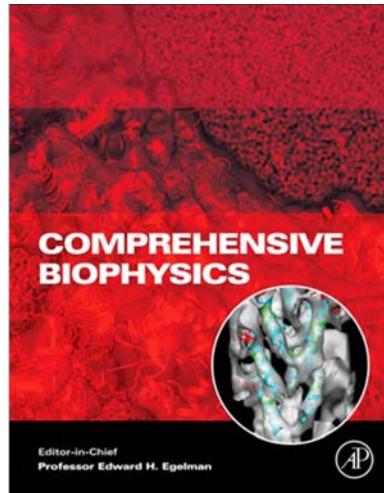


**Provided for non-commercial research and educational use only.
Not for reproduction, distribution or commercial use.**

This chapter was originally published in the *Comprehensive Biophysics*, the copy attached is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use. This includes without limitation use in instruction at your institution, distribution to specific colleagues, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited.

For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

From D. Provasi and M. Filizola, Gene Protein Coupled Receptors. In: Edward H. Egelman, editor: *Comprehensive Biophysics*, Vol 9, Simulation and Modeling, Harel Weinstein. Oxford: Academic Press, 2012. pp. 123-148.

ISBN: 978-0-12-374920-8

© Copyright 2012 Elsevier B.V.

Academic Press.

9.8 Gene Protein Coupled Receptors

D Provasi and M Filizola, Mount Sinai School of Medicine, New York, NY, USA

© 2012 Elsevier B.V. All rights reserved.

9.8.1	Structure and Dynamics of GPCR Monomers	124
9.8.1.1	Homology Modeling	125
9.8.1.1.1	Choice of suitable templates	126
9.8.1.2	Performance in docking studies	126
9.8.1.2	<i>Ab initio/de novo</i> Modeling	128
9.8.1.2.1	Modeling of the entire receptors/TM region	128
9.8.1.2.2	Modeling of loops	129
9.8.1.3	Community-Wide Assessment of GPCR Modeling	130
9.8.1.4	Simulations of Inactive Structures	131
9.8.1.4.1	Starting conformation and equilibration phase	132
9.8.1.4.2	Choice of the membrane	132
9.8.1.4.3	Choice of the force field	132
9.8.1.4.4	Simulation setup	132
9.8.1.4.5	Most simulated GPCRs	133
9.8.1.5	Modeling Activated States	134
9.8.1.5.1	Elastic networks and normal mode analyses	135
9.8.1.5.2	Systematic or MC sampling	136
9.8.1.5.3	Enhanced sampling algorithms	137
9.8.2	Structure and Dynamics of GPCR Dimers/Oligomers	137
9.8.2.1	Sequence-Based Bioinformatics Approaches	138
9.8.2.1.1	The evolutionary trace method	138
9.8.2.1.2	Correlated mutation analysis	138
9.8.2.1.3	Methods that detect tree-determinant positions	140
9.8.2.1.4	Hidden-Site class model of evolution	140
9.8.2.2	Protein-Protein Docking Methods	140
9.8.2.3	Modeling Based on Experimental Data	140
9.8.2.4	Molecular Dynamics Simulations	140
9.8.2.4.1	All-atom simulations in explicit environments	141
9.8.2.5	Coarse-Grained Simulations	141
9.8.2.6	Elastic Network Analysis	142
9.8.2.7	Dimerization-Disrupting Mutations	142
Acknowledgments		143
References		143

Abbreviations

3-D	three-dimensional	ECEPP	Empirical Conformational Energy Program for Peptides
ANM	anisotropic network model	ENM	elastic network model
CAPRI	Critical Assessment of Prediction of Interactions	GNM	Gaussian network model
CHARMM	Chemistry at HARvard Molecular Mechanics	GPCR	G protein-coupled receptor
CMA	correlated mutation analysis	GPCR-OKB	G Protein Coupled Receptor-Oligomerization Knowledge Base
CV	collective variable	Gt	G protein transducin
DEER	double electron-electron resonance	IC	intracellular
DHA	docusahexanoic acid	MC	Monte Carlo
DMPC	1,2-dimyristoyl- <i>sn</i> -glycero-3-phosphocholine	MD	molecular dynamics
DOPC	1,2-dioleoyl- <i>sn</i> -glycero-3-phosphocholine	NLX	naloxone
DPPC	1,2-dipalmitoyl- <i>sn</i> -glycero-3-phosphocholine	NMA	normal mode analysis
EC	extracellular	PDB	protein data bank
		PLPC	1-palmitoyl-2-lauroyl- <i>sn</i> -glycero-3-phosphocholine
		POPC	1-palmitoyl-2-oleoyl- <i>sn</i> -glycero-3-phosphocholine

RMSD	root mean square deviation	SDPC	1-stearoyl-2-docosahexaenoyl- <i>sn</i> -glycero-3-phosphocholine
SB	Schiff-base	SDPE	1-stearoyl-2-docosahexaenoyl- <i>sn</i> -glycero-phosphoethanolamine
SCM	subtractive correlated mutation	TM	transmembrane
SCV	scaled collective variable		
SCP-ISM	Screened Coulomb Potential-Implicit Solvent Model		

Glossary

Agonist A ligand that binds to a receptor and triggers a biological response.

Antagonist A ligand that binds to a receptor and does not produce a biological response, but rather blocks the activity of agonists and inverse agonists.

Correlated mutation analysis Technique used to predict potentially functional amino acids pairs that are distant in the primary sequence but physically interact in the native structure of homologous proteins.

Homology modeling Technique used to build structural models of a 'target' protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein (the 'template').

Inverse agonist A ligand that reverses the constitutive activity of a receptor, thus producing a biological effect opposite to that of an agonist.

Ligand A molecule able to form a stable complex with a receptor.

Monte Carlo molecular modeling Statistical sampling of the conformational space of a molecular system based on

the application of the Metropolis algorithm to generate a sequential series of states with the appropriate Boltzmann probability.

Molecular dynamics Technique used to simulate the dynamics of a molecular system based on an approximate description of the interactions between its components and a numerical integration of its equations of motion.

Network model Simplified model consisting of a set of particles linked by a network of elastic springs used to study the near-native dynamics and thermodynamics of a macromolecular system.

Oligomer Macromolecular complex composed of two or more proteins (protomers), with the minimal assembly termed a 'hetero-' or 'homo-' dimer depending on whether the protomers are different or the same.

Protomer Receptor protein forming a dimer/oligomer.

Receptor A cellular protein involved in signal transduction that can display activity in the presence or absence of a bound molecule (ligand).

9.8.1 Structure and Dynamics of GPCR Monomers

G protein-coupled receptors (GPCRs) are a large class of membrane proteins acting as receptors for a variety of extracellular (EC) chemical stimuli, and thus regulating a diverse range of intracellular (IC) signaling pathways mediated mainly, though not exclusively, by guanine nucleotide-binding proteins (G proteins). In turn, these pathways form an intricate network involved in critical biological tasks including, but not limited to, homeostasis, embryonic development, learning, and memory, through regulation of metabolic enzymes, ion channels, and transcriptional factors. Given the central role of GPCRs in the overall biology of the cell, they undoubtedly represent the most relevant group of therapeutic targets.¹

GPCRs, whose overall molecular topology consists of seven transmembrane (TM) helices connected by three EC and three IC loops, have been shown to exist in different oligomeric states, ranging from monomers, dimers, or even higher-order oligomers (see Panetta and Greenwood (2008)² and Milligan (2008)³ for reviews). This first section of the chapter will focus on the structure and dynamics of GPCR monomers while the following section will revolve around GPCR dimers/oligomers.

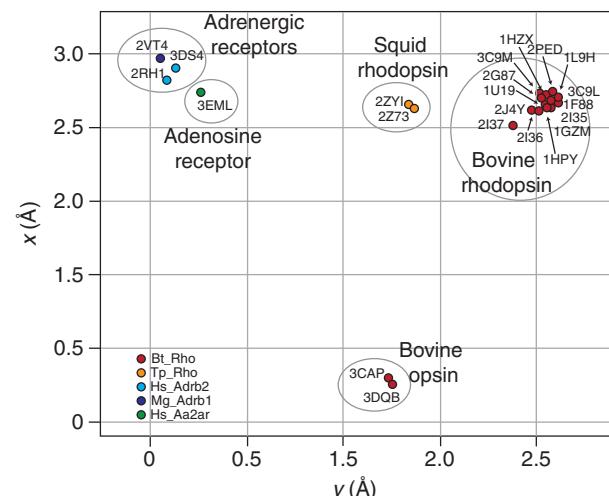
Although in recent years high-resolution crystallographic information has significantly improved our structural understanding of the GPCR family, the number of available crystal structures is still very limited. After the crystal structure of *Bos*

taurus (bovine) rhodopsin (Bt_Rho),⁴ the sole GPCR containing a covalently bound ligand, became available in 2000, only a few more crystal structures of rhodopsin corresponding to inactive or early photo-intermediate states of the receptor populated the literature in the 2000–2007 period.^{5–13} The first crystal structures of a nonrhodopsin GPCR for diffusible ligands became available in 2007, featuring an engineered form of *Homo sapiens* (human) $\beta 2$ adrenergic receptor (Hs_Adrb2) with T4 lysozyme swapped for the IC3 loop of the receptor.^{14–17} These structures were followed by the crystal structures of native *Todarodes pacificus* (squid) rhodopsin (Tp_Rho),^{18,19} ligand-free bovine opsin (both with and without the C-terminal peptide of the G α subunit),^{20,21} and two additional nonrhodopsin GPCR constructs, including a *Meleagris gallopavo* (turkey) $\beta 1$ adrenergic mutant (Mg_Adrb1),²² and an engineered *Homo sapiens* A2A adenosine receptor (Hs_Aa2ar).²³ A summary of the structural data available for intact TM regions of GPCRs is provided in Table 1.

Notably, the crystal structures listed in Table 1 present subtle differences, with the exception of the structures of ligand-free opsin, which have been suggested to contain structural features specific of activated states.²¹ The details of these structural differences have been reported in several review articles (e.g., see Filizola (2010)²⁴), and are summarized in Section 9.8.1.1.1. Figure 1 shows a multidimensional scaling plot obtained using the TM root mean square deviation (RMSD) among different GPCR structures. Thus, each

Table 1 Summary of the data available for 24 crystal structures of GPCRs.

Subfamily	Subtype	Species	Protein Database code	Notes
Rhodopsin/ opsin	Rhodopsin (Bt_Rho)	<i>Bos taurus</i> (Bovine)	1F88, 1HZX, 1L9H, 1U19, 1GZM, 2J4Y, 2PED, 3C9L, 3C9M	Inactive states
	Rhodopsin (Bt_Rho)		2I35, 2I36, 2I37; 2HPY; 2G87,	Early photointermediate states
	Opsin		3CAP	Ligand-free protein
	Opsin		3DQB	Ligand-free protein in complex with transducin C-terminal peptide
Adrenergic receptors	Rhodopsin (Tp_Rho)	<i>Todarodes pacificus</i> (Squid)	2ZIY, 2Z73	Inactive states
	$\beta 2$ -Adrenergic receptor (Hs_Adrb2)	<i>Homo sapiens</i> (Human)	2RH1, 3D4S	Engineered form with T4 lysozyme
	$\beta 2$ -Adrenergic receptor (Hs_Adrb2)		2R4R, 2R4S	In complex with immunoglobulin Fab Fragments
	$\beta 1$ -Adrenergic receptor (Mg_Adrb1)	<i>Meleagris gallopavo</i> (Turkey)	2VT4	Thermostabilized mutant
Adenosine	A2A adenosine (Hs_Aa2ar)	<i>Homo sapiens</i> (Human)	3EML	Engineered form with T4 lysozyme

**Figure 1** Two-dimensional (2-D) embedding of the structural dissimilarities among available crystal structures of G protein-coupled receptors (GPCRs) with an intact transmembrane (TM) region. The points, which represent the different crystal structures, are located in the 2-D plane so as to have their distances approximately matching the TM root mean square deviation (RMSD).

structure is represented by a point in a 2-dimensional Euclidean plane, in such a way that the Euclidean distance among different points matches as close as possible the TM RMSD. As expected, all the available structures of Bt_Rho form a compact cluster, with the exception of the photo-activated deprotonated intermediate crystal structure 2I37. The structures of Hs_Adrb2, Mg_Adrb1, and Hs_Aa2ar also form a cluster, although a less compact one. The Tp_Rho structure is midway between the bovine rhodopsin and the adrenergic receptors, and only slightly closer to the former. The most evident conformational differences in Figure 1 appear between the inactive structures (of either rhodopsin or nonrhodopsin receptors) and the ligand-free bovine opsin structures (either with or without the transducin terminal peptide), which have

been suggested to contain structural details of activated receptors.

Owing to the difficulty in obtaining crystal structures of GPCRs in either inactive or (especially) active forms, accurate structural modeling is still highly desirable for the majority of GPCRs. Over the years, structural modeling of GPCRs has mostly been obtained through homology modeling using available crystal structures as suitable templates, but various *ab initio* approaches have also been tested and put forward by a few research groups. Sections 9.8.1.1–9.8.1.3 will review both the approaches and the main issues related to the generation and the assessment of either homology or *ab initio* molecular models of GPCRs. These studies mostly refer to static inactive conformations of GPCRs. Computational methods that have been applied in the course of the years to study the dynamical properties of GPCRs in an attempt to dissect the conformational changes associated with their activation are reviewed in Sections 9.8.1.4 and 9.8.1.5.

9.8.1.1 Homology Modeling

Building a three-dimensional (3-D) molecular model of a GPCR with unknown tertiary structure (the so-called target GPCR) by homology modeling consists of assigning the atomic coordinates of the experimental 3-D structure of a related homologous GPCR (the so-called template GPCR) to equivalent positions in the target structure based on a certain sequence alignment. While homology modeling may be successful in generating useful 3-D models of highly homologous GPCR regions such as the TM regions, possible errors in these homology models stem from a poor performance in the packing of side chains, shifts in the relative positioning of the helical segments, and a poor performance of homology modeling techniques for the prediction of GPCR regions with low sequence identity, such as the IC and EC loop regions. These errors are mostly due to incorrect alignment between template and target structures,²⁵ thus making large fractions of sequence similarity, as well as the presence of highly conserved

Table 2 TM bundle structural plasticity. Pairwise root mean square deviation (RMSD) values (in Å) for the five available inactive crystal structures of G protein-coupled receptors (GPCRs), that is, Bt_Rho, Hs_Adrb2, Mg_Adrb1, Tp_Rho, and Hs_Aa2ar. Rows one to seven report the value of the structural deviation calculated for each individual TM helix, after global superposition of the two receptors in a pair. The last column refers to the RMSD values calculated for the entire TM region (TM1-7).²⁷

	<i>TM1</i>	<i>TM2</i>	<i>TM3</i>	<i>TM4</i>	<i>TM5</i>	<i>TM6</i>	<i>TM7</i>	<i>TM1-7</i>
Bt_Rho vs. Tp_Rho	1.3	2.7	1.5	1.5	4.0	0.9	1.3	1.4
Bt_Rho vs. Hs_Adrb2	3.0	1.8	1.9	2.0	3.3	1.6	1.7	1.7
Bt_Rho vs. Mg_Adrb1	3.3	2.0	1.9	2.2	3.3	1.4	1.8	1.8
Bt_Rho vs. Hs_Aa2ar	1.6	3.5	3.0	2.4	4.1	1.6	1.7	1.8
Hs_Adrb2 vs. Tp_Rho	2.8	2.6	1.3	2.1	2.6	1.7	1.4	1.6
Hs_Adrb2 vs. Mg_Adrb1	1.0	0.6	0.5	0.6	0.5	0.5	0.4	0.6
Hs_Adrb2 vs. Hs_Aa2ar	2.8	3.0	2.2	1.7	2.5	1.8	1.4	1.4
Tp_Rho vs. Mg_Adrb1	3.2	2.8	1.5	2.3	2.5	1.5	1.4	1.6
Tp_Rho vs. Hs_Aa2ar	1.0	4.6	2.5	1.9	2.8	1.6	1.1	1.4
Hs_Aa2ar vs. Mg_Adrb1	3.0	3.0	2.0	2.0	2.6	1.8	1.5	1.4
Average RMSD	2.3 ± 0.9	2.7 ± 1.0	1.8 ± 0.7	1.9 ± 0.5	2.8 ± 1.0	1.4 ± 0.4	1.4 ± 0.4	

Data from Mobaré, J. C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: Which structural template to use? *J. Med. Chem.* **2009**, *52*, 5207–5216.

residues in class A GPCRs, prerequisites to accurate modeling. In particular, for accurate homology models of class A GPCRs, it is extremely important to use fully conserved residues within the TM regions to direct the sequence alignment. Notably, these residues were used by Ballesteros and Weinstein (1995)²⁶ to define a two-component (e.g., N1.N2) generic numbering scheme for class A GPCRs to allow for easy comparison between different class A GPCR subtypes. Specifically, in this numbering scheme, the first number refers to the helix number, and the second number to the position in the helix relative to its most conserved residue, to which a value of 50 is assigned.

9.8.1.1.1 Choice of suitable templates

The Filizola lab published a quantitative analysis of the structural differences among all currently available inactive crystal structures of GPCRs.²⁷ As reported in **Table 2**, the TM regions of the two adrenergic receptors with available crystal structures (Hs_Adrb2 and Mg_Adrb1) are the most structurally similar (0.6 Å after superposition of 208 TM C_α atoms), whereas the least similar TM regions are those of rhodopsin (Bt_Rho) with either Mg_Adrb1 or the Hs_Aa2ar receptor (1.8 Å after superposition of 192 or 186 residues, respectively). To investigate the helices with the largest or smallest structural divergences, the RMSD of each TM helix of the available inactive crystal structures of GPCRs, after global TM domain superposition, was calculated. The results (see **Table 2**) pointed to TM6 and TM7 as the most structurally conserved helices (both helices with a RMSD of 1.4 ± 0.4 Å across the 10 pairs of inactive structures), and to the structures of TM2 and TM5 as the least similar ones (2.7 ± 1.0 and 2.8 ± 1.0 Å, respectively).

The same structural features can be appreciated in **Figure 2**, where a graphical comparison of the main structural differences between the receptors with known crystal structures is depicted.

Analysis also highlighted that the EC region of available inactive structures of GPCRs is more divergent than the cytoplasmic one.²⁷ In particular, the outer region of TM1 through TM4 shows a higher divergence than the inner region of the

same helices, while the inner region of helices TM5 through TM7 is more structurally different than the outer region. Of note, the divergence in the outer region of TM1-TM4 might be attributed to the specificity of ligand binding, whereas the divergence of the inner region of TM5-TM7 is probably due to crystal artifacts, although the possibility of a difference in G protein binding cannot be completely ruled out. Using 30% sequence identity as an arbitrary, yet data-supported, criterion²⁸ to obtain models of membrane proteins with a C_α RMSD of 2 Å or less with respect to the crystal structure, the authors calculated the percentage of GPCRs from the human genome that could be modeled successfully using any of the available GPCR structures as a template for homology modeling. Compared to other available GPCR crystal structures, the adrenergic receptors, which belong to the class A GPCR subfamily of amine receptors, could be used as reliable templates for a larger – yet still small – number (16–18%) of human class A GPCRs. This result is most likely due to the predominance of amine receptors in class A GPCRs, and confirms the need for more experimental structures of GPCRs. This type of analysis was pushed further, and the authors investigated which class A GPCR would be most beneficial to homology modeling of GPCRs if its structure was solved experimentally. Thus, restricting the target sequences to the nonorphan nonoliphactory class A GPCRs, somatostatin receptor 5 and 4 and high-affinity interleukin-8 receptors resulted in being the templates with the potentially highest predictive power as they might be able to produce high-resolution models for ~30% of nonorphan nonoliphactory class A GPCRs. Restricting the targets to class A orphan or olfactory receptors, the most predictive templates resulted to be the olfactory receptor family 10 (subfamily A5), the olfactory receptor family 4 (subfamily Q3), and the olfactory receptor family 1 (subfamily E1) for ~80% of these GPCR targets.

