Published in final edited form as:
Methods. 2011 December ; 55(4): 405–414. doi:10.1016/j.ymeth.2011.12.005.

Characterizing and Predicting the Functional and Conformational Diversity of Seven-Transmembrane Proteins

Ravinder Abrol[a,*], **Soo-Kyung Kim**[a], **Jenelle K. Bray**[b], **Adam R. Griffith**[a], and **William A. Goddard III**[a]

[a]Materials and Process Simulation Center (MC 139-74), California Institute of Technology, Pasadena, CA 91125

[b]Simbios, NIH Center for Biomedical Computation, Stanford University, Stanford, CA 94305

Abstract

The activation of seven-transmembrane receptors (7TMRs) allows cells to sense their environment and convert extracellular signals (like hormone binding) into intracellular signals (through G protein-coupled and/or β arrestin-coupled pathways). A single 7TMR is capable of transducing a wide spectrum of physiological responses inside a cell by coupling to these pathways. This intracellular pleiotropic action is enabled by multiple conformations exhibited by these receptors. Developments in membrane protein structure determination technologies have led to a rapid increase in crystal structures for many 7TMRs. Majority of these receptors have been crystallized in their inactive conformation and, for some, one of the many active conformations has also been crystallized. Given the topological constraints of a lipid bilayer that results in a single fold of seven almost parallel TM helices connected by mostly unstructured loops, these structures exhibit a diversity of conformations not only across the receptors but also across the different functional forms for receptors with structures for one of the functionally active conformations. Here we present a method to characterize this conformational diversity in terms of **T**rans**M**embrane **H**elix **TOP**ology (TMHTOP) parameters (TMHTOP) and how to use these helix orientation parameters to predict functionally-distinct multiple conformations for these receptors. The TMHTOP parameters enable a quantification of the structural changes that underlie 7TM activation and also sheds a unique mechanistic light on the pleiotropic nature of these receptors. It provides a common language to describe the 7TMR activation mechanisms as well as differences across many receptors in terms of visually intuitive structural parameters. Protein structure prediction methods can use these parameters to describe 7TMR conformational ensembles, which coupled to experimental data can be used to develop testable hypotheses for the structural basis of 7TMR functions.

Keywords

G protein-coupled receptors; GPCRs; Conformational ensemble; Functional selectivity; Protein structure prediction; Transmembrane Helix Topology

*Corresponding author; abrol@wag.caltech.edu.

1. Introduction

The seven-transmembrane receptors (7TMRs) are integral membrane proteins that form the largest superfamily in the human genome with ~800 receptors identified, including ~370 non-sensory receptors organized in 5 families (GRAFS): Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2, and Secretin [1, 2]. A variety of bioactive molecules, including biogenic amines, peptides, and hormones modulate the activity of 7TMRs to effect regulation of essential physiological processes (e.g. neurotransmission, cellular metabolism, secretion, cell growth, immune defense, and differentiation) through G protein-coupled and/or β arrestin-coupled pathways [3]. The coupling to G proteins is the origin of 7TMRs being commonly referred to as GPCRs (G protein-coupled receptors). A structural understanding of their function will have a tremendous and broad impact, asthe dysregulation of these receptors often plays an important role in disease pathologies [2, 4].

The pleiotropic nature of these receptors in response to different extracellular signals is the result of different ensembles of protein conformations that are stabilized in apo-wild-type, apo-mutant, and ligand-bound forms, leading to different functional states of the receptors [3]. The current structural, thermodynamic, and functional knowledge of 7TMRs suggests an emerging conformational-ensemble picture like one shown in Fig. 1, which provides a schematic of lowest energy 7TMR conformations in different scenarios. "**I**" refers to the most stable (lowest energy) inactive conformation for a receptor. "C" refers to the conformations besides the "**I**" conformation that are thermally accessible for constitutively active receptors. "**G**" refers to conformations that can exclusively couple to G proteins. "**β**" refers to conformations that can exclusively couple to β arrestins and "**G/β**" refers to conformations that can couple to G proteins as well as β arrestins. This figure will look different for different receptors and is drawn to provide a conceptual "conformational ensemble" framework to think about 7TMRs and what can potentially happen to their ensembles in mutant forms and/or in the presence of different ligands. It would be impractical to list all the possibilities. This figure suggests that it is possible to stabilize functionally different conformations, e.g., in the presence of an agonist B plus G protein (6^{th} column in Fig. 1), a "**G**" conformation (active) may be stabilized (as observed for β2 adrenergic receptor bound to Gs or Gs-mimic[5, 6]), whereas an agonist A by itself (4^{th} column in Fig. 1) may still stabilize the inactive conformation (as observed for the β2 adrenergic receptor bound to an agonist [7]) and in the presence of an agonist A, the mutant Y of the receptor (7^{th} column in Fig. 1) may stabilize a "**β**" conformation (that exclusively couples to β arrestins). This is inherently complex, but it does promise the possibility of a specific ligand or mutant or ligand-mutant combination stabilizing a conformation that may be responsible for a specific function. Thus, a full structural insight into the pleiotropic function of 7TMRs will require the functional characterization of their multiple conformations to highlight how different ligands or mutations modulate (stabilize and/or induce) conformations to cause their effects [3, 8–11]. To complement these functional studies, any protein modeling approach for 7TMRs should be able to predict these ensembles of protein conformations and how they evolve in the presence of ligands and mutations.

