1996**, *257*, 342.
(566) Madabushi, S.; Gross, A. K.; Philippi, A.; Meng, E. C.; Wensel, T. G.; Lichtarge, O. *J. Biol. Chem.* **2004**, *279*, 8126.
(567) Schulz, A.; Grosse, R.; Schultz, G.; Gudermann, T.; Schoneberg, T. *J. Biol. Chem.* **2000**, *275*, 2381.
(568) Bakker, R. A.; Dees, G.; Carrillo, J. J.; Booth, R. G.; Lopez-Gimenez, J. F.; Milligan, G.; Strange, P. G.; Leurs, R. *J. Pharmacol. Exp. Ther.* **2004**, *311*, 131.
(569) Thummer, R. P.; Campbell, M. P.; Dean, M. K.; Frusher, M. J.; Scott, P. D.; Reynolds, C. A. *J. Mol. Neurosci.* **2005**, *26*, 113.
(570) Filizola, M.; Olmea, O.; Weinstein, H. *Protein Eng.* **2002**, *15*, 881.
(571) Gomes, I.; Jordan, B. A.; Gupta, A.; Trapaidze, N.; Nagy, V.; Devi, L. A. *J. Neurosci.* **2000**, *20*, RC110.
(572) Cvejic, S.; Devi, L. A. *J. Biol. Chem.* **1997**, *272*, 26959.
(573) del Sol Mesa, A.; Pazos, F.; Valencia, A. *J. Mol. Biol.* **2003**, *326*, 1289.
(574) Soyer, O. S.; Dimmic, M. W.; Neubig, R. R.; Goldstein, R. A. *Biochemistry* **2003**, *42*, 14522.
(575) Nemoto, W.; Toh, H. *Proteins* **2005**, *58*, 644.
(576) Filizola, M.; Weinstein, H. *FEBS J.* **2005**, *272*, 2926.
(577) Casciari, D.; Seeber, M.; Fanelli, F., submitted for publication.
(578) Chen, R.; Li, L.; Weng, Z. *Proteins* **2003**, *52*, 80.
(579) Doyle, D. A.; Morais Cabral, J.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. *Science* **1998**, *280*, 69.
(580) Chang, G.; Spencer, R. H.; Lee, A. T.; Barclay, M. T.; Rees, D. C. *Science* **1998**, *282*, 2220.
(581) Bass, R. B.; Strop, P.; Barclay, M.; Rees, D. C. *Science* **2002**, *298*, 1582.
(582) Essen, L.; Siegert, R.; Lehmann, W. D.; Oesterhelt, D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11673.
(583) Ciruela, F.; Burgueno, J.; Casado, V.; Canals, M.; Marcellino, D.; Goldberg, S. R.; Bader, M.; Fuxe, K.; Agnati, L. F.; Lluis, C.; Franco, R.; Ferre, S.; Woods, A. S. *Anal. Chem* **2004**, *76*, 5354.
(584) Canals, M.; Burgueno, J.; Marcellino, D.; Cabello, N.; Canela, E. I.; Mallol, J.; Agnati, L.; Ferre, S.; Bouvier, M.; Fuxe, K.; Ciruela, F.; Lluis, C.; Franco, R. *J. Neurochem.* **2004**, *88*, 726.
(585) Fanelli, F. *Mol. Cell. Endocrinol.*, accepted for publication.
(586) Woolf, P. J.; Linderman, J. J. *J. Theor. Biol.* **2004**, *229*, 157.
(587) Woolf, P. J.; Linderman, J. J. *Biophys. Chem.* **2003**, *104*, 217.
(588) Clapham, D. E. *Nature* **1996**, *379*, 297.
(589) Coleman, D. E.; Sprang, S. R. *Trends Biochem. Sci.* **1996**, *21*, 41.
(590) Hamm, H. E.; Gilchrist, A. *Curr. Opin. Cell Biol.* **1996**, *8*, 189.
(591) Bunemann, M.; Frank, M.; Lohse, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 16077.