9.8.1.1.2 Performance in docking studies

Accurate structures/homology models of GPCRs are highly desirable for successful virtual screening aimed at computer drug discovery. Scientists have compared the results from automatic flexible ligand-rigid protein docking experiments

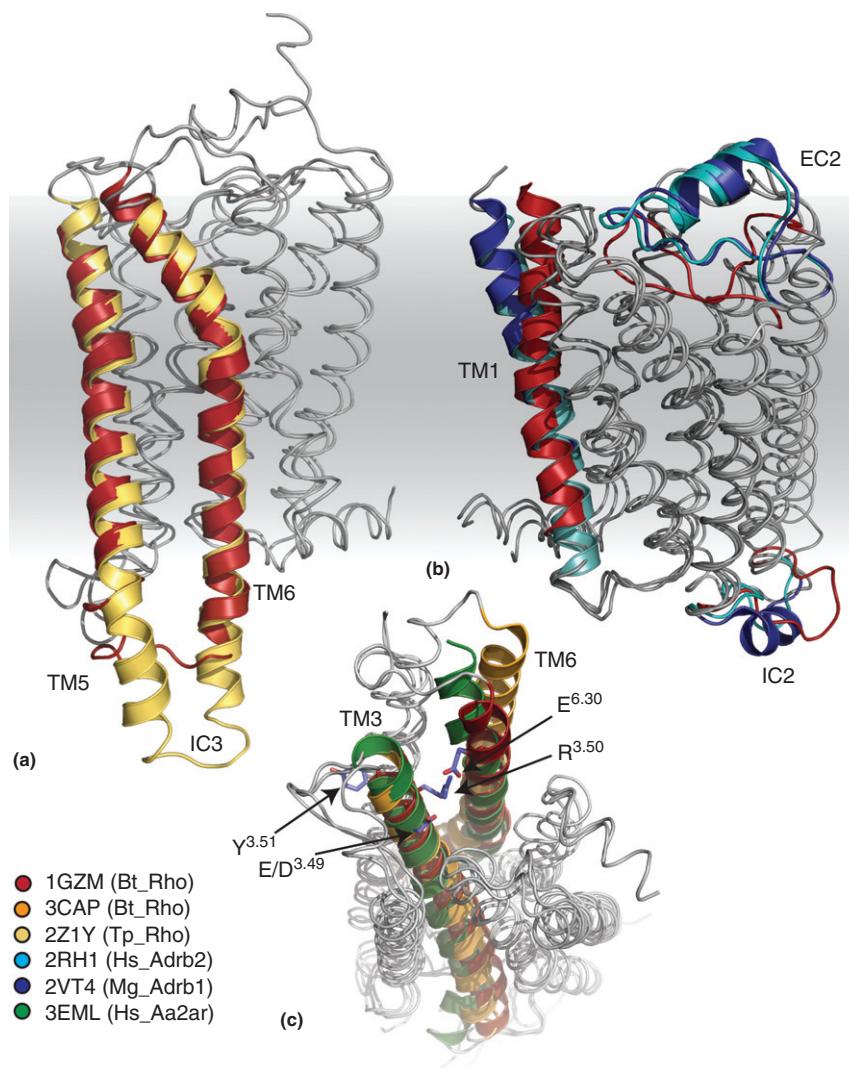


Figure 2 Representative crystal structures of G protein-coupled receptors (GPCRs). The structures are compared after superimposition of the transmembrane (TM) backbone, and their differences are highlighted by thick ribbon representations. Panel (a) shows an overlap between the two rhodopsin crystal structures from different organisms (Bt_Rho in red, Tp_Rho in yellow); panel (b) shows a structural comparison between Bt_Rho (in red) and adrenergic receptors (Hs_Adrb2 in cyan and Mg_Adrb1 in dark blue); panel (c) shows a structural comparison between Bt_Rho (in red), the opsin crystal structure (in orange), and Hs_Aa2ar (in green). The D/ERY motif at the cytoplasmic end of TM3 and the 'ionic lock' with the E residue in TM6 are represented as sticks.

carried out on crystal structures or homology models of Hs_Adrb2.^{27,29–32} For the docking of inverse agonist carazolol, an extensive set of Hs_Adrb2 models was used,²⁷ either obtained by single-template modeling using the crystal structures of Bt_Rho, Tp_Rho, Mg_Adrb1, or Hs_Aa2ar, or by multiple-template modeling using a combination of all currently available GPCR structures. For globular proteins, the use of multiple templates in homology modeling has been suggested to increase the accuracy of the resulting models, since the different templates would capture the structural variability and the evolutionary divergence of the involved proteins. However, in some cases the use of the best single model template (which of course is not possible to select *a priori*) may outperform multiple template models.³³ In an application to Hs_Adrb2, it was found that suitable docking solutions were obtained only when the Hs_Adrb2 3-D model was built

using either the experimental crystal structure or the closely related template Mg_Adrb1.²⁷ Hs_Adrb2 models obtained from templates with low sequence identity (e.g., Bt_Rho), as well as multiple templates, resulted to be less accurate when compared to the receptor crystal structure.

Independent docking experiments using the Hs_Adrb2^{29–31} or Hs_Aa2Ar³² crystal structures confirmed the increased docking performance of experimental structures with respect to homology models based on low sequence identity templates. For instance, Costanzi²⁹ tested the docking of carazolol into the Hs_Adrb2 receptor by comparing the experimental crystal structure of the adrenergic receptor with two different docking experiments using rhodopsin-based models of Hs_Adrb2. The first docking experiment used an Hs_Adrb2 model featuring an EC2 loop obtained via homology modeling using EC2 of rhodopsin as a template. In this model, the loop obstructed the

ligand entry to the binding site. The second experiment used a Hs_Adrb2 model featuring an *ab initio* EC2 loop, which more closely resembled the actual conformation of the EC2 in the Hs_Adrb2 crystal, thus resulting in an increased accessibility of the Hs_Adrb2 ligand-binding pocket. Unlike the latter case, the rhodopsin-like EC2 loop prevented correct docking of carazolol at Hs_Adrb2, confirming an essential role for both the EC2 loop and the side-chain orientations in obtaining a docked position of carazolol into Hs_Adrb2 similar to the corresponding crystal structure.

The crystal structure of Hs_Adrb2 was also tested in its ability to screen for new ligands docking at this receptor among approximately one million commercially available compounds.³⁴ A physics-based docking scoring function (with DOCK code,³⁵ version 3.5.54) was used to score all binding poses based on a ranking function that included dispersion and electrostatic interactions. The Coulomb energy of the complex was estimated using the ligand charges and the electrostatic field of the protein calculated by DelPhi and a correction for the ligand desolvation was introduced, following an approach described by Shoichet et al. (1999).³⁶ Among the 25 highest-ranking molecules tested experimentally, six compounds showed binding affinities in the low micromolar range, with the best molecule binding in the nanomolar range. In agreement with the use of an inactive crystal structure of Hs_Adrb2, five out of six molecules were found to be inverse agonists.³⁴ Notably, changes in the side-chain rotamers of residues in the binding pocket of Hs_Adrb2 (i.e., S5.43 and S5.46) were shown to yield successful docking of agonists at the carazolol-bound receptor conformation.³⁷

9.8.1.2 *Ab initio/de novo* Modeling

When considering GPCRs with very low sequence identity to any of the currently available crystal structures, *ab initio* (first-principles) or *de novo* (knowledge-based) approaches are possible alternatives to obtain reliable GPCR models. As of 2011, several *ab initio* and *de novo* approaches have been used to either build the entire GPCR structure or the TM region or the loop regions alone. These approaches are briefly described in Sections 9.8.1.2.1 and 9.8.1.2.2.

9.8.1.2.1 Modeling of the entire receptors/TM region

Complete 3-D models of more than 900 human GPCRs were obtained *ab initio* using a threading assembly refinement methodology implemented in the TASSER code by the Skolnick group in 2006.³⁸ The approach consists of identifying a template by threading, and of assembling the tertiary structure using tertiary restraints predicted in the threading phase. The code³⁸ takes into account the membrane environment by including a potential term that increases the propensity of the hydrophilic face of predicted TM regions to face the interior of the bundle. The original method was validated using rhodopsin, yielding a predicted model that exhibited a C α RMSD of 4.6 Å (2.1 Å for the TMs) with respect to the crystal structure.

The popular Rosetta approach has also been modified to build membrane protein structures. After the testing of a first coarse-grained modeling approach,³⁹ a more accurate hybrid

framework was reported where intraprotein interactions were treated with atomic resolution while the lipid-protein interactions were treated implicitly.⁴⁰ In this latter approach, the energy functional is written as a sum of weighed terms where each term represents a physically based interaction type and the weights are fitted *ex-post* in order to optimize the predicted sequence of residues that would minimize the energy for a validation set of 18 membrane protein crystal structures.⁴¹

In the protein, hydrogen bonds are treated explicitly, including a detailed description of the weak CH-O and bifurcated hydrogen bonds. Notably, the inclusion of these hydrogen bonding terms improved the performance of the energy functional significantly, as testified by a larger gap between native and non-native docked conformations. This fact underlines the role of these energy terms (unusual and less important in globular proteins) in stabilizing interfaces such as that of the glycine zipper (like the GXXXG motif at the glycophorin A interfaces), and featuring tightly packed small polar residues such as serine and threonine, that are known to play a key role in the conformation of kinked helices in TM proteins.⁴² The energy terms that describe the rotamer energy, the attractive and repulsive portions of the Lennard-Jones potential energies and knowledge-based electrostatic pair energies are identical to the Rosetta full-atom energy function for water-soluble proteins, while the solvation energy resulting from the burial of a residue in the protein is described implicitly using the solvent model IMM1. This model consists of an extension of Lazaridis' EEF1 approach to membrane proteins.⁴³ Additionally, a term that accounts for the change in solvation free energy upon transferring each isolated atom of the protein from the bulk solvent to the membrane bilayer is added to the energy functional. The membrane is represented using three planar phases: pure water and pure hydrocarbon for the nonpolar core of the membrane and an interfacial phase for the hydrocarbon-polar head group region.

The Rosetta energy functional is able to reach excellent accuracies (<2.5 Å) in the *ab initio* prediction of small TM protein structures, showing that the prediction of the packing between TM helices requires an accurate description of the interactions in the protein and between the protein and the lipid membrane. To the authors' knowledge, however, the method has not been extensively applied to the structural prediction of GPCRs. In an early test on rhodopsin, the coarse-grained method mentioned above³⁹ yielded a model with an RMSD of 10.2 Å to the native structure TM region, while no application of the all-atom improvement was reported so far. The all-atom method, however, could discriminate the near native structure of halorhodopsin from a nonnative one (this receptor is not a GPCR, but shares the same seven TM topology). Additionally, the approach was used to predict the packing of kinked helices for rhodopsin, modeling the helix far from the kinks with ideal helix fragments and docking them to native protein templates.

The authors were among the first investigators⁴⁴ to develop approaches to build *de novo* models of the TM region of GPCRs in the absence, at the time (late 1990s), of any GPCR crystal structure. In particular, the algorithm that was developed⁴⁴ used the low-resolution electron density map of rhodopsin⁴⁵ to infer relative positions and tilt angles of TM helices. This information was coupled to hydrophobicity

Table 3 Loop prediction accuracy in G protein-coupled receptors (GPCRs). Structural comparison between the best scoring results of different computational approaches for long intracellular (IC) loops of Bt_Rho (13 residues), Hs_Adbr2 (12 residues), Mg_Adbr1 (12 residues), and Hs_Aa2ar (13 residues) and for Bt_Rho IC3 (23 residues).⁵⁹

		Modloop	Dope	Loopy	Plop	Rosetta
IC2	Bt_Rho ^a	12.9 (12.0)	9.6 (8.6)	5.1 (6.3)	5.1 (6.0)	2.9 (2.6)
	Hs_Adbr2 ^b	7.9 (8.2)	9.6 (9.5)	6.0 (8.2)	7.1 (9.2)	5.0 (1.5)
	Mg_Adbr1 ^c	11.0 (8.2)	6.9 (6.9)	4.0 (4.1)	4.4 (5.7)	3.3 (4.2)
	Hs_Aa2ar ^d	4.8	6.9	5.2	3.3	1.0
IC3	Bt_Rho ^a	13.4 (19.2)	10.7 (6.0)	10.8 (9.0)	14.0 (10.3)	5.8 (3.9)

Data from Provasi, D., Bortolato, A., and Filizola, M. Exploring molecular mechanisms of ligand recognition by opioid receptors with metadynamics. *Biochemistry*, **2009**, *48*, 10020–10029. Copyright by American Chemical Society.

^aFor Bt_Rho, the first value refers to the root mean square deviation (RMSD; in Å) from the rhodopsin crystal structure 1U19, while the structural deviation from the 1GZM crystal structure is given in brackets.

^bFor Hs_Adbr2, the first value is the RMSD from the 2RH1 crystal structure and in brackets is the RMSD from the Mg_Adbr1 crystal structure (2VT4).

^cFor Mg_Adbr1, the first value is the RMSD from the crystal structure 2VT4, and in brackets is the RMSD from the Hs_Adbr2 crystal structure 2RH1.

^dFor Hs_Aa2ar, the predictions are compared to the 3ELM crystal structure.

values calculated for each ideal TM alpha-helix, which were rotated along their own axes in such a way that their hydrophobic moments were parallel to the bisector formed between three consecutive helices in the TM bundle. Similar principles have since been used in combination with an optimization strategy based on the sequential scanning of relative orientations of TM helices.⁴⁶ This approach was validated using rhodopsin (RMSD on the alpha carbon coordinates of 2.85 Å from the rhodopsin crystal structure), and applied to other GPCRs as well, including Adrb2,^{47,48} olfactory,^{49–51} dopamine D2,⁵² muscarinic M1,⁵³ and chemokine-1 (CCR1)⁵⁴ receptors.

Briefly, following the assembly of TM helices in a bundle using inferences from the low-resolution electron density map of rhodopsin, and hydrophobicity profiles for each TM helix, helical bends and kinks were first optimized either by running short (200 ps) runs on each helix in vacuum prior repositioning into the bundle, or by using the fast torsional Newton-Euler inverse mass operator method dynamics (for 500 ps). Side chains were then modeled using SCWRL 3.0. An initial systematic search of the rotational energy of each TM helix (in the context of the others kept fixed but with flexible side chains) was performed in an attempt to optimize energetically the entire TM bundle. The process was repeated until no changes were observed. Following this step, the TM bundle was inserted in an explicit lipid patch and further equilibrated.

9.8.1.2.2 Modeling of loops

As for globular proteins, homology modeling might not be very useful when applied to loop regions of GPCRs, given their low sequence identity and variable length due to insertions and deletions. Once again, *ab initio* methods might constitute a better alternative approach. A number of fairly reliable and fast *ab initio* algorithms have been reported in the literature to model the loop regions of globular proteins. These include two different loop modeling algorithms of Modeller 9v3,⁵⁵ Loopy,⁵⁶ Plop,⁵⁷ and Rosetta 2.2.0.⁵⁸ To test the ability of these methods to predict accurate structures of GPCR loops, results obtained for long loop regions of available GPCR structures using the aforementioned loop prediction algorithms have been compared.⁵⁹ Specifically, calculations were performed on the long IC2 and IC3 loops (13 and 23 amino acids, respectively) of Bt_Rho, the IC2 loop (12 amino acids)

of Hs_Adbr2, the IC2 loop (12 amino acids) of Mg_Adbr1, and the IC2 loop of Hs_Aa2ar. **Table 3** compares the lowest-energy structural predictions obtained from calculations with the corresponding structures from crystallography (see protein data bank (PDB) codes in the table legend).

Not surprisingly, the majority of these algorithms, which were optimized and validated using data sets of globular proteins, do not seem to perform very well in the case of GPCRs, as suggested by the large RMSD between predicted and experimental structures. However, the lowest-energy conformational minima generated by the Rosetta algorithm appear in general to be quite close to known crystallographic structures.

The above results are based on static comparisons with crystallographic structures. Whether the conformations identified are indeed representative of low-free energy conformations in GPCRs under physiological conditions and/or in complex with their G protein partner or other interacting signaling protein remains to be tested. Moreover, the plasticity of these regions is likely to play a substantial role in the biological properties of GPCRs, and starting conformational predictions from *ab initio* modeling methods should probably be further investigated by methods that can describe system dynamics. For instance, evidence suggests that the dynamic properties of EC loops of GPCRs might play a key role in the recognition of ligands (the impact of inaccurate loop structures on the docking performance at GPCRs was mentioned in Section 9.8.1.1.2).

Aiming to identify absolute free energy minima of GPCR loops, a Monte Carlo (MC) simulation approach using a temperature annealing protocol was proposed.⁶⁰ The method assumes that, due to the funneled topology of the free-energy surface of the loop conformations, the structures in the native ensemble are likely to be found by the random MC simulation starting from any conformation. These ideas have been implemented in a protocol that exploits the optimal search efficiency in the scaled collective variables (SCVs) framework.⁶¹ Briefly, the modeling proceeds by a series of well-defined steps. First, a segment with the sequence of the loop to be modeled is attached by its C terminal to the appropriate TM helix of the GPCR under study, while the N terminal is left free to move, and an annealing simulation allows sampling of the

loop conformations. An ensemble from these structures is then considered for the next steps. Thus, using a harmonic potential, the free N-terminal end of the loop segment is gradually brought in contact with the appropriate end of the TM helix to which the loop is supposed to be covalently connected. This is achieved by using an MC search in the space of the SCVs,⁶¹ which allows for a more exhaustive sampling. In all these simulations, the potential of the system is given by a sum of bonded and van der Waals contributions (using parameters from the Chemistry at HARvard Molecular Mechanics (CHARMM)⁶² PAR22 force field), while the Coulomb interaction is described using the electrostatic and self-energy contribution in the Screened Coulomb Potential-Implicit Solvent Model (SCP-ISM).⁶³ Finally, a nonelectrostatic cavity term proportional to the solvent accessible surface area is included. Using this computational strategy, it was shown that the crystal structure of the EC2 loop of Bt_Rho is in the native ensemble, while the calculations for the long IC3 loop showed that at least two of the crystal structures (**1U19** and **1GZM**) presented structural features of the native ensemble and at least one crystal structure (**1L9H**) did not.⁶⁴ The strategy was also applied to study the conformational preferences of the EC2 loop of the dopamine D2 receptor. Although no experimental structural data are available for this receptor to confirm the results from computation, the structural prediction was supported by available biochemical data.⁶⁵

In 2010, a new algorithm was proposed and used to predict the EC loops of all available GPCRs structures.⁶⁶ As in the aforementioned MC-based approach, this new algorithm is able to generate structural ensembles containing the available X-ray crystal structures. The modeled ensembles contained conformations of the loops with significant plasticity, suggesting possible large-scale movements that might play a role in the biological activity of the protein. Incidentally, the approach yielded conformations different from crystal structures that could still be used to reproduce the docking of ligands in the adrenergic and adenosine receptors accurately, confirming the influence of the conformations of EC loops on the binding sites of GPCR ligands.