This also has profound implications for the ability to control receptor pharmacology. The lack of experimental structures for most human 7TMRs (none were available until 2007) has led to the use of high-throughput screening (HTS) and virtual ligand screening (VLS) techniques in drug design, which do not require protein structures. A large proportion (43%) of drug candidates in clinical trials fail due to lack of efficacy, toxicity and side-effects [12]. Some side effects result from drugs binding to other 7TMRs or other subtypes of the same receptor (e.g., pramipexole which is used to treat Parkinson's disease [as a dopamine D2 receptor agonist] can produce behavioral side-effects such as compulsive gambling

associated with dopamine D3 receptor) [13]. Such side effects might be minimized by designing D2 selective agonists, which requires atomic-level structures for both D2 and D3 receptors. Some side effects are caused by the target receptor when ligand activates both beneficial and undesirable signaling pathways. An example is the niacin receptor GPR109A. Niacin is therapeutically beneficial as an antilypolytic agent (via G protein mediated pathways), but also causes cutaneous flushing which has been directly linked to the activation of β arrestin 1 pathways [14]. An analog of this molecule will be highly desirable that doesn't affect the G protein pathways but destabilizes the coupling to β arrestin 1 and blocks that pathway. This picture states that different ligands (and mutants) can stabilize different 7TMR conformations by changing its wild-type energy landscape, which can be further modulated by the presence of G proteins or β arrestins. Such detailed understanding of the effects of a drug candidate on different signaling pathways can assist in designing drugs with minimal on-target side effects, which in turn requires a structural understanding of the different conformations that may be coupling to different pathways.

The concepts in Fig. 1 also promise that an understanding of how these structural ensembles change can, in principle, allow us to stabilize multiple functionally-distinct conformations to the point of being crystallizable. In the last few years there has been a rapid increase in the solution of crystal structures for many 7TMRs[5, 7, 15–25] due mainly to a technological revolution in membrane protein structure determination methods [26]. Crystal structures corresponding to inactive forms are now available for the β_2 adrenergic [16], A_{2A} adenosine [17], D_3 dopamine [21], $CXCR_4$ chemokine [22], and H_1 histamine receptors in humans. Inactive structures are also available for bovine rhodopsin [15, 23], turkey β_1[18], and squid rhodopsin [19, 25] receptors. The structures of bovine opsin with/without a Gα peptide [20, 24], bovine metarhodopsin II [27], agonist-bound human β_2 adrenergic receptor stabilized by a G-protein-mimicking nanobody [5] or very recently by the Gs heterotrimer [6], and agonist bound adenosine A_{2A} receptor [28] are the first set of structures to provide a direct glimpse of one or more of the "active" states of those receptors. The topological comparison of many of these crystal structures and their implications for 7TMR activation has been reviewed extensively [29–33]. Structure determination efforts are moving towards the stabilization of 7TMRs in different functional conformations (e.g., bound to agonists, G proteins or β arrestins), and structural computational biology can help by mapping the energy landscape sampled by 7TMRs during their life-cycle and characterizing the critical conformations along the way to link with those that are observed in experimental structures to develop testable hypotheses for their structure-function relationships. A characterization of the above mentioned structural diversity of 7TMR conformational ensembles in terms of a few intuitive parameters can help in analyzing the experimental structures and also assist in the modeling of these receptors in different functional forms, allowing protein structure prediction and modeling to play an increasingly important role in providing detailed structural information about 7TMR activation and ligand binding.

Membrane proteins have been the focus of structure prediction and dynamics simulations for some years [34]. The interaction of these proteins with their lipid environment is critical for folding and studies have attempted to quantify this interaction on an absolute thermodynamic basis [35] by providing thermodynamic costs for the insertion of TM helical amino acids into the lipid bilayer [36]. The three-dimensional structure of these α-helical membrane proteins is also affected by interhelical interactions (H-bonds and salt-bridges) [35]. An accurate structure prediction methodology needs to be able to sample and describe these interhelical interactions thoroughly.

Many methods have been used to obtain model structures for membrane proteins. These methods have been reviewed elsewhere [37, 38]. For 7TMRs the main approach has been homology modeling and most studies have used constraints based on mutation and binding

experiments coupled to the homologous rhodopsin structure to guide additional mutation experiments. This approach has not been sufficiently accurate for predicting binding sites of ligands [39]. Our group has been developing *de novo* computational approaches (not homology-based) for predicting the structures of 7TMRs and their ligand binding sites. We have been successful in predicting 7TMR structures (mostly inactive, some active) validated by known mutation and ligand SAR data [40–49], as reviewed here [50]. These approaches sampled receptor conformations using only TM helix rotation angles, which (we have found) is inadequate to predict structures for 7TMRs distant (in sequence space) from a template or to predict multiple conformations for a given 7TMR.

To address this we have developed a method to characterize any 7TMR helix bundle conformation in terms of relative orientation describing TMHTOP parameters for all seven TM helices. This analysis method allows for the comparison of any 7TMR conformation against any other 7TMR conformation. This characterization of differences across all known templates (all 7TMR crystal conformations and validated predicted conformations), when correlated with sequence identity (or similarity), provides information vital for protein modeling efforts to assess the conformational space sampling that may be necessary in starting from any template. This comparison done between inactive and active conformations (that have been obtained for some 7TMRs) can provide mechanistic insight into activation in terms of the helix degrees of freedom (DOFs) that vary the most during activation. This method ignores the loops to characterize the 7TMRs for two reasons. First, secondary structure (helices or sheets) is present only in part of the loops making it difficult to define simple orientation parameters for them. Second, the positioning of loops above and below the lipid bilayer is governed by the orientations and lengths of TM helices connected to them making it difficult to characterize their orientation with respect to the membrane middle plane.