(592) Klein, S.; Reuveni, H.; Levitzki, A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3219.
(593) Bohm, A.; Gaudet, R.; Sigler, P. B. *Curr. Opin. Biotechnol.* **1997**, *8*, 480.
(594) Sunahara, R. K.; Tesmer, J. J.; Gilman, A. G.; Sprang, S. R. *Science* **1997**, *278*, 1943.
(595) Wall, M. A.; Coleman, D. E.; Lee, E.; Iniguez-Lluhi, J. A.; Posner, B. A.; Gilman, A. G.; Sprang, S. R. *Cell* **1995**, *83*, 1047.
(596) Lambright, D. G.; Sondek, J.; Bohm, A.; Skiba, N. P.; Hamm, H. E.; Sigler, P. B. *Nature* **1996**, *379*, 311.
(597) Medkova, M.; Preininger, A. M.; Yu, N. J.; Hubbell, W. L.; Hamm, H. E. *Biochemistry* **2002**, *41*, 9962.
(598) Kisselev, O. G.; Kao, J.; Ponder, J. W.; Fann, Y. C.; Gautam, N.; Marshall, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4270.
(599) Koenig, B. W.; Kontaxis, G.; Mitchell, D. C.; Louis, J. M.; Litman, B. J.; Bax, A. *J. Mol. Biol.* **2002**, *322*, 441.
(600) Kisselev, O. G.; Downs, M. A. *Structure (Cambridge)* **2003**, *11*, 367.
(601) Bourne, H. R. *Curr. Opin. Cell Biol.* **1997**, *9*, 134.
(602) Franke, R. R.; Sakmar, T. P.; Graham, R. M.; Khorana, H. G. *J. Biol. Chem.* **1992**, *267*, 14767.
(603) Klein-Seetharaman, J.; Hwa, J.; Cai, K.; Altenbach, C.; Hubbell, W. L.; Khorana, H. G. *Biochemistry* **1999**, *38*, 7938.
(604) Wess, J. *FASEB J.* **1997**, *11*, 346.
(605) Burstein, E. S.; Spalding, T. A.; Brann, M. R. *J. Biol. Chem.* **1998**, *273*, 24322.
(606) Burstein, E. S.; Spalding, T. A.; Brann, M. R. *Biochemistry* **1998**, *37*, 4052.
(607) Natochin, M.; Gasimov, K. G.; Moussaif, M.; Artemyev, N. O. *J. Biol. Chem.* **2003**, *278*, 37574.
(608) Franke, R. R.; Konig, B.; Sakmar, T. P.; Khorana, H. G.; Hofmann, K. P. *Science* **1990**, *250*, 123.
(609) Acharya, S.; Saad, Y.; Karnik, S. S. *J. Biol. Chem.* **1997**, *272*, 6519.
(610) Yamashita, T.; Terakita, A.; Shichida, Y. *J. Biol. Chem.* **2000**, *275*, 34272.
(611) Terakita, A.; Yamashita, T.; Nimbari, N.; Kojima, D.; Shichida, Y. *J. Biol. Chem.* **2002**, *277*, 40.
(612) Gilchrist, A.; Mazzoni, M. R.; Dineen, B.; Dice, A.; Linden, J.; Proctor, W. R.; Lupica, C. R.; Dunwiddie, T. V.; Hamm, H. E. *J. Biol. Chem.* **1998**, *273*, 14912.
(613) Rasenick, M. M.; Watanabe, M.; Lazarevic, M. B.; Hatta, S.; Hamm, H. E. *J. Biol. Chem.* **1994**, *269*, 21519.
(614) Garcia, P. D.; Onrust, R.; Bell, S. M.; Sakmar, T. P.; Bourne, H. R. *EMBO J.* **1995**, *14*, 4460.

(615) Neer, E. J. *Cell* **1995**, *80*, 249.
(616) Osawa, S.; Weiss, E. R. *J. Biol. Chem.* **1995**, *270*, 31052.