9.8.1.3 Community-Wide Assessment of GPCR Modeling

In the 2000s, a few community-wide experiments/competitions have been established to assess the performance of the

various algorithms designed for protein structure prediction. The most widely regarded of these competitions is undoubtedly the 'Critical Assessment of Techniques for Protein Structure Prediction' (CASP),⁶⁷ which is aimed at solving the direct folding problem through accurate predictions of the tertiary structure of proteins from their amino acid sequences. Another well-known competition is the 'Critical Assessment of Prediction of Interactions' (CAPRI),⁶⁸ which focuses on algorithms aimed at predicting the molecular structure of protein complexes. In concert with the publication of the crystal structure of Hs_Aa2ar,²³ the performance of various computational algorithms for structural prediction and ligand docking was evaluated in their application to GPCRs.

In a double-blind contest termed GPCR Dock 2008,⁶⁹ which was anonymous and before the release of the experimental coordinates of Hs_Aa2ar, 29 predictors submitted 206 structural models of the receptor in complex with its high-affinity antagonist, ZM241385. The proposed models of Hs_Aa2ar were evaluated for both accuracy of the predicted overall structure and accuracy of the predicted ligand binding pose compared to the corresponding crystal structure. Specifically, a z-score (i.e., a standardized value indicating how many standard deviations each accuracy value is above the mean of the submitted set) was calculated for each model by subtracting the z-score based on the ligand RMSD from the z-score obtained by counting the number of correct contacts in the structural prediction of the receptor.

Although most of the participating algorithms were able to correctly predict the overall topology of the receptor, they failed to predict the position of the ligand and the binding interactions very accurately, suggesting that the modeling and docking stages are generally distinct steps. More specifically, the models obtained an overall excellent average RMSD of 2.8 ± 0.5 Å for the receptor TM helices C α atoms, while the loop regions (except the short IC1 loop) were predicted with a much lower accuracy (i.e., the average RMSD for EC2 in the 10 best-scoring models was 8.3 Å), increasing the overall RMSD to 4.2 ± 0.9 Å (see **Table 4**). As demonstrated by previously reported results for the Hs_Adrb2 receptor,²⁹ the accurate prediction of the long EC loops, the proper formation of disulfide bonds, and the exact placement of the helices within the conserved TM bundle are confirmed to be crucial for a good docking performance.

Table 4 Overall performance of the 10 best-ranked protocols for Hs_Aa2ar modeling and docking as resulting from the community-wide assessment experiment Dock GPCR 2008.⁶⁹ The accuracy of the protein models is expressed in terms of root mean square deviation (RMSD) of the overall C α s, of the transmembrane (TM) region, of the EL2 loop, and of the binding pocket residues. Accuracy of the ligand poses is calculated as RMSD of the ligand atoms after superposition of the protein C α s. The combined standardized score (Z-score) used in the ranking of the experiment is reported in the last column.

	Costanzi	Katritch & Abagyan	Iam & Abagyan	Davis et al.	Maignet	Jurkowski & Elofsson	Kanou	Goddard	Bologa	Olson
All	3.0	4.0	4.1	3.5	5.1	6.2	3.5	4.3	3.4	3.5
TMs	2.5	2.7	3.6	2.1	4.1	2.9	2.8	2.5	2.5	2.3
EL2	3.8	8.9	7.3	8.4	9.1	12.7	7.1	10.7	7.2	7.5
Binding pocket	3.4	3.5	3.3	4.0	7.3	3.9	6.9	4.8	3.9	5.8
ZMA	2.8	6.2	5.7	5.8	2.6	5.3	5.4	5.0	6.7	4.8
Z-score	3.02	2.76	2.42	1.46	1.23	1.04	0.91	0.78	0.72	0.69

Data from Michino, M., Abola, E., GPCR Dock 2008 participants, Brooks, C.L.I., Dixon, J.S., Moult, J., and Stevens, R.C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock, *Nature Reviews Drug Discovery*, **2008**, 8, 455–463. Copyright by Nature.

Most of the better-ranking predictions reported by the GPCR Dock 2008 community assessment used homology modeling based on one or two adrenergic receptor templates to generate the starting models of Hs_Aa2ar.⁶⁹ Among the best ten models, only one (by the Goddard lab) was generated using a combination of homology modeling (based on the Hs_Adbr2 structure) and an unpublished optimization sampling procedure for the rotational position of the helices and the side chains, similar to the spirit of another published work.⁷⁰

The best homology model (by Costanzi and co-workers)^{71,72} was obtained using the Hs_Adbr2 crystal structure as a template, and a 'manually adjusted' sequence alignment. In this model, only a short segment of the EC2 loop was modeled using the homology template, while most of it was built independent on a given structure. Most importantly, the binding poses of ZM241385 (obtained with the Glide code by Schrödinger) were selected and ranked not only according to their docking score, but also according to available mutational data for the specific system under investigation and the previous modeling studies of the author. Taken together, these results indicated that the

predictor experience and knowledge of the system were crucial in the modeling process of Hs_Aa2ar.

9.8.1.4 Simulations of Inactive Structures

While novel algorithms to enhance the conformational sampling efficiency are developed and applied to more and more biological systems, standard molecular dynamics (MD) simulations of protein systems composed of tens of thousands of particles for timescales approaching a few microseconds have become accessible, with a number of studies addressing the dynamics of GPCR monomers in an explicit solvated lipid bilayer (see Figure 3), and for microsecond-long simulations.

Several issues should, however, be kept in mind when setting up an MD simulation of a GPCR system. Some of the considerations that emerged from the literature about (i) the relevance of the initial GPCR conformation and the equilibration phase, (ii) the choice of the membrane, (iii) the choice of the force field used to describe the different molecular components of the system, and (iv) the procedure to insert the GPCR system in the membrane will be discussed.

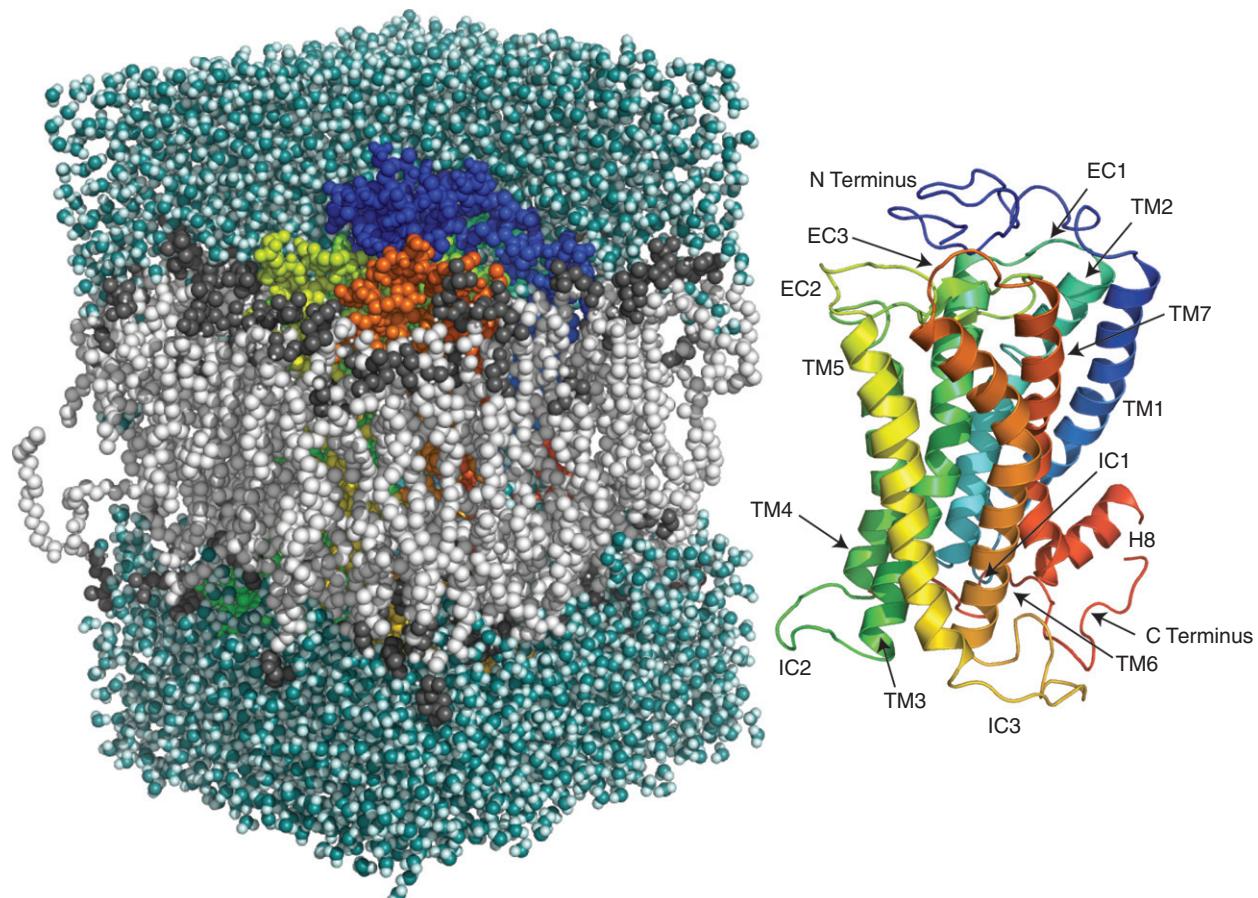


Figure 3 Monomeric form of a G protein-coupled receptor (GPCR) in an explicit lipid-water environment. A snapshot from a fully explicit molecular dynamics (MD) simulation of a GPCR monomer (left), showing a space-filling representation of the receptor atoms (colored from blue to red according to the residue number), of the lipid bilayer (light gray hydrophobic tails and dark gray polar head-groups) and of the water molecules (cyan) hydrating the membrane patch (part of the water solvating the upper side of the membrane has been removed for clarity). On the right, a cartoon representation of the GPCR highlights its main structural features.

9.8.1.4.1 Starting conformation and equilibration phase

Even if using a GPCR crystal structure as a starting point for MD simulations, an extensive equilibration phase is required (a) to obtain the correct distribution of the lipids and solvent molecules around the receptor, and (b) to correct for artifacts introduced by the artificial environment in which the protein structure was crystallized and/or by the biochemical modifications that facilitated crystallization. The results of the D. E. Shaw Research group⁷³ strongly support this observation, showing how salient structural features of GPCR inactive states that were missing in engineered forms of Hs_Adrb2 were restored by standard MD simulations of these crystal structures. Specifically, several microsecond-scale simulations of the Hs_Adrb2 crystal structure were reported where the 'ionic lock', a structural microfunctional domain known to be formed in the inactive states of most GPCRs, was reformed following MD simulations of the experimental conformation, in which it was broken.

While multiple simulations of rhodopsin were shown by Grossfield et al. (2007) to provide better convergence than a single long simulation,⁷⁴ hundreds of nanoseconds of MD simulations were required to describe rhodopsin fluctuations with a certain degree of accuracy.⁷⁵ Thus, while it is certainly difficult to give an accurate estimate of the total equilibration time needed to achieve a local equilibration of a GPCR embedded in a lipid-water environment, which might also depend on possible artifacts of the starting crystal structure, these experiments suggest that runs of the order of several hundred nanoseconds might be needed for some GPCRs.

9.8.1.4.2 Choice of the membrane

Several studies have been performed to address the role of the membrane environment and the influence of its composition on the dynamical behavior of GPCRs. For instance, Cerdomí and co-workers⁷⁶ investigated rhodopsin embedded in one-component 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), and 1-palmitoyl-2-lauroyl-sn-glycero-3-phosphocholine (PLPC) membranes, while multicomponent complex membranes were studied by Grossfield et al. (2006).⁷⁷ It is known that polyunsaturated chains, such as docosahexaenoic acid (DHA), destabilize rhodopsin and enhance the kinetics of the photocycle. MD simulations of rhodopsin⁷⁷ showed that DHA routinely forms specific interactions with a small number of residues on the protein. The presence of these tightly bound molecules weakens the interhelical packing of the GPCR bundle, providing a structural model for the effect that these lipids have on the stability of rhodopsin, and for speeding up the activation kinetics of this receptor.

Cholesterol is also known to modulate the function of GPCRs. Experiments have shown that the presence of cholesterol in the membrane stabilizes, for instance, the dark-state of rhodopsins. Although both specific interactions and non-specific effects have been suggested to play a role in the modulation of GPCR activity by cholesterol, the details of these features remain elusive. While providing evidence that the lipids near rhodopsin can preferentially reorient their unsaturated chains toward the protein surface, the 100-ns timescale simulations carried out on this receptor⁷⁸ reveal

variations in the distribution of cholesterol in the membrane. These variations complement the variations in rhodopsin transverse profile, providing a non-specific explanation for the experimental finding that cholesterol stabilizes the native dark-adapted state of rhodopsin without binding directly to the protein. Longer simulations of rhodopsins,⁷⁵ however, indicated that cholesterol can also modulate the behavior of the receptor by specific interactions with the protein. In particular, using several 100 ns long simulations, three high-cholesterol regions were identified at the EC end of TM7, at the IC extremes of TM1, TM2, and TM4, and at the IC extreme of TM2-TM3. Notably, the presence of a cholesterol strongly bound at the IC ends of TM1, TM2, and TM4, is consistent with x-ray crystallographic data on Hs_Adrb2.¹⁵

9.8.1.4.3 Choice of the force field

The improved sampling that long MD simulations ensure allows highlighting shortcomings of force-fields that would go unnoticed in shorter simulations. This was evident in DNA simulations, where long runs helped to detect major problems in the description of puckering angles, and is proving true also for membrane proteins where the need to describe simultaneously the protein and the lipid environment makes the problem more complex.⁷⁹

The ability of the CHARMM 27 force field⁸⁰ with the CMAP correction to describe both lipids and amino acids has made it a preferred choice for many research groups performing simulations on Bt_Rho and Hs_Adrb2 receptors. As an alternative to this approach, the all atom OPLS-AA force field can be used for the protein while the Berger lipids describe the environment.⁸¹ This approach has been validated by Tieleman et al. (2006)⁸² by comparing the free energy of transfer from water to cyclohexane, and by using a fully hydrated 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) bilayer with and without WALP23 peptides. One of the main advantages of this choice is the opportunity to combine an all-atom description of the protein with a united atom description of the lipid bilayer, enabling faster equilibration runs, due to the reduced number of degrees of freedom. This possibility has also become available in the CHARMM force field,⁸³ but to the best of our knowledge, it has not been validated yet for a protein embedded in the membrane. Moreover, the Berger's lipid parameters have been shown to reproduce an area per lipid close to experimental values in tensionless bilayers, in contrast with CHARMM parameters, where constant area simulations are needed in order to maintain the density of the membrane close to experimental values.

9.8.1.4.4 Simulation setup

An important step in MD simulations of membrane proteins is the preparation of a suitable starting point for the system under study, where the protein is usually embedded in a pre-equilibrated, explicit, lipid-water environment. In the majority of published simulations of GPCRs, protein insertion into the membrane is achieved by removal of overlapping lipids. An automated method⁸⁴ consisting of inserting the membrane protein in a bilayer with lipid-lipid spacing larger than a real one, subsequently reduced with alternating steps of compression and energy minimization, has been proposed. The first step of this strategy can be accomplished by either

generating a regular two dimensional lattice by replicating a single lipid or by 'blowing up' a pre-equilibrated bilayer, while the second step can be implemented as either a translation of each individual lipid or a scaling of the lipid coordinates in the membrane plane. In a practical case, to determine the sufficient number of compression and minimization steps, the surface density of the membrane can be calculated at each step, enabling comparison with a reference equilibrium density of a membrane of the same composition. While this approach clearly entails a longer preparation phase as compared with the simpler procedure of starting with an equilibrated membrane and removing the overlapping lipids, it yields an environment that matches with more accuracy the shape of the protein inclusion, reducing the subsequent length of the simulation needed to equilibrate the protein-membrane interface. This strategy has been applied for simulations of opioid receptors,⁵⁹ as well as other GPCR subtypes (e.g., rhodopsin⁸⁵).

9.8.1.4.5 Most simulated GPCRs

Rhodopsin, the protein that converts photo-stimuli into biochemical signals in the rod cell, thus enabling vision in low-light conditions, is undoubtedly the most simulated GPCR as of today. Since it was the first GPCR with an experimentally determined high-resolution structure, it has been the object of extensive MD simulation studies in several explicit membrane environments. Unlike other GPCRs, rhodopsin is covalently bound to its natural ligand, the 11-*cis* isomer of the retinal chromophore, via a protonated Schiff-base (SB) linkage to a residue in the protein TM7 helix. Being an inverse agonist, the 11-*cis*-retinal stabilizes rhodopsin in an inactive (dark) state, but undergoes a fast isomerization to 11-*trans*-retinal upon photon absorption by the receptor. This isomerization gives rise to a series of conformational changes in the receptor (see Section 9.8.1.5 for more details) leading to its activation and signal transduction mostly through coupling to G protein transducin (Gt). A number of MD simulations with timescales of a few tens of nanoseconds have been performed on the inactive rhodopsin crystal structure to investigate the early effects of retinal isomerization using the framework of standard classical MD.^{86–91} To enable study of the early events of rhodopsin activation within this framework, several researchers forced the retinal isomerization by transiently switching the dihedral energy term of the C11-C12 bond^{87,88,92,93} to a value stabilizing the 11-*trans* isomer, and analyzed tens of nanoseconds of MD simulations on these conformations.

The isomerization of the retinal chromophore and the early relaxation steps of the protein environment surrounding it were first studied with a 10-ns MD simulation,⁸⁸ using a fully explicit description of both the membrane and the protein. Following the forced isomerization of the retinal, dramatic changes in the hydrogen bonding network around the chromophore and involving E113^{3,28} (acting as a counterion of the SB base) were observed after a few hundreds of ps. The ensuing local conformational changes (including the breaking of the contacts between the beta-ionone ring of the retinal and W265^{6,48}) showed a possible mechanism by which the deformation energy internally stored was converted in non-bonded interactions between the chromophore and the protein environment. Increased mobility of some regions of the

protein were observed in the following 10 ns of simulation, chiefly in the region of TM6, supporting the hypothesis that the local conformational variations around the chromophore unlocked TM6, preparing the protein for the larger conformational changes occurring later in the activation cycle. To assess quantitatively the allosteric coupling between the retinal and TM6, a steered MD simulation was used to accelerate the dynamics of the rotation of this helix around its axis. The analysis of the work needed to induce TM6 rotation in the presence of different conformations of the retinal chromophore supported the hypothesis of a key role of the interaction between the retinal and the TM6 residue W265^{6,48} on the activation of rhodopsin. A subsequent 40 ns simulation of rhodopsin in an explicit DOPC bilayer⁸⁶ confirmed the change in the hydrogen bonding network surrounding the retinal, and its coupling through a series of local displacements with changes in the TM6 helix of the receptor. While these simulations were performed using the 1HZX crystal structure of rhodopsin, a similar methodology was employed subsequently⁹² using a refinement of the chromophore conformation based on quantum studies and a higher resolution crystal structure of rhodopsin (1L9H), which included functionally important water molecules. Notably, the more accurate description of the local energy terms of the retinal molecule allowed a study of the internal energy transfer from the isomerized bond of the chromophore to the rest of the protein through a subtle mechanism.

These calculations showed the importance of an accurate description of the chromophore and its environment. Another subtle point to be considered in simulations of rhodopsin is the description of the counterion switch mechanism that stabilizes the retinal protonated SB. Two of these models have emerged on the basis of vibrational and UV-visible spectroscopy: the protonated SB is either stabilized by E181 alone, or – as in the most recent proposals – by a complex counterion formed by E181 and E113^{3,28}. The two models were assessed using two all-atom microsecond-scale simulations of rhodopsin embedded in a complex lipid bilayer (1-stearoyl-2-docosahexaenoyl-*sn*-glycero-3-phosphocholine (SDPC), 1-stearoyl-2-docosahexaenoyl-*sn*-glycero-phosphoethanolamine (SDPE), and cholesterol) with different protonation states of E181.⁹⁴ The corresponding trajectories were used to predict 2 H NMR spectra, and their comparison with experimental data supported the complex-counterion mechanism formed by E181 and E113^{3,28}.