In the next section we will describe the method to characterize the orientation of any 7TMR conformation and then correlate the orientation parameters with the sequence-based measures like identity or similarity. These correlations provide valuable insight into the conformational diversity across different 7TMRs and how these parameters can be used in protein structure prediction workflows aimed at 7TMRs. We have used these orientation parameters to develop and test rigorous conformational sampling methods to provide an ensemble of thermally accessible apo and holo conformations, corresponding to different functional states of the receptor. To this effect we have developed a method called GEnSeMBLE (GPCR Ensemble of Structures in the Membrane BiLayer Environment) that uses these orientation parameters to sample TM helix bundle conformations and predict low energy conformations for man 7TMRs (see, for example, ref. [46]). Some results of the application of this method will be highlighted to show the usefulness of the orientation parameters.

2. Methods

The seven TM helix topology of 7TMRs presents unique advantages and challenges for the quantification of sequence-structure relationships. Many comparative modeling programs can predict structures of globular proteins (with 30% or higher sequence identity to a crystallized protein) to a reasonable accuracy, as the belief is that a major fraction of structural folds are now known for globular proteins. Same cannot be said for membrane proteins in general. 7TMRs, however, can be thought of having one structural fold that crosses the membrane bilayer seven times in the form of seven TM helices interconnected by intracellular and extracellular loops. In spite of the boundary conditions imposed the membrane and lengths of loops, these seven helices can be packed together in innumerable ways. We will describe below how to quantify the topology of these TM helix packings in

absolute terms using TMHTOP orientation parameters describing each of the seven TM helices, which can be used to characterize different receptor structures or different conformations of the same receptor.

2.1 TM Helix Orientation Coordinate System

To quantify the relationship between sequence and structure for 7TMRs, we need to characterize the known structures using some standard geometrical parameters. As crystal structures don't provide absolute membrane orientation of 7TMRs, we use the OPM (Orientation of Proteins in Membrane, [51]) database, which aligns each newly deposited membrane protein structure to an implicit membrane maximizing the free energy of insertion of the membrane protein into the membrane. The middle of the membrane corresponds to the z=0 plane, which we designate asthe hydrophobic plane. Each 7TMR structure can then be characterized by the six orientation parameters of the seven helices as defined in Fig. 2, which shows how the helix position and orientation are defined in the frame of this z=0 hydrophobic plane.

Helix position (R) on the hydrophobic plane is defined by the first two coordinates x and y. Third parameter h corresponds to the hydrophobic center (HPC) residue from the helix that will be positioned on the (z=0) hydrophobic plane. In practice, this hydrophobic center residue is a non-integer to signify that the positioning of a TM helix in the z=0 plane could lie between two residues. Fourth parameter is the angle θ that refers to the helix tilt angle signifying how much a helix tilts with respect to the z-axis or the membrane normal. This angle can be directly associated with determining the hydrophobic match of the helix with the membrane bilayer. Fifth parameter is the angle ϕ, called the helix sweep angle, which signifies the azimuthal angle of the helix for a fixed value of the helix tilt angle θ. So, the sweep angle ϕ (for a fixed θ) samples a helix degree of freedom (DOF) that preserves the hydrophobic match with the membrane. The last orientation parameter is the helix rotation angle η which corresponds to the helix rotation angle about its (as yet undefined) axis. This angle requires the specification of an η residue (ηRES) which is used as a reference to define this rotation angle uniquely. For class A 7TMRs, this residue is chosen to be the most conserved residue in each helix (given by Ballesteros numbering scheme X.50[52]) where X is the TM helix in question. The exception to this is TM helix 3, for which the residue position 3.32 is used as ηRES because residue position 3.50 is located at the bottom of this helix, which may not provide a consistent η rotation angle representing the whole helix. The helix DOF controlled by this rotation angle most directly controls which face of a specific helix may be interacting with another helix in the TM bundle.

The helixtilt angle θ, sweep angle ϕ, and rotation angle η require a unique definition of the helical axis which needs to account for the reality of bent helices found in integral membrane proteins, as Proline residues are commonly found in their TM helices. We use a helical axis that corresponds to the least moment of inertia vector obtained by the eigensolution of the moment of inertia matrix for the helix using only heavy backbone atoms. We rotate the membrane-aligned 7TMRs from the OPM database in the x-y plane such that the helical axis of TM helix 3 goes through the origin, and that of TM helix 2 intersects the x-axis. This transformation coupled to the convention that extracellular face of the receptor points in the +z direction results in a unique orientation reference frame for 7TMRs, which will allow for a consistent and quantitative comparison of the helix orientation parameters across different receptors or across different conformations of the same receptor. As an example, Table 1 shows the six TM helix orientation parameters for the first crystallized human β2 adrenergic receptor structure [16]. It also lists the ηRES used to define the helix rotation angle η for each TM helix.