(617) Blahos, J., 2nd; Mary, S.; Perroy, J.; de Colle, C.; Brabet, I.; Bockaert, J.; Pin, J. P. *J. Biol. Chem.* **1998**, *273*, 25765.
(618) Hamm, H. E.; Deretic, D.; Arendt, A.; Hargrave, P. A.; Koenig, B.; Hofmann, K. P. *Science* **1988**, *241*, 832.
(619) Nishimura, S.; Kandori, H.; Maeda, A. *Biochemistry* **1998**, *37*, 15816.
(620) Martin, E. L.; Rens-Domiano, S.; Schatz, P. J.; Hamm, H. E. *J. Biol. Chem.* **1996**, *271*, 361.
(621) Aris, L.; Gilchrist, A.; Rens-Domiano, S.; Meyer, C.; Schatz, P. J.; Dratz, E. A.; Hamm, H. E. *J. Biol. Chem.* **2001**, *276*, 2333.
(622) Mazzoni, M. R.; Hamm, H. E. *J. Biol. Chem.* **1996**, *271*, 30034.
(623) Bae, H.; Anderson, K.; Flood, L. A.; Skiba, N. P.; Hamm, H. E.; Graber, S. G. *J. Biol. Chem.* **1997**, *272*, 32071.
(624) Onrust, R.; Herzmark, P.; Chi, P.; Garcia, P. D.; Lichtarge, O.; Kingsley, C.; Bourne, H. R. *Science* **1997**, *275*, 381.
(625) Natochin, M.; Granovsky, A. E.; Muradov, K. G.; Artemyev, N. O. *J. Biol. Chem.* **1999**, *274*, 7865.
(626) Cai, K.; Itoh, Y.; Khorana, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4877.
(627) Itoh, Y.; Cai, K.; Khorana, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4883.
(628) Ernst, O. P.; Meyer, C. K.; Marin, E. P.; Henklein, P.; Fu, W. Y.; Sakmar, T. P.; Hofmann, K. P. *J. Biol. Chem.* **2000**, *275*, 1937.
(629) Marin, E. P.; Krishna, A. G.; Zvyaga, T. A.; Isele, J.; Siebert, F.; Sakmar, T. P. *J. Biol. Chem.* **2000**, *275*, 1930.
(630) Herrmann, R.; Heck, M.; Henklein, P.; Kleuss, C.; Hofmann, K. P.; Ernst, O. P. *J. Biol. Chem.* **2004**, *279*, 24283.
(631) Wang, X.; Kim, S. H.; Ablonczy, Z.; Crouch, R. K.; Knapp, D. R. *Biochemistry* **2004**, *43*, 11153.
(632) Hamm, H. E.; Deretic, D.; Hofmann, K. P.; Schleicher, A.; Kohl, B. *J. Biol. Chem.* **1987**, *262*, 10831.
(633) Lichtarge, O.; Bourne, H. R.; Cohen, F. E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 7507.
(634) Tesmer, J. J.; Berman, D. M.; Gilman, A. G.; Sprang, S. R. *Cell* **1997**, *89*, 251.
(635) Oliveira, L.; Paiva, A. C.; Vriend, G. *Protein Eng*. **1999**, *12*, 1087.
(636) Mahmoudian, M. *J. Mol. Graphics* **1994**, *12*, 22.
(637) Sutcliffe, M. J.; Haneef, I.; Carney, D.; Blundell, T. L. *Protein Eng*. **1987**, *1*, 377.
(638) Sutcliffe, M. J.; Hayes, F. R.; Blundell, T. L. *Protein Eng*. **1987**, *1*, 385.
(639) Ausiello, G.; Cesareni, G.; Helmer-Citterich, M. *Proteins* **1997**, *28*, 556.
(640) Slusarz, R.; Ciarkowski, J. *Acta Biochim. Pol.* **2004**, *51*, 129.
(641) Fanelli, F.; Dell'Orco, D., submitted for publication.
(642) Freire, E. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11680.
(643) Laughlin, R. B.; Pines, D.; Schmalian, J.; Stoikovic, B. P.; Wolynes, P. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 32.

CR000095N