Coupling between local conformational changes at the retinal binding site and long-range conformational changes in rhodopsin were also observed in a 150 ns long simulation.⁹³ In this time frame, the isomerization of retinal yielded tilt and kink angle changes in some of the helices that were well beyond their fluctuation range in the equilibrium inactive state. In particular, the simulation highlighted changes occurring in the orientation of TM5 and TM6 helices, following a closer interaction between the retinal and A169^{4,58} on TM4, and a shift in the SB counterion E113^{3,28}, which also reflected changes in side chain interactions with the retinal.

Considerations similar to those reported above can be drawn from a careful qualitative analysis of the simulation trajectories. A more quantitative approach was followed in

Kong and Karplus (2007),⁸⁷ where a long atomistic simulation was used to calculate equal-time correlations of the fluctuating interaction energy of residue pairs. This was defined as the correlation function

$$C(i_1j_1|i_2j_2) = \frac{\sum_t (E_{i_1j_1}(t) - \bar{E}_{i_1j_1})(E_{i_2j_2}(t) - \bar{E}_{i_2j_2})}{\sum_t \sqrt{(E_{i_1j_1}(t) - \bar{E}_{i_1j_1})^2 (E_{i_2j_2}(t) - \bar{E}_{i_2j_2})^2}},$$

where E_{ij} is the sum of the nonbonded energy terms between residues i and j and the average is calculated over all the frames t of the trajectory. The resulting correlation function of the interaction energy was used to outline the details of a network of interactions that couples the retinal pocket of rhodopsin to the binding site of Gt, at the cytoplasmic interface. Of note, the well-known microfunctional domains at the cytoplasmic end of TM3 (i.e., the D(E)RY motif) and at the C-terminal half of TM7 (i.e., the NPXXY motif), were found to be strongly correlated (with respect to the interaction energy) with retinal, confirming the suitability of the approach. To elucidate the role of each TM helix in the signal-transduction pathway further, the authors performed the correlations of the fluctuating interaction energy on a set of simulations in which each TM helix was restrained. Based on the hypothesis that restrictions on TM helices relevant to the signal transduction process would reduce the correlation between the retinal and Gt binding sites, the approach drew attention to interdigitating side chains within TM6 and TM7 helices.

Despite the large amount of computational work on rhodopsin, none of the published simulation studies provide statistical conformational sampling and/or accurate thermodynamic information, due to the reduced timescales that can be achieved compared to the experimental timescale of rhodopsin activation and the limitations on the number of long simulations that can be carried out. Thus, although these published studies shed some light on early events of the visual cascade of rhodopsin, the molecular mechanism by which early photoactivated states of rhodopsin transform into fully active states remains unclear. An overview of the methods that have been used in the past few years to model fully active states of rhodopsin is provided in Section 9.8.1.5.

In addition to rhodopsin, the Hs_Adrb2 receptor has also been object of extensive simulations. Before the crystal structure was obtained, a number of simulations, mostly aimed at model refinement, were carried out both for the ligand-free and ligand-bound receptor. More recently, long simulations of Hs_Adrb2 have been reported in the literature, using the receptor crystal structure as a starting point. For instance, 600-ns MD simulations were carried out, which quantified receptor flexibility in a membrane bilayer environment, and showed how the binding of an agonist ligand, epinephrine (also called adrenaline), can induce conformational changes to the ligand-binding pocket, as well as the neighboring helices.⁹⁵ Long MD simulations of Hs_Adrb2 with the partial inverse agonists carazolol also appeared in the literature (e.g., see Dror et al. (2009)⁷³). These simulations restored the broken ‘ionic lock’ in the Hs_Adrb2 crystal structure simulated with or without the ligand. Notably, Hs_Adrb2 initial conformations in which an *ab initio* IC3 loop was substituted to the T4L in the crystal structure, also formed the ‘ionic lock’, at

variance with the simulations with the T4L, reconciling the experimental structures with the observation that the ‘ionic lock’ is formed in inactive structures. Along the same line, mutations interfering with the ‘ionic lock’ formation led to the separation of TM3 and TM6. Reconstitution of the ‘ionic lock’ was also achieved by long simulations of Mg_Arbdb1⁹⁶ and, most recently, Hs_Aa2ar.⁹⁷

Very long MD simulations of GPCRs with unknown experimental structure are rare, probably due to the low accuracy of the initial conformation. Results from relatively short (less than 50 ns) MD simulations in explicit membranes, were reported in the literature for several GPCRs, including a 6-ns simulation on CXCR4,⁹⁸ a 20-ns simulation of the δ opioid receptor,⁹⁹ a 2-ns simulation of the μ-opioid receptor,¹⁰⁰ and a 10-ns simulation of the 5-HT4 serotonin receptor.¹⁰¹ In contrast, results of microsecond timescale simulations have been reported for the cannabinoid CB2 receptor.¹⁰² These simulations allowed exploration of the initial stages of ligand binding to the receptor, and of the activation. Specifically, relative movement of TM3 and TM6 was observed after ligand binding, which occurred through the lipid bilayer and an entry point between TM6 and TM7. The receptor conformational changes occurring upon activation were inferred by essential dynamics analysis of the simulation trajectory. Notably, unlike an equivalent analysis on rhodopsin, a single eigenvector was found to be able to describe the motion of the CB2 receptor upon the breaking of a salt bridge and the subsequent opening of the IC surface of the receptor.

9.8.1.5 Modeling Activated States

As mentioned, all experimentally available GPCR crystal structures refer to inactive receptor forms. The absence of crystal structures of active GPCR forms has spurred a research focus on computational prediction of activated 3-D models of GPCRs. In spite of the importance of the topic and the great effort that has been poured into this investigation, both the molecular details of the activation mechanism of GPCRs and the structural features of resulting activated conformations remain elusive. Furthermore, very few predictions deriving from currently available activated models of GPCRs have been validated experimentally. At the very least, the structural similarity between these GPCR active models and the crystal structures of either the isolated native bovine opsin¹⁰³ or the complex between opsin and the C-terminal peptide derived from the α-subunit of the G protein transducin²¹ should be assessed. In fact, although it cannot be ruled out that currently available opsin structures represent stable conformations past the stage of regenerability,¹⁰⁴ these structures contain some of the structural features that have often been attributed to active states of GPCRs.¹⁰⁵ Specifically, they show a relatively large tilt of TM6, and a smaller motion of TM7 at the cytoplasmic side of the receptor, that together seem to open a cleft for G protein binding, in agreement with both earlier and recent spectroscopic studies.^{106,107} Thus, it is tempting to speculate that these opsin structures may provide a new framework for building and assessing putative models of transformation from inactive to active receptor forms for a more effective use in rational drug design.

As expected, most of the computational studies to build activated states of GPCRs have been performed on rhodopsin. While it is believed that different GPCRs share a common activation mechanism, it should be kept in mind that, unlike all the other members of the GPCR family, the activation of rhodopsin is triggered by a photo-induced conformational change in its covalently bound chromophore, and occurs through transformation of the receptor into different photo-intermediate states with characteristic ultraviolet (UV)/visible absorption spectra.

The active form of rhodopsin bound to 11-*trans*-retinal, MetII (MII), exists in two forms in a pH-dependent equilibrium, MIIa and MIIb, with only the latter capable of activating Gt.¹⁶ Two protonation-dependent switches occur during the transition from the inactive to the active MII states of rhodopsin: 1) Disruption of an inactivating salt bridge between the all-*trans*-retinylidene SB and a complex counterion composed of E113^{3,28} and E181 (EC2 loop) by internal proton transfer to the E113^{3,28} during the transition to MIIa; and 2) proton uptake at E134^{3,49} in the E(D)RY motif, following separation of the cytoplasmic ends of TM3 and TM6 relative to each other^{106,108} upon TM6 outward motion during the transition from MIIa to MIIb,¹⁰⁹ and leading to a MIIbH⁺ state. This movement is accompanied by the rupture of the salt bridge, or 'ionic lock', between the charged side chains of E247^{6,30} in TM6 and the E134^{3,49}/R135^{3,50} pair in the E(D)RY motif of rhodopsin. Breaking of this interaction has repeatedly been shown to occur upon activation of a number of different GPCRs (e.g., for a review see Kobilka and Deupi (2007)¹¹⁰).

Available crystal structures of the first photo-intermediates of rhodopsin^{7,8} showed that the TM helix bundle of the receptor does not undergo significant conformational changes during early activation events. No large structural rearrangements, but only alterations in portions of IC2 and IC3 loops, and in the density corresponding to the β -ionone ring of 11-*trans*-retinal, were also observed in the structural model of a rhodopsin intermediate containing a deprotonated SB based on low-resolution diffraction data from illuminated rhodopsin crystals.¹¹ However, the possibility cannot be ruled out that the crystal lattice limits the magnitude of the changes observed in this and other light-activated rhodopsin structures.¹¹¹ Notably, a high-resolution double electron-electron resonance (DEER) spectroscopy analysis of distance changes between pairs of nitroxide side chains introduced in helices at the cytoplasmic surface of rhodopsin suggested a 5-Å outward movement of TM6 and smaller, albeit significant, movements of TM1, TM7, and the C-terminal portion following H8 upon activation.¹¹²

In contrast to the early photoactivated states of rhodopsin, no high-resolution crystal structures of either photoactivated MIIb or the complex between MIIbH⁺ and Gt have been solved. The crystal structures of ligand-free opsin at low pH,¹⁰³ or bound to a carboxy-terminal peptide of Gt,²¹ have been proposed to resemble a MIIb active state. A dramatic rigid-body movement is observed in these structures at the cytoplasmic end of TM6, which appears to be shifted more than 6 Å outward, and closer to TM5. In this new conformation of TM6, the 'ionic lock' is broken, and new interactions are formed between R135^{3,50} and Y223^{5,58}, and between E247^{6,30} and K231^{5,66}. This rearrangement of the 'ionic lock' residues is

different from that seen in the crystal structures of a deprotonated intermediate of rhodopsin,¹¹ β 2-adrenergic receptor,¹⁴ β 1-adrenergic receptor,²² and adenosine A2A receptor.²³

Given the lack of high-resolution structural information about activated forms of rhodopsin, several investigators have applied various computational strategies (see details in the following subsections) to predict activated models of this and other GPCRs.

9.8.1.5.1 Elastic networks and normal mode analyses

The physical and mechanical principles behind functional conformational changes in proteins are complex and as yet, not fully understood. Nevertheless, two different types of motion can, in principle, contribute to such conformational changes: (i) Local elastic oscillations around a (meta)stable configuration, and (ii) more extensive plastic conformational transitions, involving local unfolding and the breaking and forming of nonlocal contacts that go beyond the quasi-harmonic picture.^{113–115} Given the relatively small scale of conformational changes associated with the activation of GPCRs (see previous section), it might be reasonable to describe GPCR activation using only a quasi-harmonic framework. This assumption led to a number of studies in which normal mode analysis (NMA) techniques were applied to derive activated states of GPCRs. Although, in general, NMA refers to the analysis of the local topology of the phase space around a (meta)stable free-energy minimum through the spectral decomposition of the Hessian matrix, and without any reference to the molecular model used to calculate the forces, most studies in this field have been applied to simplified models wherein an elastic network accounts for the interaction between structural elements of the receptor.

Such studies are based on two orders of observations (for reviews see Bahar and Rader (2005),¹¹⁶ Ma (2005),¹¹⁷ and Tama and Brooks (2006)¹¹⁸). First, structural changes upon ligand binding usually occur predominantly in directions that correspond to combinations of a few low-frequency modes.^{119–123} Second, very simple quasi-harmonic, topology-based, models can reproduce the size of fluctuations at single-residue resolution as evidenced by experimental B-factors,^{124–126} hydrogen exchange rates,¹²⁵ and comparison to atomistic potential NMA.¹²⁷ Elastic network models (ENM) are a simple and natural implementation of this idea, based on the assumption that the topology of inter-residue contacts in the native structure (which as previously outlined, is a major factor defining equilibrium dynamics) is known. The interactions between residues in close proximity (based on a cutoff distance criterion between C α atoms that usually identify the network junctions) are represented by harmonic potentials with a uniform spring constant. The successful reproduction of experimental data stems from the observation that low-frequency motions, also referred to as 'global' modes, are insensitive to the details of the models and energy parameters used in NMA, at least as long as inter-residue contacts are maintained. In an early application of these ideas to GPCRs (Figure 4), two closely related ENMs (i.e., the Gaussian network model (GNM) and the anisotropic network model (ANM)) were applied to rhodopsin monomers.¹²⁸

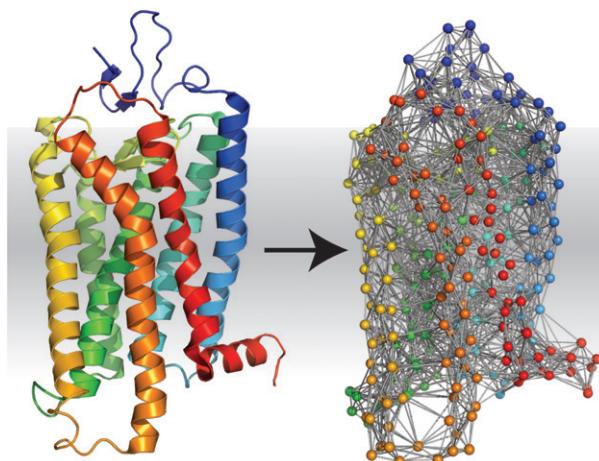


Figure 4 Elastic network model (ENM) of a G protein-coupled receptor (GPCR). The example shows the reduction of a fully atomistic representation of Bt_Rho (on the left) to its ENM (on the right). The internal dynamics of the protein can be studied by modeling the system with a set of beads positioned at the C α centers of the crystal structure and connecting beads closer than a given cutoff (10 Å in this figure) by springs with a given elastic constant.

The protein was described according to a single-parameter potential where only residues closer than a predetermined cutoff value of 0.7 nm were considered to be interacting. In the GNM case the potential is

$$V_{GNM} = \frac{\gamma}{2} \sum_{i=1}^N ||\mathbf{R}_{ij} - \mathbf{R}_{ij}^{(0)}||^2 \theta(r_c - |\mathbf{R}_{ij}^{(0)}|),$$

while in the ANM, the potential takes on a slightly different form in which the norm of the displacement is replaced by the squared variation of distance from the native conformation. In the case of the rhodopsin dark state (modeled using the crystal structure), the two models were shown to yield very similar low-frequency modes.¹²⁸ The statistical mechanics of the protein can then be obtained analytically.

In order to generate a model for an activated rhodopsin, the authors displaced the molecule along the most cooperative direction, identified as the normal mode corresponding to the low-frequency eigenvalue λ_k

$$\mathbf{R}_i^{(k)} = \mathbf{R}_i^{(0)} \pm \frac{s}{\sqrt{\lambda_k}} \mathbf{u}_i^{(k)},$$

where $\mathbf{u}_i^{(k)}$ labels the eigenvector of the system associated with this cooperative motion. An arbitrary total displacement was used, and the total energy, as estimated using an all-atom model, was used to choose between the 'positive' and 'negative' directions.

While in this application the low-frequency collective modes of the dark state of rhodopsin were used as a predictive tool (assuming that the slow modes point toward the activated state) and to determine structural details that can be compared to experimental information, NMA can also be used as an analytical tool. Thus, a protocol was presented¹²⁹ that uses constraints from the experimental observations on activated GPCRs to generate activated models of rhodopsin and

subsequently analyzes the conformational change between the two conformations using the spectral decomposition from the GNM. Since proteins are predisposed to undergo conformational fluctuations that are relevant to their biological functions,¹³⁰ several low-frequency normal modes in rhodopsins were identified as contributing to transition toward the active conformations, though none of the normal modes has a dominant contribution. Despite numerous insightful applications, ENMs have limitations, including lack of residue specificity, atomic detail, or side-chain motions. In order to combine the fast exploration of the phase space allowed by the ENMs with the accuracy of an all-atom simulation, Isin et al. (2008)¹³¹ devised a combined approach in which a small number of low-frequency and cooperative modes obtained with ANM were used to define steering directions for an all-atom MD. The displacement procedure applied in Isin et al. (2006)¹²⁸ was employed subsequently to displace the system along multiple directions. The resulting structure was used as a starting point for a new ANM calculation in an iterative manner.

Another application of multiple scale simulations, was proposed in a study similar to Niv et al. (2006).¹²⁹ When trying to generate models of the activated states of rhodopsins using constraints from experiments, it should be remembered that the experimental probes are sensitive to two different classes of structural changes: (i) The first group includes the motion associated with the isomerization of retinal and the perturbations occurring in the binding pocket upon activation, while (ii) the second group includes the global movements of the TM helices.

As in Niv et al. (2006),¹²⁹ the Lumi x-ray structure was driven¹³² toward a putative MII structure by imposing distance restraints obtained from two groups of experimental measurements on two different scales: Global restraints, based on experimental information on the overall shape of the protein backbone, and local restraints, deriving from experiments involving local distance changes. Reckoning that global restraints are most effective when applied to low-resolution models, since they could be satisfied by unphysical local deformations in all-atom models, these constraints were imposed on an ENM with few degrees of freedom. In contrast, the second set of constraints was imposed on an all-atom model.

9.8.1.5.2 Systematic or MC sampling

One of the *ab initio* systematic approaches described in Section 9.8.1.2.1 for the generation of 3-D models of GPCRs without the use of homology modeling techniques^{46–48} has been modified to build activated conformations by coupling energy estimates to a systematic scanning of the rotational position of the TM helices involved in the binding of agonists.^{133,134} In a first step, the helices of a starting model (which could be either a crystal of an inactive structure or a model) are individually rotated and the ligand-binding energy, the interhelical hydrogen bonds, and the ligand-receptor hydrogen bonds are calculated for each value of the angular position. This step allows restriction of the subsequent optimization search to only helices that are significantly perturbed by the presence of the ligand. In the second step, a systematic simultaneous search is performed spanning all combinations of orientations for the helices identified in the

first step. The systematic nature of this approach allows conformations separated from the initial state by high-energy barriers to be visited, and to build a complete energy surface in the subset of conformational changes included in the systematic search. This approach – named LITiCon – has been applied to both Adrb2¹³⁴ and rhodopsin.¹³³

In the case of Adrb2,¹³⁴ a series of agonists with different efficacy (norepinephrine, salbutamol, dopamine, and catechol), as well as the inverse agonist ICI-118551, were docked into the inactive receptor, and LITiCon was applied to the complexes. Notably, the predicted minimal energy models showed structural features in agreement with experimental evidence. Specifically, norepinephrine and dopamine were found to break the ‘ionic lock’ and to induce the ‘rotamer toggle switch’, while the weaker salbutamol and catechol ligands activated one of the switches, but not both.

The same approach was also used to study active-state structural properties of bovine rhodopsin.¹³³ The resulting model shows the well-established conformational changes of TM5 and TM6, along with the formation of inter-helical contacts stabilizing the receptor in the active state. The breaking of the ‘ionic lock’ and the activation of the toggling switch of the side chain rotamer of W265^{6,48} are also correctly predicted by this method.