2.2 TM Helix Orientation Parameters for Inactive 7TMR Conformations

The TMHTOP helix orientation parameters for all seven helices in a 7TMR provide a quantitative view of its topology in the membrane and can be directly compared against another receptor to quantify topological deviations across receptors. Of the six orientation parameters, five parameters (x,y,θ, ϕ, η) can be directly compared across different receptors to capture the topological diversity. The sixth parameter h specifies the hydrophobic residues aligned on the membrane middle plane and is less insightful in capturing 7TM helix bundle differences. Furthermore, we reduce the helix position parameters x,y into a single parameter R that captures the distance from origin of the point of intersection of the helical axis and the hydrophobic plane. So, the deviation of this distance from origin between two 7TMRs *i* and *j* for a TM helix *k* is given by:

$$\Delta R_k^{i,j} = \sqrt{\left[\left(x_k^i - x_k^j\right)^2 + \left(y_k^i - y_k^j\right)^2\right]} \quad (1)$$

The corresponding deviation of the helix orientation angles is given by:

$$\Delta\alpha_k^{i,j} = \left(\alpha_k^i - \alpha_k^j\right), \text{ where } \alpha = \theta, \phi, \text{ or } \eta \quad (2)$$

This analysis has been applied to inactive conformations of 7TMRs that have been crystallized so far, which is expected to provide useful insight into the topological diversity across these different 7TMRs. The inactive conformations used in this comparison are the following: bovRhod (pdbid: 1u19 [23]), humBeta2 (pdbid: 2rh1 [16]), turBeta1 (pdbid: 2vt4, chainB [18]), humA2A (pdbid: 3eml[17]), squRhod (pdbid: 2z73[19]), humD3 (pdbid: 3pbl[21]), humCXCR4 (pdbid: 3odu[22]), and humH1 (pdbid: 3rze[53]). All conformations were pre-aligned to humBeta2 OPM conformation (2rh1) to capture absolute differences across conformations and to reduce the potential for noise coming from different OPM orientations.

The panels in Figure 3 attempt to capture the topological differences across this set of inactive receptor conformations. Figure 3A shows the deviations in helix position (ΔR) as a function of the corresponding deviations in helix rotation angle (Δη) across all eight inactive receptor conformations in green and across the four aminergic receptors' (humBeta2, turBeta1, humD3, and humH1) conformations in red. Each deviation point in Figure 3 is marked by the TM helix number that it corresponds to, so the panels contain a lot of information. Figures 3B and 3C show the deviations in helix tilt angle (Δθ) and helix sweep angle (Δϕ) respectively as a function of the corresponding deviations in helix rotation angle (Δη). Figures 3D and 3E show the deviations in helix position (ΔR) and helix tilt angle (Δθ) respectively as a function of the corresponding deviations in helix sweep angle (Δϕ). Figure 3F shows the deviations in helix position (ΔR) as a function of the corresponding deviations in helix tilt angle (Δθ). So, all pairwise correlations of these deviations are represented in Figure 3. The deviations across aminergic receptors (shown in red points) are smaller than those across all eight receptors as expected. Based on Figures 3A, 3D, and 3F, TM5 shows the biggest deviation in its position ΔR (upto 6Å) across the eight receptors, whereas it is TM1 that shows the biggest deviation in its position (upto about 4Å) across the aminergic receptors. Figures 3B, 3E, and 3F show that TM6 displays biggest deviation in its tilt angle (Δθ) across aminergic as well as all receptors, with deviation among aminergic receptors (upto 5°) significantly smaller than that among all receptors (upto 15°). Figures 3C, 3D, and 3E show that TM4 and TM6 display the biggest deviations in their helix sweep angles (Δϕ) for both sets of receptors (upto about 80°). Figures 3A, 3B, and 3C show that TM2 and TM5 display the biggest deviations in their helix rotation angles (Δη) among all receptors (upto 60°), whereas it is TM6 that shows the biggest variability among aminergic receptors (upto

30°). Overall, this analysis suggests that inactive 7TMR conformations exihibit large topological diversity, which will be difficult to capture by homology modeling. Even aminergic receptors display non-negligible topological differences, some of which might be important for capturing the properties associated with their inactive conformations like binding to different antagonists and inverse agonists.

2.3 TM Helix Orientation Parameters for Active 7TMR Conformations

The analysis described in the previous section has also been applied to the active conformations of 7TMRs that have been captured recently in some crystal structures. The topologies of these active conformations have also been compared to the inactive conformations analyzed earlier. Three 7TMRs have been crystallized so far in multiple functionally distinct conformations: Bovine Rhodopsin, Human β_2 Adrenergic Receptor, and Human Adenosine A_{2A} Receptor. This limited number of available active conformations is not enough to make conclusions about GPCR activation mechanisms, but the following analysis provides an introduction to the type of analysis enabled by TMHTOP orientation parameters.

For Bovine Rhodopsin, the following active conformations were compared to the inactive 1U19 (pdbid) conformation [23] and to each other: bovine Opsin (pdbid: 3cap [20]), rhodopsin mutant E113Q with Gα C-terminal peptide (pdbid: 2×72 [54]), and metarhodopsin II (pdbid: 3pqr [27]). For Human β2 Adrenergic Receptor, the following active and inactive conformations were compared to the inactive 2rh1 (pdbid) conformation [16] and to each other: humBeta2 bound to an inverse agonist (pdbid: 3d4s [55]), humBeta2 bound to inverse agonists and a neutral antagonist (pdbids: 3nya, 3ny8, 3ny9 [56]), humBeta2 bound to an irreversible agonist (pdbid: 3pds [7]), humBeta2 bound to an agonist along with a nanobody (pdbid: 3p0g [5]), and humBeta2 bound to the Gs protein heterotrimer (pdbid: 3sn6[6]). For Human Adenosine A_{2A} Receptor, the following active conformations were compared to the inactive 3eml (pdbid) conformation [17] and to each other: humA2A bound to agonist UK432097 (pdbid: 3qak [28]), humA2A bound to agonist adenosine (pdbid: 2ydo [57]), and humA2A bound to agonist NECA (pdbid: 2ydv [57]). In all cases, the different conformations were pre-aligned to the corresponding reference inactive conformation from the OPM database, which was 1u19 for bovRhod, 2rh1 for humBeta2 and 3eml for humA2A. The results are shown in Figure 4, with rhodopsin comparisons in blue, β_2 comparisons in green and A_{2A} comparisons in red.