Instead of being confined to the rotations around a fixed axis, a more general search can be performed by scanning the complete degrees of freedom of the helices – but yet considering them rigid bodies. This approach has been applied to rhodopsin.¹³⁵ Using a simple Empirical Conformational Energy Program for Peptides (ECEPP)/2 force field with rigid valence geometry, the interhelical potential energy was minimized in this study using a complete scan of the ‘global’ parameters describing the three translations along the three Cartesian directions and the three rotations¹³⁶ of the helices.

9.8.1.5.3 Enhanced sampling algorithms

Several approaches can be applied to reduce the computational cost of MD simulation to the levels of feasibility. For instance, enhanced sampling algorithms can be applied, as long as they are correctly taken into account when extracting the free energy from the simulations.¹³⁷ Provasi and Filizola (2010) explored the possibility of combining sampling enhancement with the careful use of restraining potentials to investigate the thermodynamics of putative activation pathways of bovine rhodopsin.⁸⁵ Specifically, the conformational transition from the crystal structure of a photo-activated deprotonated state of rhodopsin (PDB: 2I37) to the low pH crystal structure of opsin (PDB: 3CAP) in the presence of 11-trans-retinal was induced, using adiabatic biased MD simulations.¹³⁸ The thermodynamics of the system, using a path collective variable (CV) approach based on metadynamics was then explored.¹³⁹ Metadynamics is an enhanced sampling algorithm that allows the increase of the efficiency of the exploration of the phase space and the reconstruction of the free-energy surfaces of biological systems by adding a non-Markovian (history-dependent) bias to the interaction potential in the space defined by a few CVs.

During the simulation, a non-Markovian bias is added to the simulation, pushing the system away from the regions of

the phase space that it has already visited, by applying a force given by

$$\mathbf{f}_{\text{Bias}}(\mathbf{r}) = -\nabla \sum_{t' \leq t} W \exp\left(-\frac{(s(\mathbf{r}) - s_{t'})^2}{2\sigma_s^2}\right)$$

Here, the phase space is described by a single CV, s , though the algorithm works similarly in the case of multidimensional phase spaces, and the bias is formed by adding Gaussian ‘discouraging potentials’ described by the parameters W and σ_s . It can be rigorously shown that the bias potential

$$\lim_{t \rightarrow \infty} U_{\text{Bias}}(s) = \sum_{t'} W \exp\left(-\frac{(s - s_{t'})^2}{2\sigma_s^2}\right)$$

converges to an estimator of the free-energy.¹⁴⁰

Using a variant of this method termed well-tempered metadynamics,¹⁴¹ the occurrence of at least two putative active states between the crystallographic structures of rhodopsins 2I37 and 3CAP was predicted. Structural analysis of these two minima suggested ways to stabilize one state over the other,⁸⁵ which might facilitate crystallization and/or understanding of mechanisms underlying rhodopsin function.

Incidentally, metadynamics was also applied to study the binding and the possible entry pathways of a nonselective antagonist (naloxone; NLX) from the solvent into a well-accepted alkaloid binding pocket of a δ opioid receptor.⁵⁹ The delta opioid receptor was modeled using homology with the Hs_Adrb2, and simulated using a fully explicit environment modeled with a hydrated DPPC lipid bilayer. In order to describe the binding process of NLX to the receptor binding pocket, two distances were used as CVs that accounted for the position of NLX relative to the center of the binding pocket, and for the conformation of EL2 that can assume either a ‘closed’ conformation impeding the access to the binding pocket, or an ‘open’ conformation.

Moreover, in order to limit the regions of the phase space sampled by the ligand during the simulation, a steep restraint was added that maintained the ligand within a conical region centered on the protein center and comprising the binding pocket. The simulation was able to distinguish between the docked and the undocked as well as metastable bound intermediates. After correction for the restraints, the calculated free energy of binding of NLX to the receptor was fairly close to experimental values.⁵⁹

9.8.2 Structure and Dynamics of GPCR Dimers/Oligomers

Despite the recognized ability of monomeric GPCRs to activate heterotrimeric G proteins,¹⁴² several GPCR oligomers have been shown to exhibit altered pharmacological profile of possible clinical relevance (e.g., the δ-κ opioid receptor,¹⁴³ serotonin 2A-metabotropic glutamate receptor 2,¹⁴⁴ κ opioid receptor-chemokine CXCR4 receptor,¹⁴⁵ gonadotropin GnRH-GnRH receptor,¹⁴⁶ cannabinoid CB1-adenosine A2A receptor,¹⁴⁷ prostaglandin EP1-β2 adrenergic receptor,¹⁴⁸ angiotensin AT1-bradykinin B2 receptor,¹⁴⁹ and

γ -aminobutyric acid GABA_{B1}-GABA_{B2} receptor¹⁵⁰ complexes). This evidence has diverted attention from conventional drug design through inhibition of monomeric GPCRs at orthosteric ligand-binding sites to the more complex goal of designing new drugs targeting oligomeric complexes. Notwithstanding this interest in GPCR oligomers as new drug targets, our understanding of GPCR oligomerization at a molecular level remains very limited. Several important questions need to be answered to explore this new venue efficiently. For instance, it is not clear: 1) Which is the preferred structural arrangement of GPCR protomers within oligomers; 2) what molecular determinants are responsible for dimerization/oligomerization; 3) what is the stability of dimerization interfaces; 4) what is the percentage of homomers vs. heteromers; 5) which is the minimum functional unit in most GPCR oligomers; 6) how many binding pockets need to be occupied and how many protomers must undergo conformational changes associated with activation; 7) which molecular mechanisms and dynamics govern oligomer assembly or disassembly; and 8) how many G proteins are activated by a GPCR oligomer. Given the significant interest in new oligomeric GPCR targets, a compilation of current computational and experimental information on GPCR oligomers will be made available to the scientific community through GPCR-OKB (G Protein Coupled Receptor-Oligomerization Knowledge Base), an information management system that has been developed based on an interdisciplinary assessment of the requirements for a database that will be intuitive and user-friendly, yet providing the pertinent elements of information needed to study GPCR dimerization/oligomerization.¹⁵¹

The following subsections provide an overview of results from computational methods applied to GPCR dimers/oligomers with the goal of improving knowledge of their structure and dynamics, thus facilitating the design and interpretation of current physiological and pharmacological experiments on GPCRs. Specifically, 1) sequence-based and structure/docking-based approaches applied to predict likely interfaces of GPCR oligomerization; 2) both all-atom and coarse-grained simulations of GPCR oligomers in their comparison to simulations of receptor monomers; and 3) methods to predict dimerization-disrupting mutations are reviewed. For simplicity, only Ballesteros and Weinstein's generic sequence numbering²⁶ is reported.

9.8.2.1 Sequence-Based Bioinformatics Approaches

The current prevailing mode of association³ of the TM helices of GPCR monomers into dimers is termed 'contact dimerization', and it corresponds to the packing of the TM helices of two different protomers via otherwise lipid-exposed interfaces.¹⁵² In the presence of seven TM helices in each protomer, there is a large number of possibilities (at least 49 for heterodimers and 28 for homodimers) for GPCR protomers to interact. Sequence-based bioinformatics approaches have extensively been applied to GPCRs by several research groups (see Filizola and Weinstein (2005),¹⁵³ Reggio (2006),¹⁵⁴ and Vohra et al. (2007)¹⁵⁵ for reviews), in an effort to reduce this number of possibilities. These methods are based on the assumption that evolutionarily conserved patterns in multiple

alignments of GPCR sequences may correspond to common structural and functional features. Thus, the application of these methods to GPCRs may be limited by the scarcity of sequences available for most receptor subfamilies, and the consequent insufficient statistical significance of the sequence data set for the multiple sequence alignment used in the calculations. A general overview of the different sequence-based methods, with a focus on the results of their application to GPCRs follows.

9.8.2.1.1 The evolutionary trace method

The evolutionary trace method was first described by Lichtarge et al. (1996)¹⁵⁶ and proposed as a powerful technique to predict functionally important residues in proteins of known structure based on the assumption that functionally important residues have a lower mutation rate compared to other residues, which is punctuated by mutations that cause divergence. Thus, in addition to residues conserved within a subfamily that are different between subfamilies, evolutionary trace residues include residues conserved across an entire sequence family (see Figure 5).

The analysis of hundreds of aligned GPCR sequences from classes A, B, and C by an enhanced evolutionary trace method using MC techniques¹⁵⁷ (entropy method¹⁵⁸) suggested TM2, TM3, TM5, and TM6 as candidates for the dimerization interface of GPCRs since lipid-exposed faces of these helices exhibited potential functional sites in common to each GPCR family or subfamily of receptors.

A more recent application of the evolutionary trace method to a multiple sequence alignment of opsin, bioamine, olfactory, and chemokine class A GPCRs¹⁵⁹ allowed identification of clusters of residues that might be responsible for 'global' and 'class-specific' functions. Thus, a generic signal transduction mechanism was inferred to occur through position 4.47, which was later mutated in chemokine receptors, and suggested to affect receptor homodimerization.¹⁶⁰

9.8.2.1.2 Correlated mutation analysis

Correlated mutation analysis (CMA) relies on the identification of simultaneous mutations in multiple sequence alignments, and was originally described as a powerful tool to predict inter-residue interactions in homologous proteins,¹⁶¹ including GPCRs.¹⁶² The observation that correlated mutations tend to accumulate at protein interfaces^{162,163} inspired the use of this method to predict GPCR dimerization/oligomerization interfaces^{160,164–166} and/or interfaces between GPCRs and G proteins.^{159,164,167–169} The underlying idea is that, to preserve an interaction between two proteins, sequence changes that occur during evolution at one protein interface must be compensated by changes in the interacting protein (see Figure 6).

While the original CMA algorithm had been proven useful for highlighting a small number of functional residues,¹⁶² it failed to provide unambiguous identification of large contact areas such as protein-protein interfaces.¹⁵⁵ For instance, Gouldson et al.¹⁶⁴ could not discriminate among TM helices as likely interfaces of dimerization of class A GPCRs since multiple sequence alignments of these receptors yield to correlated mutations in all seven TM helices.

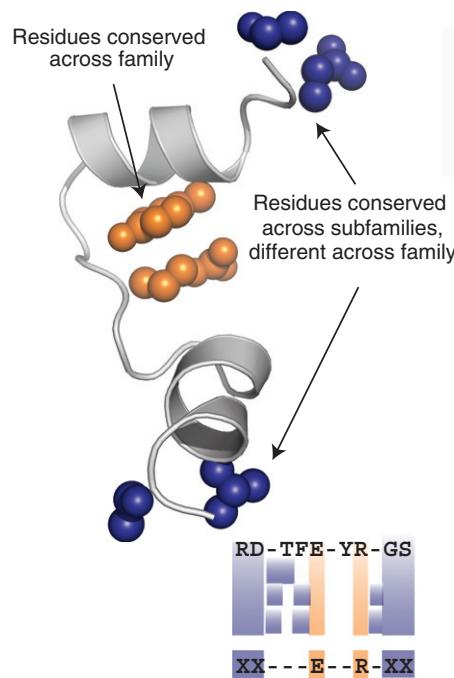


Figure 5 The evolutionary trace method. An agglomerative clustering tree is used to identify residues conserved across the whole family (in yellow) and residues conserved in each subfamilies (in blue) that change across the family.

Based on these considerations, we employed a modified version of the original CMA method using filtering algorithms based on sequence conservation and contact density.¹⁷⁰ Using this modified CMA version, and additional stringent criteria described in detail in Filizola and Weinstein (2005)¹⁵³ (e.g., focus on highly correlated pairs among lipid-exposed residues from rhodopsin crystal structure,⁴ retention of the most significant L/2 predictions,¹⁷⁰ where L is the length of the sequence of each receptor subtype; requirement for at least three lipid-exposed correlated mutations close to each other) to prune the initial lists of correlated mutations, TM1, TM4, and TM5 were identified as the most likely interfaces of opioid receptor homomers.¹⁷¹ For heteromers, an ad-hoc modified version of Olmea's CMA algorithm was developed,¹⁶⁵ termed subtractive correlated mutation (SCM) method,¹⁶⁵ to filter out the intramolecular pairs of correlated residues identified for each interacting GPCR from the complete list of intra- and intermolecular pairs of correlated residues identified in a multiple sequence alignment of concatenated GPCR sequences from the same organism.

It should be kept in mind that, while reducing the possibility of 'false positives' (i.e., incorrect predictions of residues at the interface), stringent criteria to prune the original correlated mutation lists may cause the procedure to overlook 'false negatives' (i.e., residues that should be found at an interface). In a more general application to several class A GPCRs shown to dimerize experimentally,¹⁷² a stringent CMA-based approach identified TM1 and TM4 most often as putative dimerization/oligomerization interfaces of GPCRs. Notably, several experimental approaches (e.g., receptor fragmentation, mutagenesis, bioluminescence and fluorescence resonance energy transfer, disulfide trapping, self-association, etc.) have suggested TM1 and TM4 as likely

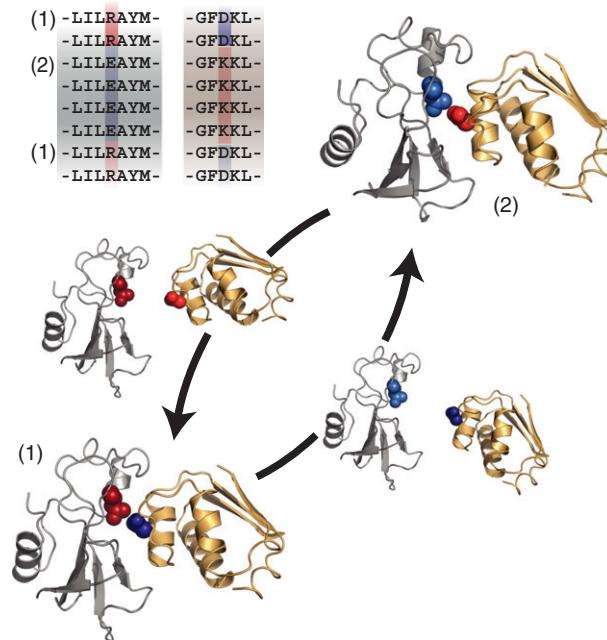


Figure 6 The correlated mutation analysis (CMA) method. Schematic representation of the CMA of a multiple sequence alignment. To preserve interactions, residue in (1) mutate in a compensatory manner to reinstate the contact in (2).

interfaces of dimerization/oligomerization in a number of GPCRs, including rhodopsin,¹⁷³ serotonin 5-HT2C,¹⁷⁴ serotonin 5-HT4,¹⁷⁵ dopamine D₂,¹⁷⁶ α 1 adrenergic,¹⁷⁷ C5a,¹⁷⁸ chemokine CCR5,¹⁶⁰ serotonin 5-HT2A-metabotropic glutamate 2,¹⁴⁴

and corticotropin releasing hormone-VT2 arginine vasotocin¹⁷⁹ receptors.

9.8.2.1.3 Methods that detect tree-determinant positions

Tree-determinant positions correspond to residues that are conserved within a subfamily of proteins, but differ between subfamilies. These positions can be identified by several algorithms^{156,180–191} in an effort to reveal important functional residues in a protein family. Among them, the 'level entropy' and the 'sequencespace automatization' methods, demonstrated to predict functionally important residues from tests on nonredundant lists of protein families, served to identify likely interfaces of dimerization in chemokine receptors¹⁶⁰ in combination with CMA.¹⁹² Specifically, the 'level entropy' method automatically identifies optimal partitions of the phylogenetic tree of a protein family based on the number of tree-determinant positions in a protein family, normalized by the number of conserved positions in each subfamily, and on inferences from the relative entropy from information theory.¹⁹³ The 'sequencespace automatization' method automatically recognizes residues with similar tendencies by using a geometric criterion based on the identification of clusters of residues in the multidimensional space.

In combination with CMA, these methods predicted TM1, TM2, and TM4 as likely interfaces of homodimerization for the chemokine receptor CCR5.¹⁶⁰ These predicted helices were used to build 100 3-D models of the CCR5 homodimer with the Global Range Molecular Matching (GRAMM) docking program.¹⁹⁴ A model with an asymmetric interface involving TM1 and TM4 helices was selected based on inferences from membrane orientation and proximity of the calculated lipid-exposed tree-determinant positions and/or correlated residues. Notably, a two-point mutation at positions 1.54 and 4.47 caused dimerization-disruption of the CCR5 homodimer, consistent with these helices forming an interface of homodimerization for this receptor.¹⁶⁰

9.8.2.1.4 Hidden-site class model of evolution

The 'hidden-site class model' of evolution¹⁹⁵ was put forward to take into account the probability that a mutation at a specific location in a protein sequence may be not functionally relevant at all locations. Thus, this method uses different substitution matrices to represent mutations at different locations in a protein sequence. Specifically, it creates 'site classes' that are each associated with a specific substitution model. Using a maximum likelihood formulation, the method calculates iteratively the assignment of locations to different site classes (unknown a priori), following optimization of the corresponding substitution models. Locations that are assigned to site classes with the slowest rate of substitution are inferred to correspond to structural or functional important positions.

This sophisticated model of evolution was applied to different subfamilies of class A GPCRs,¹⁹⁶ and identified TM4, TM5, and TM6 as likely interfaces of homodimerization based on the number of lipid-exposed locations with the slowest rate of substitution at these helices. Specifically, aminergic receptors exhibited the majority of lipid-exposed locations with the slowest rate of substitution on TM5 and TM6, whereas calculations on muscarinic, opsin, and serotonin receptors pointed

to TM4 and TM5 as likely interfaces of dimerization/oligomerization.

9.8.2.2 Protein-Protein Docking Methods

Protein-protein docking methods, which are based on a variety of selection criteria and exhaustive searches of the spatial interaction between two proteins by matching local complementary features on their surfaces, have also been applied to GPCRs in an effort to predict dimerization/oligomerization interfaces. The geometry complementarity-based GRAMM computer algorithm,¹⁹⁴ mentioned in Section 9.8.2.1.3 in reference to the CCR5 homodimer, was also applied to 5-HT4 receptor oligomers.¹⁹⁷ In this case, 40% of the most favorable 5-HT4 receptor oligomers were predicted to have TM2 and TM4 at the interface, whereas another 40% formed TM4-6/TM4-6 complexes. Structural complementarity was also used via the rigid-body docking algorithm ZDOCK¹⁹⁸ to predict TM1, TM2, and TM4 as likely interfaces of dimerization for the neuropeptidyl 1 receptor.¹⁹⁹ This strategy was tested on the smaller membrane protein system glycoporphin.²⁰⁰

9.8.2.3 Modeling Based on Experimental Data

Integrated efforts by the experimental and computational communities are necessary to increase the chances to produce accurate GPCR oligomeric models, which in turn can be used to guide novel discovery. Experimentally driven computational modeling of several GPCR oligomeric arrangements includes dopamine D2 homomers,^{176,201} and serotonin 2A/metabotropic glutamate receptor 2 complex.¹⁴⁴ Figure 7 shows the data-supported dimeric interfaces (TM1-TM1, TM4-TM4, and TM4,5-TM4,5) that have more often been reported in the literature for GPCRs.²⁴

Notably, the high-resolution crystal structure of β 2 adrenergic receptor,¹⁴ which exhibited a straighter TM1 conformation compared to rhodopsin, played a key role in building the oligomeric model of the D2 receptor supported by crosslinking data.¹⁷⁶ In fact, a rhodopsin-based model of this homomer failed to reproduce the same packing due to the different bending of the TM1 helices. Based on this model, residues at the interface within H8 were identified, which were further confirmed by experiments.¹⁷⁶ The predictive ability of the combined computational and experimental strategy in its application to dopamine D2 receptor creates exciting expectations for the structural characterization of other GPCR oligomers.

9.8.2.4 Molecular Dynamics Simulations

Although computational simulations of 3D models of GPCRs might provide useful insights into the dynamic properties of GPCR function, dynamic studies of these receptor complexes are still in their infancy. The reason for these studies lagging behind might mostly stem from: a) The large size of these systems; b) the limited timescale accessible to standard MD simulations; and c) the need for simulating several events to obtain estimates that can be compared to experiments. To overcome these problems, coarse-grained representations have

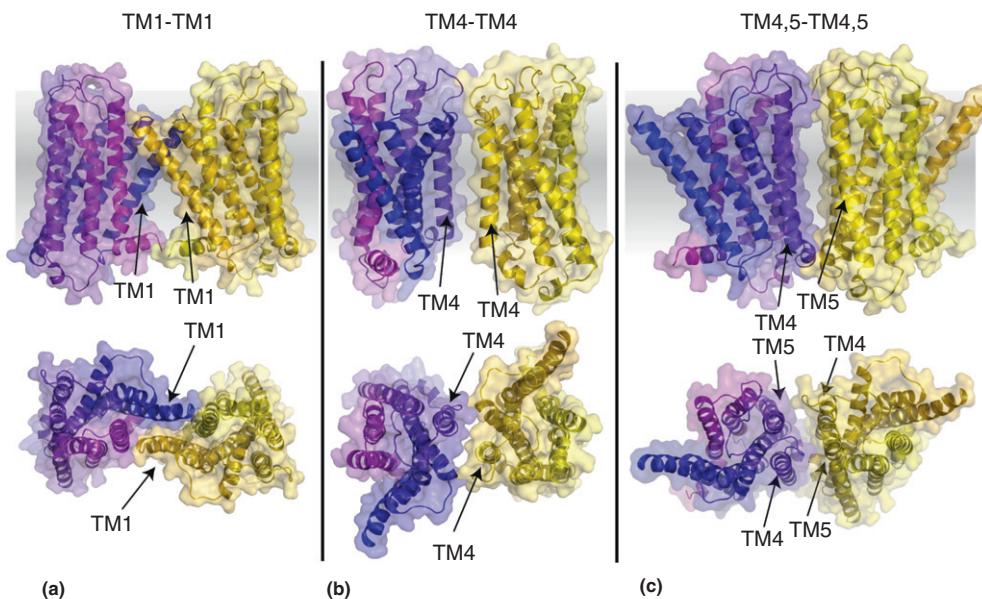


Figure 7 Structural models of putative symmetric interfaces of G protein-coupled receptor (GPCR) dimers. Experimental evidence points to different possible interfaces between protomers, shown here in both vertical and horizontal views. In particular, a dimeric interface involving TM1 helices is shown in panel (a); a dimeric interface involving TM4 helices is depicted in panel (b); panel (c) reports a third interface, where residues from both helices TM4 and TM5 are involved in interprotein contacts.

been considered, and a few applications have started to appear in the literature. The results of the different simulation methods that have been applied so far to GPCR dimers/oligomers will now be reviewed.

9.8.2.4.1 All-atom simulations in explicit environments

Given the size of dimeric or (even worse) oligomeric systems, studies using standard MD simulations of these systems in an explicit lipid-water environment have been quite scarce, due to the fact that this type of simulation is limited by the system size and submicrosecond timescales accessible using current computer hardware and algorithms.

The first nanosecond time-scale standard MD simulation of a GPCR dimer was reported by Filizola et al. in 2006.²⁰² Specifically, this study referred to the rhodopsin dimer with TM4 and TM5 helices simultaneously involved at the interface. The results of these simulations were compared to those obtained for a rhodopsin monomer simulated in the same type of bilayer membrane model, that is, an equilibrated unit cell of hydrated POPC.²⁰² An intrinsic asymmetry of the dimeric system derived from a combined essential dynamics analysis of the MD results of the rhodopsin monomer and dimer. This asymmetric behavior was particularly evident at the dimerization interface, in line with the experimentally based hypothesis²⁰¹ that a conformational rearrangement of the dimerization interface may play an important role in modulating GPCR function.

A nanosecond MD simulation of the mGluR2/5-HT2A heterodimer also forming a TM4/TM5 interface and embedded in an explicit hydrated phospholipidic bilayer has been published in the literature, and compared to the results of an analogous simulation carried out on the monomeric 5-HT2A receptor.²⁰³ Inferences from essential dynamics of these simulations pointed to an effect of the dimerization interface

on the binding site topology of individual protomers, in line with the experimentally demonstrated effect of mGluR2 on the 5-HT2A-mediated physiological response.¹⁴⁴

9.8.2.5 Coarse-Grained Simulations

Computationally efficient calculations of the dynamics of rhodopsin oligomers for larger time and length scales than accessible to atomistic models have been carried out using the MARTINI force field.²⁰⁴ This force field was developed with the aim of allowing robust transferability to a wide range of systems. A simple mapping between the all-atom system and the coarse-grained model reduces the number of degrees of freedom explicitly described in the simulation by grouping approximately four heavy atoms into one bead. The interaction between beads include both bonded (with semiempirical parameters for bonds, angles, and dihedrals) and nonbonded interactions fitted to reproduce the partitioning free-energy of amino acid side chains between water and oil phases. The model beads are comprised of four major types (polar, non-polar, apolar, and charged), and further distinguished by their hydrogen-bonding capability (acceptor, donor, both, none) or their degree of polarity (in a scale from 1 to 5). Details of the mapping for the amino acid side chains can be found in Figure 8.

Using the MARTINI force-field, systems with up to 16 rhodopsin molecules at a 1:100 protein-to-lipid ratio have been simulated in four different phospholipid environments, for 8 μs, to study the effect of the different phospholipids on the spontaneous self-assembly of rhodopsin oligomers.²⁰⁴ The results of these simulations pointed to local membrane deformation as a key component in protein-protein associations, while interactions at TM1TM2/H8, TM4/TM5, and

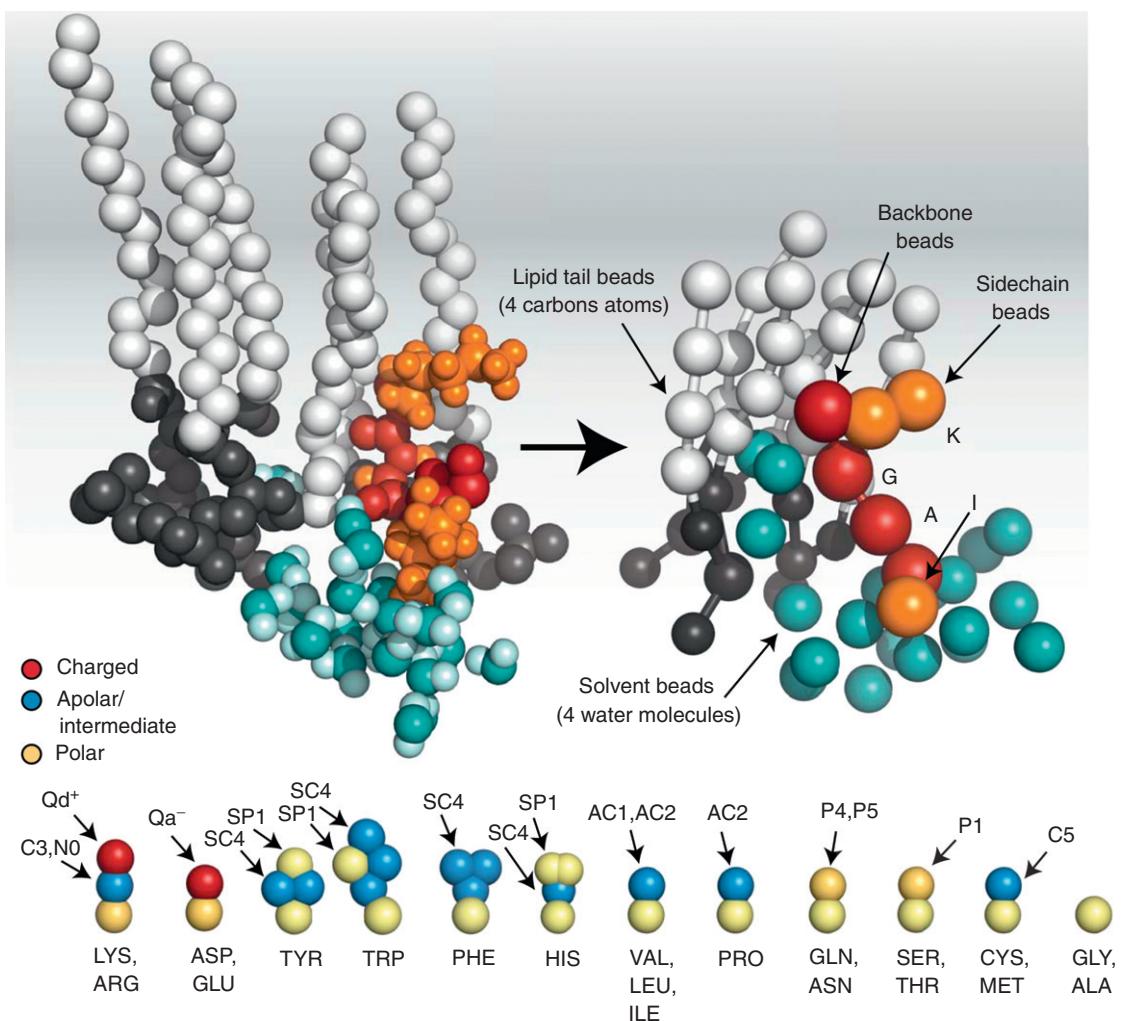


Figure 8 Coarse graining proteins in membranes. Example of the mapping between an all-atom system (left) composed of lipids (depicted in a united atoms representation with the polar head-groups in black and the hydrophobic tail in white), water (in cyan) and a peptide (with polar and nonpolar hydrogens, with the backbone atoms in red and the side chains in orange) and its coarse-grained representation (right) according to the Martini force-field.²⁰⁴ In the bottom panel the structural building blocks corresponding to the 20 amino acids in the force field. The beads are broadly divided in four types (Q for charged groups, in red; P for polar, in yellow; C for apolar and intermediate, in blue). Charged bead types are further specified as hydrogen bond donors (D) or acceptors (A); polar and intermediate groups are ranked for their polarity from 1 to 5.

TM6/TM7 appeared by the end of the microsecond simulations, indicative of the formation of stable complexes involving these helices. The next challenge of these simplified models is to allow simulations of dimeric/oligomeric arrangements of GPCRs in fully explicit representations of their environments and in the presence of G proteins (e.g., see an illustration in Figure 9).

9.8.2.6 Elastic Network Analysis

Applied ENM was used to study the overall dynamic behavior of oligomeric arrangements of rhodopsin compared to monomeric units.²⁰⁵ The results of these simulations pointed to a definite effect of oligomerization (mostly at the interface) on the dynamics of the rhodopsin monomer, and to favorable conformational rearrangements of the interfaces. Specifically, of the three activation-induced oligomeric models put forward

by experimental crosslinking data (i.e., rigid body rotation of interacting TM4s along their own helical axes, clockwise rotation of each protomer involved in intradimeric interaction along the membrane axis as viewed from the cytoplasm, and clockwise rotation of each protomer involved in intradimeric interaction followed by their sliding within the array as viewed from the cytoplasm),²⁰¹ the rigid body rotation of interacting TM4s along their own helical axes was excluded as a feasible conformational rearrangement.

9.8.2.7 Dimerization-Disrupting Mutations

Understanding the nature of the interaction between GPCR protomers at the structural and dynamic level is of extreme value as it might suggest new ways of modulating GPCR function. Thus, there is a great amount of interest in understanding how to disrupt GPCR dimerization/oligomerization.

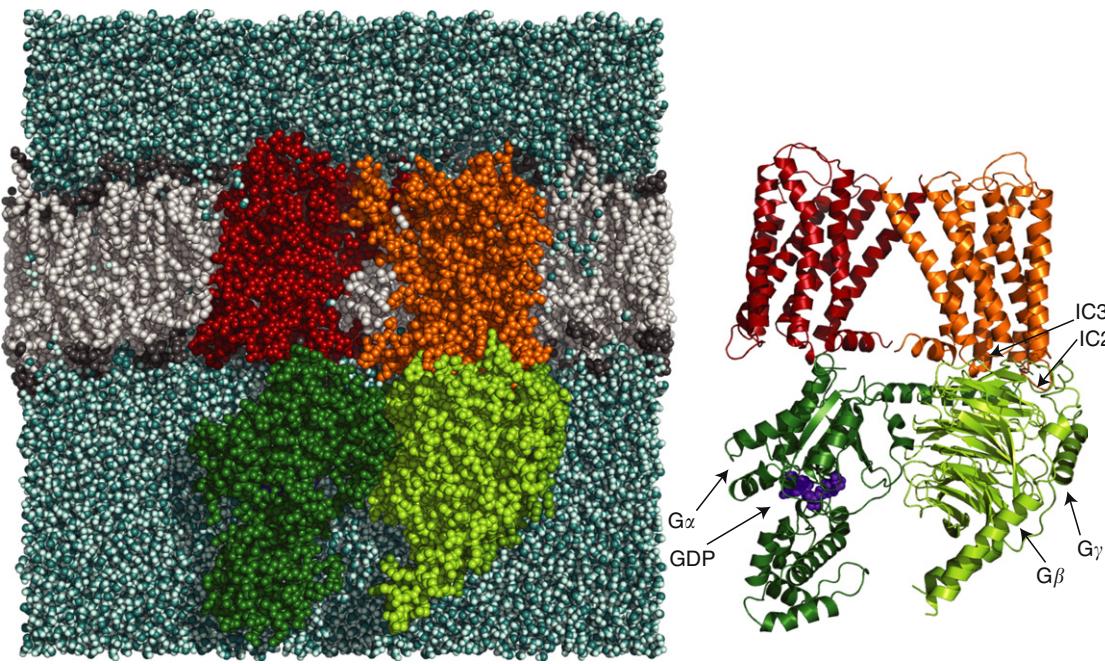


Figure 9 Structural model of a G protein-coupled receptor (GPCR) dimer in complex with its heterotrimeric G-protein. Two GPCR monomers (in red and orange) interacting through a putative TM1 interface and forming a complex with a heterotrimeric G-protein (the G α subunit is in dark green, the G β in light green, and the short G γ chain in dark green). Light and dark gray are used to represent the lipid molecules, while water molecules are depicted in cyan. The guanosine diphosphate (GDP) ligand is also shown in a purple space-filling representation inside the cartoon GPCR-G protein complex on the right.

To be able to predict dimerization-disrupting mutants in GPCRs, an *ad-hoc* computational method for the prediction of mutations that can destabilize a GPCR dimer/oligomer while maintaining the native fold of single protomers was designed.²⁰⁶ The simpler single-spanned interacting TM domain of glycophorin A was chosen for testing the efficacy of the method by comparing experimental data from mutagenesis with computational predictions. Specifically, a flexible template for a rhodopsin homodimer built around the TM4/TM5 interface was used to predict sets of three and/or five dimerization-disrupting mutations. The idea behind this choice was based on the assumption that single point mutations might not be able to disrupt the large surface of an interacting GPCR dimer. These predictions remain to be tested experimentally.

Acknowledgments

Support from NIH grants DA020032, DA026434, and DA017976 from the National Institute on Drug Abuse.

References

- [1] Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Disc.* **2006**, *5*, 993–996.
- [2] Panetta, R.; Greenwood, M. T. Physiological relevance of GPCR oligomerization and its impact on drug discovery. *Drug Discov. Today* **2008**, *13*, 1059–1066.
- [3] Milligan, G. A day in the life of a G protein-coupled receptor: The contribution to function of G protein-coupled receptor dimerization. *Br. J. Pharmacol.* **2008**, *153*(Suppl 1), S216–S229.
- [4] Palczewski, K.; Kumashiro, 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. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739–745.
- [5] Li, J.; Edwards, P. C.; Burghammer, M.; Villa, C.; Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **2004**, *343*, 1409–1438.
- [6] Nakamichi, H.; Buss, V.; Okada, T. Photoisomerization mechanism of rhodopsin and 9-cis-rhodopsin revealed by x-ray crystallography. *Biophys. J.* **2007**, *92*, L106–L108.
- [7] Nakamichi, H.; Okada, T. Crystallographic analysis of primary visual photochemistry. *Angew. Chem.* **2006**, *45*, 4270–4273.
- [8] Nakamichi, H.; Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12729–12734.
- [9] Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5982–5987.
- [10] Okada, T.; Sugihara, M.; Bondar, A. N.; Elstner, M.; Entel, P.; Buss, V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **2004**, *342*, 571–583.
- [11] Salom, D.; Lodowski, D. T.; Stenkamp, R. E.; Le Trong, I.; Golczak, M.; Jastrzebska, B.; Harris, T.; Ballesteros, J. A.; Palczewski, K. Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16123–16128.
- [12] Standfuss, J.; Xie, G.; Edwards, P. C.; Burghammer, M.; Oprian, D. D.; Schertler, G. F. Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.* **2007**, *372*, 1179–1188.
- [13] Teller, D. C.; Okada, T.; Behnke, C. A.; Palczewski, K.; Stenkamp, R. E. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry* **2001**, *40*, 7761–7772.
- [14] Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.;