Comparisons of active Rhodopsin conformations to the inactive conformation (blue points in Figure 4) show biggest movements for TM2, TM5 and TM6. The position (R) of TM2 moves the most (~5–6Å), followed by TM5 and TM6 (~3–4Å). Deviations in helix tilt angle ($\Delta\theta$) stay below 5° for all TMs. TM6 shows the biggest deviations in helix sweep angle ($\Delta\phi$ ~ 45°). TM6 also shows the biggest deviations in helix rotation angle ($\Delta\eta$ ~ 40°), followed by TM3 which shows a rotation of about 30°. Comparisons of β_2 conformations (green points in Figure 4) show TM5, TM6 and TM7 with most deviations. TM position doesn't deviate by more than 2.5Å (TM7). TM5 shows the largest helix title angle deviation ($\Delta\theta$ ~ 5°). TM6 shows the biggest deviation in helix sweep angle ($\Delta\phi$ ~ 80°) followed by TM4 (~40°). Helix rotation angle deviation ($\Delta\eta$) stays below 20° for all TMs. In the comparison of active A_{2A} conformations to the inactive one (red points in Figure 4), TM6 and TM7 stand out with the largest deviations. TM6 shows large deviation in helix sweep angle ($\Delta\phi$ ~ 80°) like in β2. TM7 shows a small deviation in helix position (ΔR~2Å) but a large deviation in helix rotation angle ($\Delta\eta$ ~ 75°). The deviations in β_2 and A_{2A} conformations behave differently from those in rhodopsin as noticed clearly in Figure 4C, which shows that helix sweep angle (ϕ) and helix rotation angle (η) deviations for TM6 are not coupled for β_2 and A_{2A} conformations, but they are for rhodopsin.

The analysis shown in Figure 4 provides topological fingerprints for active conformations which can be coupled to known function from biochemical experiments. This coupling can allow one to hypothesize about the function of any observed topologies. Different topological fingerprints would invariably be associated with some type of different downstream functional outcome, but same fingerprints may not always lead to same function. An example is the observation of "islands" of points in Figure 4 of identical TMs for a given receptor. Since these "islands" correspond to multiple active conformations that have been crystallized, they show that those active conformations have very similar TM helix topologies, e.g. blue points in Figure 4 (for rhodopsin) show islands of TM2, TM5 and TM6 deviations. These islands capture deviations of opsin, for example, as well as metaII from the dark state of rhodopsin, suggesting that opsin and metaII have very similar TM helix topologies. So, their known functional differences could be most likely ascribed to potential differences in their intracellular loops and/or their C-termini.

Figure 5 shows a minimal version of Figure 4, where only a single inactive-to-active comparison is included per receptor using the following conformations. For rhodopsin, the blue points in Figure 5 compare the dark state (pdbid: 1u19 [23]) with the metaII state (pdbid: 3pqr [27]). For β_2, the green points in Figure 5 compare the inactive conformation bound to carazolol (pdbid: 2rh1 [16]) with the recently crystallized "active" conformation bound to the full heterotrimeric Gs protein (pdbid: 3sn6 [58]). For A_{2A}, the red points in Figure 5 compare its inactive conformation bound to antagonist ZM241385 (pdbid: 3eml [17]) with the recently crystallized active conformation bound to the agonist UK432097 (pdbid: 3qak [28]). Data from this figure shows that as a consensus TM6 appears to be the most critical TM helix that moves during activation and its consensus motion occurs in the helix sweep angle (ϕ) space. This motion appears to be the only common denominator based on the three receptors with experimental active conformations. The differences across these receptors in terms of structural changes during activation might be specific to sub families of receptors more closely related to these receptors.

These analyses highlight the fact that the TMHTOP TM helix orientation parameters provide a standard way to quantify differences across various receptors and across different conformations of the same receptor, which in turn enables the generation of testable hypotheses about mechanistic insights into the function in these receptors.

Next, we will compare the differences in these orientation parameters across the receptors with their sequence identity and similarity, to provide a metric for the reliability of homology-based models vs *de novo* structures.

2.4 Sequence vs Topology

The TM regions display high sequence conservation when compared to the loop regions as is expected (see Table 2). The table shows an all-to-all comparison of the sequence identity (Table 2A) and sequence similarity (Table 2B) across all eight 7TM receptors that have been crystallized so far in one or more conformations (bovRhod: bovine Rhodopsin; humBeta2: human β2 adrenergic receptor; turBeta1: turkey β1 adrenergic receptor; humA2A: human Adenosine A_{2A} receptor; squRhod: squid Rhodopsin; humD3: human Dopamine D_3 receptor; humCXCR4: human Chemokine $CXCR_4$ receptor; humH1: human histamine H_1 receptor). Similarity is defined by using the BLOSUM62 matrix[59], where two residues are considered similar if the corresponding substitution element in the BLOSUM62 matrix is greater than zero. For each pair of sequences, two numbers are provided in Table 2 for identity and two for similarity in the form M/N, where M corresponds to identity (or similarity) for the full sequences and N corresponds to identity (or similarity) for the TM helix portions of the sequences. The TM helix portions were inferred from the HELIX fields in the corresponding PDB files of the receptors' inactive conformations. For each 7TM

receptor column in Table 2, the receptors with highest and lowest TM region measure (identity or similarity) to that 7TM receptor are highlighted with green and red cells respectively. For example, humBeta2 column in Table 2A shows that in terms of TM region sequence identity, it is closest to turBeta1 and farthest from bovRhod. A curious observation from Table 2B is that turBeta1 is closest to five out of seven sequences, whereas bovRhod and humCXCR4 are farthest from four out of seven sequences. Another useful observation from Table 2B is that aminergic receptors (humBeta2, turBeta1, humD3, and humH1) are all above 60% similarity of each other in the TM regions.