- Stevens, R. C. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007**, *318*, 1258–1265.
- [15] Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V. P.; Chien, E. Y.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. *Structure* **2008**, *16*, 897–905.
- [16] Rasmussen, S. G.; Choi, H. J.; Rosenbaum, D. M.; Kobilka, T. S.; Thian, F. S.; Edwards, P. C.; Burghammer, M.; Ratnala, V. R.; Sanishvili, R.; Fischetti, R. F.; Schertler, G. F.; Weis, W. I.; Kobilka, B. K. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **2007**, *450*, 383–387.
- [17] Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Yao, X. J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **2007**, *318*, 1266–1273.
- [18] Shimamura, T.; Hiraki, K.; Takahashi, N.; Hori, T.; Ago, H.; Masuda, K.; Takio, K.; Ishiguro, M.; Miyano, M. Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region. *J. Biol. Chem.* **2008**, *283*, 17753–17756.
- [19] Murakami, M.; Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **2008**, *453*, 363–367.
- [20] Park, J. K.; Scheerer, P.; Hofmann, K. P.; Choe, H-W.; Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **2008**, *454*, 183–187.
- [21] Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krauss, N.; Choe, H. W.; Hofmann, K. P.; Ernst, O. P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008**, *455*, 497–502.
- [22] Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G.; Tate, C. G.; Schertler, G. F. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **2008**, *454*, 486–491.
- [23] Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211–1217.
- [24] Filizola, M. Increasingly accurate dynamic molecular models of G-protein coupled receptor oligomers: Panacea or Pandora's box for novel drug discovery? *Life Sci.* **2010**, *86*(15–16), 590–597.
- [25] Sánchez, R.; Sali, A. Advances in comparative protein-structure modelling. *Curr. Opin. Struct. Biol.* **1997**, *7*, 206–214.
- [26] Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Meth. Neurosci.* **1995**, *25*, 366–428.
- [27] Mobarac, J. C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: Which structural template to use? *J. Med. Chem.* **2009**, *52*, 5207–5216.
- [28] Forrest, L. R.; Tang, C. L.; Honig, B. On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys. J.* **2006**, *91*, 508–517.
- [29] Costanzi, S. On the applicability of GPCR homology models to computer-aided drug discovery: A comparison between *in silico* and crystal structures of the beta2-adrenergic receptor. *J. Med. Chem.* **2008**, *51*, 2907–2914.
- [30] Sabio, M.; Jones, K.; Topiol, S. Use of the X-ray structure of the beta2-adrenergic receptor for drug discovery. Part 2: Identification of active compounds. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5391–5395.
- [31] Topiol, S.; Sabio, M. Use of the X-ray structure of the Beta2-adrenergic receptor for drug discovery. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1598–1602.
- [32] Yuzlenko, O.; Kiec-Kononowicz, K. Molecular modeling of A1 and A2A adenosine receptors: Comparison of rhodopsin- and beta2-adrenergic-based homology models through the docking studies. *J. Comput. Chem.* **2009**, *30*, 14–32.
- [33] Larsson, P.; Wallner, B.; Lindahl, E.; Elofsson, A. Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Sci.* **2008**, *17*, 990–1002.
- [34] Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of beta2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. USA* **2009**, *106*(16), 6843–6848.
- [35] Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269–288.
- [36] Shoichet, B. K.; Leach, A. R.; Kuntz, I. D. Ligand solvation in molecular docking. *Proteins* **1999**, *34*, 4–16.
- [37] de Graaf, C.; Rognan, D. Selective structure-based virtual screening for full and partial agonists of the beta2 adrenergic receptor. *J. Med. Chem.* **2008**, *51*, 4978–4985.
- [38] Zhang, Y.; Devries, M. E.; Skolnick, J. Structure modeling of all identified G protein-coupled receptors in the human genome. *PLoS Comput. Biol.* **2006**, *2*, e13.
- [39] Yarov-Yarovsky, V.; Schonbrun, J.; Baker, D. Multipass membrane protein structure prediction using Rosetta. *Proteins* **2006**, *62*, 1010–1025.
- [40] Barth, P.; Schonbrun, J.; Baker, D. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15682–15687.
- [41] Kühlman, B.; Baker, D. Native protein sequences are close to optimal for their structures. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10383–10388.
- [42] Deupi, X.; Olivella, M.; Govaerts, C.; Ballesteros, J. A.; Campillo, M.; Pardo, L. Ser and Thr residues modulate the conformation of pro-kinked transmembrane alpha-helices. *Biophys. J.* **2004**, *86*, 105–115.
- [43] Lazaridis, T. Effective energy function for proteins in lipid membranes. *Proteins* **2003**, *52*, 176–192.
- [44] Filizola, M.; Perez, J. J.; Carteni-Farinha, M. BUNDLE: A program for building the transmembrane domains of G-protein-coupled receptors. *J. Comput. Aided Mol. Des.* **1998**, *12*, 111–118.
- [45] Schertler, G. F.; Hargrave, P. A. Projection structure of frog rhodopsin in two crystal forms. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 11578–11582.
- [46] Trabertino, R. J.; Hall, S. E.; Vaidehi, N.; Floriano, W. B.; Kam, V. W.; Goddard, III W. A. First principles predictions of the structure and function of G-protein-coupled receptors: Validation for bovine rhodopsin. *Biophys. J.* **2004**, *86*, 1904–1921.
- [47] Freddolino, P. L.; Kalani, M. Y.; Vaidehi, N.; Floriano, W. B.; Hall, S. E.; Trabertino, R. J.; Kam, V. W.; Goddard, III W. A. Predicted 3D structure for the human beta 2 adrenergic receptor and its binding site for agonists and antagonists. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2736–2741.
- [48] Spijker, P.; Vaidehi, N.; Freddolino, P. L.; Hilbers, P. A.; Goddard, III W. A. Dynamic behavior of fully solvated beta2-adrenergic receptor, embedded in the membrane with bound agonist or antagonist. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4882–4887.
- [49] Floriano, W. B.; Vaidehi, N.; Goddard, III W. A. Making sense of olfaction through predictions of the 3D structure and function of olfactory receptor. *Chem. Sens.* **2004**, *29*, 269–290.
- [50] Hall, S. E.; Floriano, W. B.; Vaidehi, N.; Goddard, III W. A. 3D Structures for mouse I7 and rat I7 olfactory receptors from theory and odor recognition profiles from theory and experiment. *Chem. Sens.* **2004**, *29*, 595–616.
- [51] Hummel, P.; Vaidehi, N.; Floriano, W. B.; Hall, S. E.; Goddard, W. A. Test of the Binding Threshold Hypothesis for olfactory receptors: Explanation of the differential binding of ketones to the mouse and human orthologs of olfactory receptor. *Protein Sci.* **2005**, *14*, 703–710.
- [52] Kalani, M. Y.; Vaidehi, N.; Hall, S. E.; Trabertino, R.; Freddolino, P.; Kalani, M. A.; Floriano, W. B.; Kam, V.; Goddard, III W. A. Predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3815–3820.
- [53] Peng, J.; Vaidehi, N.; Hall, S.; Goddard, III W. A. The predicted 3D structures of the human M1 muscarinic acetylcholine receptor with agonist or antagonist bound. *Chem. Med. Chem.* **2006**, *8*, 878–890.
- [54] Vaidehi, N.; Schlyer, S.; Trabertino, R.; Kochanny, M.; Abrol, R.; Koovakat, S.; Dunning, L.; Liang, M.; Sharma, S.; Fox, J. M.; Floriano, W. B.; Mendonça, F. L. D.; Pease, J. E.; Goddard, III W. A.; Horuk, R. Predictions of CCR1 chemokine receptor structure and BX 471 antagonist binding followed by experimental validation. *J. Biol. Chem.* **2006**, *281*, 27613–27620.
- [55] Martí-Renom, M.; Stuart, A.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.
- [56] Xiang, Z.; Soto, C. S.; Honig, B. Evaluating conformational free energies: The colony energy and its application to the problem of loop prediction. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7432–7437.
- [57] Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins* **2004**, *55*, 351–367.
- [58] Wang, C.; Bradley, P.; Baker, D. Protein-protein docking with backbone flexibility. *J. Mol. Biol.* **2007**, *373*, 503–519.

- [59] Provasi, D.; Bortolato, A.; Filizola, M. Exploring molecular mechanisms of ligand recognition by opioid receptors with metadynamics. *Biochemistry* **2009**, *48*, 10020–10029.
- [60] Mehler, E. L.; Periole, X.; Hassan, S. A.; Weinstein, H. Key issues in the computational simulation of GPCR function: Representation of loop domains. *J. Comput. Aided Mol. Des.* **2002**, *16*, 841–853.
- [61] Noguti, T.; Go, N. Efficient Monte Carlo method for simulation of fluctuating conformations of native proteins. *Biopolymers* **1985**, *24*, 527–546.
- [62] Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
- [63] Hassan, S. A.; Mehler, E. L. A critical analysis of continuum electrostatics: The screened Coulomb potential-implicit solvent model and the study of the alanine dipeptide and discrimination of misfolded structures of proteins. *Proteins* **2002**, *47*, 45–61.
- [64] Kartagere, S.; Roy, A.; Mehler, E. L. *Ab initio* computational modeling of long loops in G-protein coupled receptors. *J. Comput. Aided Mol. Des.* **2006**, *20*, 427–436.
- [65] Shi, L.; Javitch, J. The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 440–445.
- [66] Nikiforovich, G. V.; Taylor, C. M.; Marshall, G. R.; Baranski, T. J. Modeling the possible conformations of the extracellular loops in G-protein-coupled receptors. *Proteins* **2010**, *78*, 271–285.
- [67] Moult, J.; Fidelis, K.; Kryshtafovych, A.; Rost, B.; Anna, T. Critical assessment of methods of protein structure prediction – Round VIII. *Proteins* **2009**, *77*, 1–4.
- [68] Janin, J.; Henrick, K.; Moult, J.; Eyck, L.; Sternberg, M.; Vajda, S.; Vakser, I.; Wodak, S.; CAPRI, A. Critical Assessment of PRdicted Interactions. *Proteins* **2003**, *52*, 2–9.
- [69] Michino, M.; Abola, E.; GPCR Dock 2008 participants.; Brooks, C. L. I.; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock. *Nat. Rev. Drug Disc.* **2008**, *8*, 455–463.
- [70] Li, Y. Y.; Zhu, F. Q.; Vaidehi, N.; Goddard, W. A. Prediction of the 3D structure and dynamics of human DP G-protein coupled receptor bound to an agonist and an antagonist. *J. Am. Chem. Soc.* **2007**, *129*, 10720–10731.
- [71] Costanzi, S.; Ivanov, A. A.; Tikhonova, I. G.; and Jacobson, K. A.; Structure and function of g protein-coupled receptors studied using sequence analysis, molecular modeling and receptor engineering. In *Frontiers in Drug Design and Discovery*; Caldwell, G. W.; Rahman, A. U.; Player, M. R.; and Chouday, M. I., Eds.; *3*, 63–79, 2007, Bentham Science Publishers Ltd., Oak Park, IL.
- [72] Costanzi, S.; Lambertucci, C.; Vittori, S.; Volpini, R.; Cristalli, G. 2- and 8-alkynadenosines: Conformational studies and docking to human adenosine A3 receptor can explain their different biological behavior. *J. Mol. Graphics Model* **2003**, *21*, 253–262.
- [73] Dror, R. O.; Arlow, D. H.; Borhani, D. W.; Jensen, M. O.; Piana, S.; Shaw, D. E. Identification of two distinct inactive conformations of the beta2-adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4689–4694.
- [74] Grossfield, A.; Feller, S. E.; Pitman, M. C. Convergence of molecular dynamics simulations of membrane proteins. *Proteins* **2007**, *67*, 31–40.
- [75] Khelashvili, G.; Grossfield, A.; Feller, S. E.; Pitman, M. C.; Weinstein, H. Structural and dynamic effects of cholesterol at preferred sites of interaction with rhodopsin identified from microsecond length molecular dynamics simulations. *Proteins* **2009**, *76*, 403–417.
- [76] Cordomí, A.; Perez, J. J. Molecular dynamics simulations of rhodopsin in different one-component lipid bilayers. *J. Phys. Chem. B* **2007**, *111*, 7052–7063.
- [77] Grossfield, A.; Feller, S. E.; Pitman, M. C. A role for direct interactions in the modulation of rhodopsin by omega-3 polyunsaturated lipids. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4888–4893.
- [78] Pitman, M. C.; Grossfield, A.; Suits, F.; Feller, S. E. Role of cholesterol and polyunsaturated chains in lipid-protein interactions: Molecular dynamics simulation of rhodopsin in a realistic membrane environment. *J. Am. Chem. Soc.* **2005**, *127*, 4576–4577.
- [79] Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.* **2009**, *19*, 120–127.
- [80] MacKerell, J. A.; Banavali, N.; Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **2001**, *56*, 257–265.
- [81] Berger, O.; Edholm, O.; Jahnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **1997**, *72*, 2002–2013.
- [82] Tieleman, D. P.; MacCallum, J. L.; Ash, W. L.; Kandt, C.; Xu, Z.; Monticelli, L. Membrane protein simulations with a united-atom lipid and all-atom protein model: Lipid-protein interactions, side chain transfer free energies and model proteins. *J. Phys. Condens. Matter* **2006**, *18*, S1221–S1234.
- [83] Hénin, J.; Shinoda, W.; Klein, M. L. United-atom acyl chains for CHARMM phospholipids. *J. Phys. Chem. B* **2008**, *112*, 7008–7015.
- [84] Kandt, C.; Ash, W. L.; Tieleman, D. P. Setting up and running molecular dynamics simulations of membrane proteins. *Methods* **2007**, *41*, 475–488.
- [85] Provasi, D.; Filizola, M. Putative active states of a prototypic G-protein coupled receptor from biased molecular dynamics. *Biophys. J.* **2010**, *98*(10), 2347–2355.
- [86] Crozier, P. S.; Stevens, M. J.; Forrest, L. R.; Woolf, T. B. Molecular dynamics simulation of dark-adapted rhodopsin in an explicit membrane bilayer: Coupling between local retinal and larger scale conformational change. *J. Mol. Biol.* **2003**, *333*, 493–514.
- [87] Kong, Y.; Karplus, M. The signaling pathway of rhodopsin. *Structure* **2007**, *15*, 611–623.
- [88] Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K. Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin. *Biophys. J.* **2002**, *83*, 3097–3112.
- [89] Faraldo-Gómez, J. D.; Forrest, L. R.; Baaden, M.; Bond, P. J.; Domene, C.; Patargias, G.; Cuthbertson, J.; Sansom, M. S. Conformational sampling and dynamics of membrane proteins from 10-nanosecond computer simulations. *Proteins* **2004**, *57*, 783–791.
- [90] Huber, T.; Botelho, A. V.; Beyer, K.; Brown, M. F. Membrane model for the G-protein-coupled receptor rhodopsin: Hydrophobic interface and dynamical structure. *Biophys. J.* **2004**, *86*, 2078–2100.
- [91] Schlegel, B.; Sippl, W.; Holtje, H. D. Molecular dynamics simulations of bovine rhodopsin: Influence of protonation states and different membrane-mimicking environments. *J. Mol. Mod.* **2005**, *12*, 49–64.
- [92] Lemaitre, V.; Yeagle, P.; Watts, A. Molecular dynamic simulations of retinal in rhodopsin: From the dark-adapted state towards lumirhodopsin. *Biochemistry* **2005**, *44*, 12667–12680.
- [93] Crozier, P. S.; Stevens, M. J.; Woolf, T. B. How a small change in retinal leads to G-protein activation: Initial events suggested by molecular dynamics calculations. *Proteins* **2007**, *66*, 559–574.
- [94] Martínez-Mayorga, K.; Pitman, M. C.; Grossfield, A.; Feller, S. E.; Brown, M. F. Retinal counterion switch mechanism in vision evaluated by molecular simulations. *J. Am. Chem. Soc.* **2006**, *128*, 16502–16503.
- [95] Huber, T.; Menon, S.; Sakmar, T. P. Structural basis for ligand binding and specificity in adrenergic receptors: Implications for GPCR-targeted drug discovery. *Biochemistry* **2008**, *47*, 11013–11023.
- [96] Vanni, S.; Neri, M.; Tavernelli, I.; Rothlisberger, U. Observation of “ionic lock” formation in molecular dynamics simulations of wild-type beta 1 and beta 2 adrenergic receptors. *Biochemistry* **2009**, *48*, 4789–4797.
- [97] Lyman, E.; Higgs, C.; Kim, B.; Lupyan, D.; Shelley, J. C.; Farid, R.; Voth, G. A. A role for a specific cholesterol interaction in stabilizing the Apo configuration of the human A(2A) adenosine receptor. *Structure* **2009**, *17*, 1660–1668.
- [98] Trent, J. O.; Wang, Z. X.; Murray, J. L.; Shao, W.; Tamamura, H.; Fujii, N.; Peiper, S. C. Lipid bilayer simulations of CXCR4 with inverse agonists and weak partial agonists. *J. Biol. Chem.* **2003**, *278*, 47136–47144.
- [99] Aburi, M.; Smith, P. E. Modeling and simulation of the human delta opioid receptor. *Protein Sci.* **2004**, *13*, 1997–2008.
- [100] Zhang, Y.; Sham, Y. Y.; Rajamani, R.; Gao, J.; Portoghesi, P. S. Homology modeling and molecular dynamics simulations of the mu-opioid receptor in a membrane-aqueous system. *Chem. Bio. Chem.* **2005**, *6*, 1–7.
- [101] Rivail, L.; Chipot, C.; Maigret, B.; Bestel, I.; Sicsic, S.; Tarek, M. Large-scale molecular dynamics of a G protein-coupled receptor, the human 5-HT4 serotonin receptor, in a lipid bilayer. *J. Mol. Struct. THEOCHEM* **2007**, *817*, 19–26.
- [102] Lynch, D.; Hurst, D.; Reggio, P.; Grossfield, A.; Pitman, M. C. Atomic level description of GPCR activation revealed by microsecond time scale molecular dynamics. *Biophys. J.* **2009**, *96*, 365a.
- [103] Park, J. H.; Scheerer, P.; Hofmann, K. P.; Choe, H. W.; Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsins. *Nature* **2008**, *454*, 183–187.
- [104] Mustafi, D.; Palczewski, K. Topology of class A G protein-coupled receptors: Insights gained from crystal structures of rhodopsins, adrenergic and adenosine receptors. *Mol. Pharmacol.* **2009**, *75*, 1–12.