This comparison when done for a new 7TM receptor (which may be a structure prediction target) with sequences of known experimental structures provides first hints of good structural templates to use for the application of protein structure prediction methodologies, be they *de novo* methods or homology based methods. One may expect conservatively that the homology based methods have a good chance of working at least across aminergic receptors. We will use an example later to compare homology-based and *de novo* predicted structures using two templates to suggest a criteria to choose homology-based or *de novo* based methods for reliable structure prediction of 7TMRs.

In order to correlate the sequence variability of the crystallized 7TMRs with their TM helix geometries, we analyzed all-to-all (across these 7TMR systems) maximum absolute deviations in helix position *R*, and helix orientation angles θ, ϕ, η over all helices against the corresponding sequence identity and similarity. The maximum absolute deviations are defined between a 7TMR*i* and another 7TMR*j* by:

$$max\left[\Delta R_k^{i,j}, k=1\ldots 7\right] \tag{3}$$

$$max\left[\Delta \alpha_k^{i,j}, k=1\ldots 7\right], \text{where } \alpha=\theta,\phi,\eta; \tag{4}$$

The index *k* is for the 7 TM helices. Figure 6 shows this maximum all-to-all deviation (for all *i*-*j* pairs of 7TMRs) as a function of identity or similarity between 7TMR*i* and 7TMR*j*, to highlight the maximum variability seen in these eight systems for any TM helix and to find correlation to the sequence homology.

Table 3 shows two different correlation coefficients to quantify these correlations. All coefficients show an inverse correlation because one expects largest deviations for the least similar receptors. The deviation in TM helix tilt angle (θ) shows the least correlation to sequence homology, be it TM sequence identity or similarity. This is not surprising because this angle is strongly influenced by hydrophobic matching with the membrane in addition to interhelical interactions. The deviation in TM helix rotation angle (η) shows the highest correlation with sequence homology. This angle has the strongest influence on interhelical interactions as opposed to the other two angles, so it can be expected that this angle shows the strongest correlation to TM sequence similarity or identity.

2.5 Homology vs *de novo* methods for structure prediction of 7TMRs

As Table 1 showed, the eight receptor sequences analyzed above cover a broad sequence space in class A 7TMRs with sequence identity ranging from 69% down to 19%, and sequence similarity ranging from 83% down to 35%. Now, given a receptor with unknown structure, which of the eight receptor sequences provides the ideal template for its homology modeling or as a reasonable starting structure for its de novo structure prediction? Using the sequence homology information for this target receptor against the eight available template sequences, Figure 6 allows one to quantify the conformational space sampling necessary

within any structure prediction protocol. This comparison of observed changes in TMHTOP orientation parameters with sequence similarity allows one to estimate changes that might be expected in similar situations. These estimations are only going to improve as more 7TMR conformations are observed experimentally. This has strong implications for the structure prediction of 7TMRs. Many homology modeling based methods use molecular dynamics (MD) simulations to sample accessible conformational space, but these methods need to expand their conformational sampling capability to capture the diversity of distinct conformations observed for GPCRs. They are expected to work for only the most closely related receptors (TM sequence similarity > 60%), e.g., predicting humBeta2 structure from humBeta1 template, but not for predicting humBeta2 structure from humA2A template.

GEnSeMBLE (GPCR Ensemble of Structures in Membrane BiLayer Environment) [60] is one such method that uses the homology-based model only as a starting structure but then performs local but rigorous conformational sampling in the space specified by the three TMHTOP helix orientation angles (θ,ϕ,η). It is a de novo structure prediction method that implements a complete sampling of this space using a highly efficient but fast BiHelix algorithm[61], that splits up a seven-helix bundle problem into many two-helix bundle problems. The extent of sampling that may be necessary is pre-determined by the analysis of sequence similarity and changes in TMHTOP parameters seen so far (Figure 6). An application of this method is provided below to demonstrate that TMHTOP analysis need not be done in isolation but can be coupled to structure prediction methods to estimate the conformational space sampling that may be necessary, as such sampling can be computationally prohibitive. Similar de novo effort is required to predict active conformations starting from the inactive conformations as homology modeling and/or MD simulations cannot be expected to sample the conformational changes necessary for activation, especially because the active conformations have higher energy than the inactive conformations for the apo receptors.

To demonstration this structure prediction guidance provided by TMHTOP analysis, we predicted the structure of humA2A receptor starting with the humBeta2 receptor as a template. These sequences share a similarity of 50% in their TM regions, which is high by soluble proteins' standards and gives confidence in homology models for those soluble proteins. This will allow for the direct comparison of homology based models and those predicted by de novo methods. The starting structure for humA2A was built by placing crystal humA2A TM helices in the humBeta2 template. Using original crystal helices allows one to remove any noise in the procedure coming from the different TM helix shapes (between humA2A and humBeta2) and provides a clearer demonstration of the comparison between homology and de novo models. So, the starting structure can be considered the homology structure and has a CαRMSD of 2.1Å relative to the humA2A crystal TM bundle. The above mentioned three orientation angles will be used to sample the conformational space of ±10° in θ (with Δθ=10°), ±30° in ϕ (with Δϕ=15°), and ±30° in η (with Δη=15°) (with indications from Figure 6 for changes in TMHTOP parameters) using this homology model as a starting strucure. Even this limited space corresponds to ~13 trillion TM bundle conformations. The above mentioned GEnSeMBLE method of de novo structure prediction that samples θ,ϕ,η conformational space (described above) was applied to this TM bundle to see if the procedure can identify the native humA2A crystal conformation starting from the humBeta2 template. After the sampling procedure, the lowest energy conformation had an improved CαRMSD of 1.4Å relative to the crystal structure. Figure 7A compares this and the starting structure with the crystal structure.