- [105] Schwartz, T. W.; Hubbell, W. L. Structural biology: A moving story of receptors. *Nature* **2008**, *455*, 473–474.
- [106] Hubbell, W. L.; Altenbach, C.; Hubbell, C. M.; Khorana, H. G. Rhodopsin structure, dynamics, activation: A perspective from crystallography, site-directed spin labeling, sulfhydryl reactivity, and disulfide cross-linking. *Adv. Protein Chem.* **2003**, *63*, 243–290.
- [107] Kusnetzow, A. K.; Altenbach, C.; Hubbell, W. L. Conformational states and dynamics of rhodopsin in micelles and bilayers. *Biochemistry* **2006**, *45*, 5538–5550.
- [108] Farrens, D. L.; Altenbach, C.; Yang, K.; Hubbell, W. L.; Khorana, H. G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **1996**, *274*, 768–770.
- [109] Knierim, B.; Hofmann, K. P.; Ernst, O. P.; Hubbell, W. L. Sequence of late molecular events in the activation of rhodopsin. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20290–20295.
- [110] Kobilka, B. K.; Deupi, X. Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **2007**, *28*, 397–406.
- [111] Ridge, K. D.; Palczewski, K. Visual rhodopsin sees the light: Structure and mechanism of G protein signaling. *J. Biol. Chem.* **2007**, *282*, 9297–9301.
- [112] Altenbach, C.; Kusnetzow, A. K.; Ernst, O. P.; Hofmann, K. P.; Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7439–7444.
- [113] Best, R. B.; Chen, Y. G.; Hummer, G. Slow protein conformational dynamics from multiple experimental structures: The helix/sheet transition of arc repressor. *Structure* **2005**, *13*, 1755–1763.
- [114] Maragakis, P.; Karplus, M. Large amplitude conformational change in proteins explored with a plastic network model: Adenylate kinase. *J. Mol. Biol.* **2005**, *352*, 807–822.
- [115] Miyashita, O.; Wolynes, P. G.; Onuchic, J. N. Simple energy landscape model for the kinetics of functional transitions in proteins. *J. Phys. Chem. B* **2005**, *109*(5), 1959–1969.
- [116] Bahar, I.; Rader, A. J. Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* **2005**, *15*, 586–592.
- [117] Ma, J. Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* **2005**, *13*, 373–380.
- [118] Tama, F.; Brooks, III C. L. Symmetry, form, and shape: Guiding principles for robustness in macromolecular machines. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 115–133.
- [119] Cui, Q.; Li, G.; Ma, J.; Karplus, M. A normal mode analysis of structural plasticity in the biomolecular motor F(1)-ATPase. *J. Mol. Biol.* **2004**, *340*, 345–372.
- [120] Hayward, S.; Kitao, A.; Berendsen, H. J. Model-free methods of analyzing domain motions in proteins from simulation: A comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins* **1997**, *27*, 425–437.
- [121] Krebs, W. G.; Alexandrov, V.; Wilson, C. A.; Echols, N.; Yu, H.; Gerstein, M. Normal mode analysis of macromolecular motions in a database framework: Developing mode concentration as a useful classifying statistic. *Proteins* **2002**, *48*, 682–695.
- [122] McCommon, J. A.; Gelin, B. R.; Karplus, M.; Wolynes, P. G. The hinge-bending mode in lysozyme. *Nature* **1976**, *262*, 325–326.
- [123] Tama, F.; Valle, M.; Frank, J.; Brooks, III C. L. Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9319–9323.
- [124] Lu, M.; Ma, J. The role of shape in determining molecular motions. *Biophys. J.* **2005**, *89*, 2395–2401.
- [125] Bahar, I.; Wallqvist, A.; Covell, D.; Jernigan, R. Correlation between native state hydrogen exchange and cooperative residue fluctuations from a simple model. *Biochemistry* **1998**, *37*, 1067–1073.
- [126] Tirion, M. M. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Physic. Rev. Lett.* **1996**, *77*, 1905–1908.
- [127] van Vlijmen, H.; Karplus, M. Normal mode calculations of icosahedral viruses with full dihedral flexibility by use of molecular symmetry. *J. Mol. Biol.* **2005**, *350*, 528–542.
- [128] Isin, B.; Rader, A. J.; Dhiman, H. K.; Klein-Seetharaman, J.; Bahar, I. Predisposition of the dark state of rhodopsin to functional changes in structure. *Proteins* **2006**, *65*, 970–983.
- [129] Niv, M. Y.; Skrabanek, L.; Filizola, M.; Weinstein, H. Modeling activated states of GPCRs: The rhodopsin template. *J. Comput. Aided Mol. Des.* **2006**, *20*, 437–448.
- [130] Tobi, D.; Bahar, I. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18908–18913.
- [131] Isin, B.; Schulter, K.; Tajkhorshid, E.; Bahar, I. Mechanism of signal propagation upon retinal isomerization: Insights from molecular dynamics simulations of rhodopsin restrained by normal modes. *Biophys. J.* **2008**, *95*, 789–803.
- [132] Tikhonova, I. G.; Best, R. B.; Engel, S.; Gershengorn, M. C.; Hummer, G.; Costanzi, S. Atomistic insights into rhodopsin activation from a dynamic model. *J. Am. Chem. Soc.* **2008**, *130*, 10141–10149.
- [133] Bhattacharya, S.; Hall, S. E.; Vaidehi, N. Agonist-induced conformational changes in bovine rhodopsin: Insight into activation of G-protein-coupled receptors. *J. Mol. Biol.* **2008**, *382*, 539–555.
- [134] Bhattacharya, S.; Hall, S. E.; Lia, H.; Vaidehi, N. Ligand-stabilized conformational states of human $\beta 2$ adrenergic receptor: Insight into G-protein-coupled receptor activation. *Biophys. J.* **2008**, *94*, 2027–2042.
- [135] Nikiforovich, G. V.; Marshall, G. R. Three-dimensional model for meta-II rhodopsin, an activated G-protein-coupled receptor. *Biochemistry* **2003**, *42*, 9110–9120.
- [136] Nikiforovich, G. V.; Galaktionov, S.; Balodis, J.; Marshall, G. R. Novel approach to computer modeling of seven-helical transmembrane proteins: Current progress in the test case of bacteriorhodopsin. *Acta Biochim. Pol.* **2001**, *48*, 53–64.
- [137] Roux, B. Statistical mechanical equilibrium theory of selective ion channels. *Biophys. J.* **1999**, *77*, 139–153.
- [138] Marchi, M.; Ballone, P. Adiabatic bias molecular dynamics: A method to navigate the conformational space of complex molecular systems. *J. Chem. Physics* **1999**, *110*, 3697–3702.
- [139] Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566.
- [140] Bussi, G.; Laio, A.; Parrinello, M. Equilibrium free energies from non-equilibrium metadynamics. *Phys. Rev. Lett.* **2006**, *96*, 090601.
- [141] Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **2008**, *100*, 020603.
- [142] Chabre, M.; le Maire, M. Monomeric G-protein-coupled receptor as a functional unit. *Biochemistry* **2005**, *44*, 9395–9403.
- [143] Waldhoer, M.; Fong, J.; Jones, R. M.; Lunzer, M. M.; Sharma, S. K.; Kostenis, E.; Portoghesi, P. S.; Whistler, J. L. A heterodimer-selective agonist shows *in vivo* relevance of G protein-coupled receptor dimers. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9050–9055.
- [144] Gonzalez-Maeso, J.; Ang, R. L.; Yuen, T.; Chan, P.; Weisstaub, N. V.; Lopez-Gimenez, J. F.; Zhou, M.; Okawa, Y.; Caillado, L. F.; Milligan, G.; Gingrich, J. A.; Filizola, M.; Meana, J. J.; Sealfon, S. C. Identification of a serotonin/glutamate receptor complex implicated in psychosis. *Nature* **2008**, *452*, 93–97.
- [145] Finley, M. J.; Chen, X.; Bardi, G.; Davey, P.; Geller, E. B.; Zhang, L.; Adler, M. W.; Rogers, T. J. Bi-directional heterologous desensitization between the major HIV-1 co-receptor CXCR4 and the kappa-opioid receptor. *J. Neuroimmunol.* **2008**, *197*, 114–123.
- [146] Leanos-Miranda, A.; Ulloa-Aguirre, A.; Janovick, J. A.; Conn, P. M. In vitro coexpression and pharmacological rescue of mutant gonadotropin-releasing hormone receptors causing hypogonadotropic hypogonadism in humans expressing compound heterozygous alleles. *J. Clin. Endocrinol. Metab.* **2005**, *90*, 3001–3008.
- [147] Carriba, P.; Ortiz, O.; Patkar, K.; Justinova, Z.; Stroik, J.; Themann, A.; Muller, C.; Woods, A. S.; Hope, B. T.; Ciruela, F.; Casado, V.; Canela, E. I.; Lluis, C.; Goldberg, S. R.; Moratalla, R.; Franco, R.; Ferre, S. Striatal adenosine A2A and cannabinoid CB1 receptors form functional heteromeric complexes that mediate the motor effects of cannabinoids. *Neuropsychopharmacology* **2007**, *32*, 2249–2259.
- [148] McGraw, D. W.; Mihlbachler, K. A.; Schwab, M. R.; Rahman, F. F.; Small, K. M.; Almoosa, K. F.; Liggett, S. B. Airway smooth muscle prostaglandin-EP1 receptors directly modulate beta2-adrenergic receptors within a unique heterodimeric complex. *J. Clin. Invest.* **2006**, *116*, 1400–1409.
- [149] AbdAlla, S.; Lother, H.; el Massiry, A.; Quitterer, U. Increased AT(1) receptor heterodimers in preeclampsia mediate enhanced angiotensin II responsiveness. *Nat. Med.* **2001**, *7*, 1003–1009.
- [150] Brauner-Osborne, H.; Krosgaard-Larsen, P. Functional pharmacology of cloned heterodimeric GABAB receptors expressed in mammalian cells. *Br. J. Pharmacol.* **1999**, *128*, 1370–1374.
- [151] Skrabanek, L.; Murcia, M.; Bouvier, M.; Devi, L.; George, S. R.; Lohse, M. J.; Milligan, G.; Neubig, R.; Palczewski, K.; Parmentier, M.; Pin, J. P.; Vriend, G.

- G.; Javitch, J. A.; Campagne, F.; Filizola, M. Requirements and ontology for a G protein-coupled receptor oligomerization knowledge base. *BMC Bioinform.* **2007**, *8*, 177.
- [152] Schulz, A.; Gross, R.; Schultz, G.; Gudermann, T.; Schoneberg, T. Structural implication for receptor oligomerization from functional reconstitution studies of mutant V2 vasopressin receptors. *J. Biol. Chem.* **2000**, *275*, 2381–2389.
- [153] Filizola, M.; Weinstein, H. The study of G-protein coupled receptor oligomerization with computational modeling and bioinformatics. *FEBS J.* **2005**, *272*, 2926–2938.
- [154] Reggio, P. H. Computational methods in drug design: Modeling G protein-coupled receptor monomers, dimers, and oligomers. *AAPS J.* **2006**, *8*, E322–E336.
- [155] Vohra, S.; Chintapalli, S. V.; Illingworth, C. J.; Reeves, P. J.; Mullineaux, P. M.; Clark, H. S.; Dean, M. K.; Upton, G. J.; Reynolds, C. A. Computational studies of family A and family B GPCRs. *Biochem. Soc. Trans.* **2007**, *35*, 749–754.
- [156] Lichtarge, O.; Bourne, H. R.; Cohen, F. E. An evolutionary trace method defines binding surfaces common to protein families. *J. Mol. Biol.* **1996**, *257*, 342–358.
- [157] Upton, G.; Fingleton, B. *Spatial Data Analysis by Example*; Wiley: Chichester, UK, 1985; pp 9–104.
- [158] 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. Dimerization of G-protein-coupled receptors. *J. Med. Chem.* **2001**, *44*, 4595–4614.
- [159] Madabushi, S.; Gross, A. K.; Philippi, A.; Meng, E. C.; Wensel, T. G.; Lichtarge, O. Evolutionary trace of G protein-coupled receptors reveals clusters of residues that determine global and class-specific functions. *J. Biol. Chem.* **2004**, *279*, 8126–8132.
- [160] 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. Identification of amino acid residues crucial for chemokine receptor dimerization. *Nat. Immunol.* **2004**, *5*, 216–223.
- [161] Gobel, U.; Sander, C.; Schneider, R.; Valencia, A. Correlated mutations and residue contacts in proteins. *Proteins* **1994**, *18*, 309–317.
- [162] Oliveira, L.; Paiva, A. C. M.; Vriend, G. A common motif in G-protein coupled seven transmembrane helix receptors. *J. Comput. Aided Mol. Des.* **1993**, *7*, 649–658.
- [163] Pazos, F.; Helmer-Citterich, M.; Ausiello, G.; Valencia, A. Correlated mutations contain information about protein-protein interaction. *J. Mol. Biol.* **1997**, *271*, 511–523.
- [164] Gouldson, P. R.; Dean, M. K.; Snell, C. R.; Bywater, R. P.; Gkoutos, G.; Reynolds, C. A. Lipid-facing correlated mutations and dimerization in G-protein coupled receptors. *Protein Eng.* **2001**, *14*, 759–767.
- [165] Filizola, M.; Olmea, O.; Weinstein, H. Prediction of heterodimerization interfaces of G-protein coupled receptors with a new subtractive correlated mutation method. *Protein Eng.* **2002**, *15*, 881–885.
- [166] Filizola, M.; Olmea, O.; Weinstein, H. Using correlated mutation analysis to predict the heterodimerization interface of GPCRs. *Biophys. J.* **2002**, *82*, 2307.
- [167] Oliveira, L.; Paiva, A. C.; Vriend, G. Correlated mutation analyses on very large sequence families. *ChemBioChem.* **2002**, *3*, 1010–1017.
- [168] Horn, F.; van der Wenden, E. M.; Oliveira, L.; IJzerman, A. P.; Vriend, G. Receptors coupling to G proteins: Is there a signal behind the sequence? *Proteins* **2000**, *41*, 448–459.
- [169] Moller, S.; Vilo, J.; Croning, M. D. Prediction of the coupling specificity of G protein coupled receptors to their G proteins. *Bioinformatics* **2001**, *17*, S174–S181.
- [170] Olmea, O.; Valencia, A. Improving contact predictions by the combination of correlated mutations and other sources of sequence information. *Fold. Des.* **1997**, *2*, S25–S32.
- [171] Filizola, M.; Weinstein, H. Structural models for dimerization of G-protein coupled receptors: The opioid receptor homodimers. *Biopolymers* **2002**, *66*, 317–325.
- [172] Filizola, M.; Guo, W.; Javitch, J. A.; Weinstein, H. Oligomerization Domains of G-Protein Coupled Receptors: Insights into the Structural Basis of GPCR Association. In *Contemporary Clinical Neuroscience: The G Protein-Coupled Receptor Handbook*; Devi, L. A., Ed.; Humana Press Inc.: Totowa, NJ, 2005; pp 243–265.
- [173] Liang, Y.; Fotiadis, D.; Filipek, S.; Saperstein, D. A.; Palczewski, K.; Engel, A. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *J. Biol. Chem.* **2003**, *278*, 21655–21662.
- [174] Mancia, F.; Assur, Z.; Herman, A. G.; Siegel, R.; Hendrickson, W. A. Ligand sensitivity in dimeric associations of the serotonin 5HT2c receptor. *EMBO Rep.* **2008**, *9*, 363–369.
- [175] Berthouze, M.; Rivail, L.; Lucas, A.; Ayoub, M. A.; Russo, O.; Sicsic, S.; Fischmeister, R.; Berque-Bestel, I.; Jockers, R.; Lezoualc'h, F. Two transmembrane Cys residues are involved in 5-HT4 receptor dimerization. *Biochem. Biophys. Res. Commun.* **2007**, *356*, 642–647.
- [176] Guo, W.; Urizar, E.; Kralkova, M.; Mobarac, J. C.; Shi, L.; Filizola, M.; Javitch, J. A. Dopamine D2 receptors form higher order oligomers at physiological expression levels. *EMBO J.* **2008**, *27*, 2293–2304.
- [177] Lopez-Gimenez, J. F.; Canals, M.; Pediani, J. D.; Milligan, G. The alpha1b-adrenoceptor exists as a higher-order oligomer: Effective oligomerization is required for receptor maturation, surface delivery, and function. *Mol. Pharmacol.* **2007**, *71*, 1015–1029.
- [178] Klco, J. M.; Lassere, T. B.; Baranski, T. J. C5a receptor oligomerization. I. Disulfide trapping reveals oligomers and potential contact surfaces in a G protein-coupled receptor. *J. Biol. Chem.* **2003**, *278*, 35345–35353.
- [179] Mikhailova, M. V.; Blanett, J.; Jacobi, S.; Mayeux, P. R.; Cornett, L. E. Transmembrane domain IV of the Gallus gallus VT2 vasotocin receptor is essential for forming a heterodimer with the corticotrophin releasing hormone receptor. *J. Biomed. Opt.* **2008**, *13*, 031208.
- [180] Livingstone, C. D.; Barton, G. J. Protein sequence alignments: A strategy for the hierarchical analysis of residue conservation. *Bioinformatics* **1993**, *9*, 745–756.
- [181] Casari, G.; Sander, C.; Valencia, A. A method to predict functional residues in proteins. *Nat. Struct. Biol.* **1995**, *2*, 171–178.
- [182] Dorit, R. L.; Ayala, F. J. ADH evolution and the phylogenetic footprint. *J. Mol. Evol.* **1995**, *40*, 658–662.
- [183] Andrade, M. A.; Casari, G.; Sander, C.; Valencia, A. Classification of protein families and detection of the determinant residues with an improved self-organizing map. *Biol. Cybern.* **1997**, *76*, 441–450.
- [184] Zhang, B.; Rychlewski, L.; Pawlowski, K.; Fetrow, J. S.; Skolnick, J.; Godzik, A. From fold predictions to function predictions: Automation of functional site conservation analysis for functional genome predictions. *Protein Sci.* **1999**, *8*, 1104–1115.
- [185] Goh, C. S.; Bogan, A. A.; Joachimiak, M.; Walther, D.; Cohen, F. E. Co-evolution of proteins with their interaction partners. *J. Mol. Biol.* **2000**, *299*, 283–293.
- [186] Landgraf, R.; Xenarios, I.; Eisenberg, D. Three-dimensional cluster analysis identifies interfaces and functional residue clusters in proteins. *J. Mol. Biol.* **2001**, *307*, 1487–1502.
- [187] Armon, A.; Graur, D.; Ben-Tal, N. ConSurf: An algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J. Mol. Biol.* **2001**, *307*, 447–463.
- [188] Madabushi, S.; Yao, H.; Marsh, M.; Kristensen, D. M.; Philippi, A.; Sowa, M. E.; Lichtarge, O. Structural clusters of evolutionary trace residues are statistically significant and common in proteins. *J. Mol. Biol.* **2002**, *316*, 139–154.
- [189] Lichtarge, O.; Sowa, M. E. Evolutionary predictions of binding surfaces and interactions. *Curr. Opin. Struct. Biol.* **2002**, *12*, 21–27.
- [190] Pupko, T.; Bell, R. E.; Mayrose, I.; Glaser, F.; Ben-Tal, N. Rate4Site: An algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics* **2002**, *18*(Suppl 1), S71–S77.
- [191] del Sol Mesa, A.; Pazos, F.; Valencia, A. Automatic methods for predicting functionally important residues. *J. Mol. Biol.* **2003**, *326*, 1289–1302.
- [192] Pazos, F.; Valencia, A. In silico two-hybrid system for the selection of physically interacting protein pairs. *Proteins* **2002**, *47*, 219–227.
- [193] Shannon, C.; Weaver, W. *Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, 1963.
- [194] Vakser, I. A.; Jiang, S. Strategies for modeling the interactions of transmembrane helices of G protein-coupled receptors by geometric complementarity using the GRAMM computer algorithm. *Methods Enzymol.* **2002**, *343*, 313–328.
- [195] Koshi, J. M.; Goldstein, R. A. Context-dependent optimal substitution matrices. *Protein Eng.* **1995**, *8*, 641–645.
- [196] Soyer, O. S.; Dimmic, M. W.; Neubig, R. R.; Goldstein, R. A. Dimerization in amineergic G-protein-coupled receptors: Application of a hidden-site class model of evolution. *Biochemistry* **2003**, *42*, 14522–14531.
- [197] Soulier, J. L.; Russo, O.; Giner, M.; Rivail, L.; Berthouze, M.; Ongeri, S.; Maigret, B.; Fischmeister, R.; Lezoualc'h, F.; Sicsic, S.; Berque-Bestel, I. Design and synthesis of specific probes for human 5-HT4 receptor dimerization studies. *J. Med. Chem.* **2005**, *48*, 6220–6228.

- [198] Chen, R.; Li, L.; Weng, Z. ZDOCK: An initial-stage protein-docking algorithm. *Proteins* **2003**, *52*, 80–87.
- [199] Casciari, D.; Dell'Orco, D.; Fanelli, F. Homodimerization of neurotensin 1 receptor involves helices 1, 2, and 4: Insights from quaternary structure predictions and dimerization free energy estimations. *J. Chem. Inform. Model* **2008**, *48*, 1669–1678.
- [200] Dell'Orco, D.; De Benedetti, P. G.; Fanelli, F. In silico screening of mutational effects on transmembrane helix dimerization: Insights from rigid-body docking and molecular dynamics simulations. *J. Phys. Chem. B* **2007**, *111*, 9114–9124.
- [201] Guo, W.; Shi, L.; Filizola, M.; Weinstein, H.; Javitch, J. A. Crosstalk in G protein-coupled receptors: Changes at the transmembrane homodimer interface determine activation. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17495–17500.
- [202] Filizola, M.; Wang, S. X.; Weinstein, H. Dynamic models of G-protein coupled receptor dimers: Indications of asymmetry in the rhodopsin dimer from molecular dynamics simulations in a POPC bilayer. *J. Comput. Aided Mol. Des.* **2006**, *20*, 405–416.
- [203] Bruno, A.; Guadix, A. E.; Costantino, G. Molecular dynamics simulation of the heterodimeric mGluR2/5HT(2A) complex. An atomistic resolution study of a potential new target in psychiatric conditions. *J. Chem. Inform. Model* **2009**, *49*, 1602–1616.
- [204] Periole, X.; Huber, T.; Marrink, S.-J.; Sakmar, T. P. G protein-coupled receptors self-assemble in dynamics simulations of model bilayers. *J. Am. Chem. Soc.* **2007**, *129*(33), 10126–10132.
- [205] Niv, M. Y.; Filizola, M. Influence of oligomerization on the dynamics of G-protein coupled receptors as assessed by normal mode analysis. *Proteins* **2008**, *71*, 575–586.
- [206] Taylor, M. S.; Fung, H. K.; Rajgaria, R.; Filizola, M.; Weinstein, H.; Floudas, C. A. Mutations affecting the oligomerization interface of G-protein-coupled receptors revealed by a novel de novo protein design framework. *Biophys. J.* **2008**, *94*, 2470–2481.