The three receptor structures shown in Figure 7A were used in a docking exercise using MSCDock [62] to predict the binding site for humA2A ligand ZM241385. The predicted binding site of this ligand for the de novo predicted receptor structure (Figure 7C) shows

much better agreement with the co-crystal pose (Figure 7D). It forms a strong hydrogen bond with N253(TM6), consistent with the experimental structure. The docked ligand in the homology structure takes a very different pose, and does not form any hydrogen bond with N253(TM6). This shows that a homology model 2.1Å away in CαRMSD from the crystal structure was not accurate enough to predict the ligand binding site in the absence of any other experimental data (e.g. mutagenesis).

3. Conclusions

The number of experimental 7TMR structures has exploded since 2007, and now the structures not only cover a large sequence space for class A 7TMRs, but for many receptors more than one functionally distinct conformation. These receptors are characterized by a single-fold of 7 TM helices crossing the membrane seven times. We have presented an intuitive geometric set of TMHTOP helix orientation parameters that can quantify the topologies of all these single-fold structures using a standard set of definitions, which will allow for direct visual and quantitative comparison of the structures and conformations in terms of these parameters. As more 7TMR structures are solved, the comparison of these parameters across all available experimental structures will provide a quantitative view of the differences in topology exhibited by these critical receptors. These topological differences coupled to sequence homology information can guide the application of homology-based or de novo structure prediction methods and make the resulting structures more reliable. The same orientation parameters can describe the structural changes underlying activation of these receptors and shed critical light on the activation mechanisms to develop and test novel hypotheses of receptor function. Given the diversity of roles played by 7TMRs, the availability of these topological comparisons and associated functional hypotheses provides a common language and broad strategy to investigate the fundamental pleiotropic nature of 7TMRs.

References

1. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB. Mol Pharmacol. 2003; 63:1256–1272. [PubMed: 12761335]
2. Lagerstrom MC, Schioth HB. Nat Rev Drug Discov. 2008; 7:339–357. [PubMed: 18382464]
3. Kenakin T, Miller LJ. Pharmacol Rev. 2010; 62:265–304. [PubMed: 20392808]
4. Lundstrom K. Curr Protein Pept Sci. 2006; 7:465–470. [PubMed: 17073697]
5. Rasmussen SG, Choi HJ, Fung JJ, Pardon E, Casarosa P, Chae PS, Devree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI, Kobilka BK. Nature. 2011; 469:175–180. [PubMed: 21228869]
6. Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK. Nature. 2011; 477:549–555. [PubMed: 21772288]
7. Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK. Nature. 2011; 469:236–240. [PubMed: 21228876]
8. Strange PG. Biochem Pharmacol. 1999; 58:1081–1088. [PubMed: 10484065]
9. Vauquelin G, Van Liefde I. Fund Clin Pharmacol. 2005; 19:45–56.
10. Kobilka BK, Deupi X. Trends Pharmacol Sci. 2007; 28:397–406. [PubMed: 17629961]
11. Hoffmann C, Zurn A, Bunemann M, Lohse MJ. Brit J Pharmacol. 2008; 153:S358–S366. [PubMed: 18059316]
12. Schuster D, Laggner C, Langer T. Curr Pharm Design. 2005; 11:3545–3559.
13. Bostwick JM, Hecksel KA, Stevens SR, Bower JH, Ahlskog JE. Mayo Clin Proc. 2009; 84:310–316. [PubMed: 19339647]

14. Walters RW, Shukla AK, Kovacs JJ, Violin JD, DeWire SM, Lam CM, Chen JR, Muehlbauer MJ, Whalen EJ, Lefkowitz RJ. J Clin Invest. 2009; 119:1312–1321. [PubMed: 19349687]
15. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M. Science. 2000; 289:739–745. [PubMed: 10926528]
16. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. Science. 2007; 318:1258–1265. [PubMed: 17962520]
17. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC. Science. 2008; 322:1211–1217. [PubMed: 18832607]
18. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF. Nature. 2008; 454:486–491. [PubMed: 18594507]
19. Murakami M, Kouyama T. Nature. 2008; 453 363-U333.
20. Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP. Nature. 2008; 454:183–187. [PubMed: 18563085]
21. Chien EYT, Liu W, Zhao QA, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC. Science. 2010; 330:1091–1095. [PubMed: 21097933]
22. Wu BL, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC. Science. 2010; 330:1066–1071. [PubMed: 20929726]
23. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. Journal of molecular biology. 2004; 342:571–583. [PubMed: 15327956]
24. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP. Nature. 2008; 455 497-U430.
25. Shimamura T, Hiraki K, Takahashi N, Hori T, Ago H, Masuda K, Takio K, Ishiguro M, Miyano M. J Biol Chem. 2008; 283:17753–17756. [PubMed: 18463093]
26. Blois TM, Bowie JU. Protein Sci. 2009; 18:1335–1342. [PubMed: 19536805]
27. Choe HW, Kim YJ, Park JH, Morizumi T, Pai EF, Krauss N, Hofmann KP, Scheerer P, Ernst OP. Nature. 2011; 471 651-U137.
28. Xu F, Wu HX, Katritch V, Han GW, Jacobson KA, Gao ZG, Cherezov V, Stevens RC. Science. 2011; 332:322–327. [PubMed: 21393508]
29. Kobilka B, Schertler GFX. Trends Pharmacol Sci. 2008; 29:79–83. [PubMed: 18194818]
30. Mustafi D, Palczewski K. Mol Pharmacol. 2009; 75:1–12. [PubMed: 18945819]
31. Rosenbaum DM, Rasmussen SG, Kobilka BK. Nature. 2009; 459:356–363. [PubMed: 19458711]
32. Hanson MA, Stevens RC. Structure. 2009; 17:8–14. [PubMed: 19141277]
33. Worth CL, Kleinau G, Krause G. Plos One. 2009; 4
34. Fleishman SJ, Unger VM, Ben-Tal N. Trends Biochem Sci. 2006; 31:106–113. [PubMed: 16406532]
35. White SH. Adv Protein Chem. 2006; 72:157. [PubMed: 16581376]
36. Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, Lerch-Bader M, Nilsson I, White SH, von Heijne G. Nature. 2007; 450 1026-U1022.
37. Fanelli F, De Benedetti PG. Chem Rev. 2005; 105:3297–3351. [PubMed: 16159154]
38. Fleishman SJ, Ben-Tal N. Curr Opin Struc Biol. 2006; 16:496–504.
39. Yarov-Yarovoy V, Schonbrun J, Baker D. Proteins-Structure Function and Bioinformatics. 2006; 62:1010–1025.
40. Bray JK, Goddard WA 3rd. J Mol Graph Model. 2008; 27:66–81. [PubMed: 18499489]
41. Goddard WA 3rd, Abrol R. J Nutr. 2007; 137:1528S–1538S. discussion 1548S. [PubMed: 17513420]
42. Goddard WA, Kim SK, Li YY, Trzaskowski B, Griffith AR, Abrol R. J Struct Biol. 2010; 170:10–20. [PubMed: 20079848]
43. Heo J, Han SK, Vaidehi N, Wendel J, Kekenes-Huskey P, Goddard WA 3rd. Chembiochem. 2007; 8:1527–1539. [PubMed: 17647204]

44. Heo J, Vaidehi N, Wendel J, Goddard WA 3rd. J Mol Graph Model. 2007; 26:800–812. [PubMed: 17728165]
45. Kim SK, Li YY, Park C, Abrol R, Goddard WA. ChemMedChem. 2010; 5:1594–1608. [PubMed: 20683923]
46. Kim S-K, Riley L, Abrol R, Jacobson KA, Goddard WA 3rd. Proteins: Structure, Function, and Bioinformatics. 2011 In Press (Published online: 15 Feb 2011).
47. Li Y, Zhu F, Vaidehi N, Goddard WA 3rd, Sheinerman F, Reiling S, Morize I, Mu L, Harris K, Ardati A, Laoui A. J Am Chem Soc. 2007; 129:10720–10731. [PubMed: 17691773]
48. Peng JY, Vaidehi N, Hall SE, Goddard WA 3rd. ChemMedChem. 2006; 1:878–890. [PubMed: 16902941]
49. Vaidehi N, Schlyer S, Trabanino RJ, Floriano WB, Abrol R, Sharma S, Kochanny M, Koovakat S, Dunning L, Liang M, Fox JM, de Mendonca FL, Pease JE, Goddard WA 3rd, Horuk R. J Biol Chem. 2006; 281:27613–27620. [PubMed: 16837468]
50. Abrol R, Goddard WA. Abstr Pap Am Chem S. 2007; 234
51. Lomize MA, Lomize AL, Pogozheva ID, Mosberg HI. Bioinformatics. 2006; 22:623–625. [PubMed: 16397007]
52. Ballesteros, JA.; Weinstein, H. Methods in Neuroscience. Sealfon, SC.; Conn, PM., editors. San Diego: Academic Press; 1995. p. 366-428.
53. Stevens RC, Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Iwata S. Nature. 2011; 475 65-U82.
54. Standfuss J, Edwards PC, D'Antona A, Fransen M, Xie G, Oprian DD, Schertler GF. Nature. 2011; 471:656–660. [PubMed: 21389983]
55. Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EY, Velasquez J, Kuhn P, Stevens RC. Structure. 2008; 16:897–905. [PubMed: 18547522]
56. Wacker D, Fenalti G, Brown MA, Katritch V, Abagyan R, Cherezov V, Stevens RC. J Am Chem Soc. 2010; 132:11443–11445. [PubMed: 20669948]
57. Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG, Tate CG. Nature. 2011; 474:521–525. [PubMed: 21593763]
58. Rasmussen SG, Devree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK. Nature. 2011
59. Henikoff S, Henikoff JG. Proc Natl Acad Sci U S A. 1992; 89:10915–10919. [PubMed: 1438297]
60. Abrol, R.; Griffith, AR.; Bray, JK.; Goddard, WA, 3rd. Methods in Molecular Biology "Complementary experimental and computational techniques to study membrane protein structure, dynamics, and interactions". Vaidehi, N.; Klein-Seetharaman, J., editors. New York: Humana; 2011.
61. Abrol R, Bray JK, Goddard WA 3rd. Proteins: Structure, Function, and Bioinformatics. 2011 In Press.
62. Cho AE, Wendel JA, Vaidehi N, Kekenes-Huskey PM, Floriano WB, Maiti PK, Goddard WA. Journal of Computational Chemistry. 2005; 26:48–71. [PubMed: 15529328]

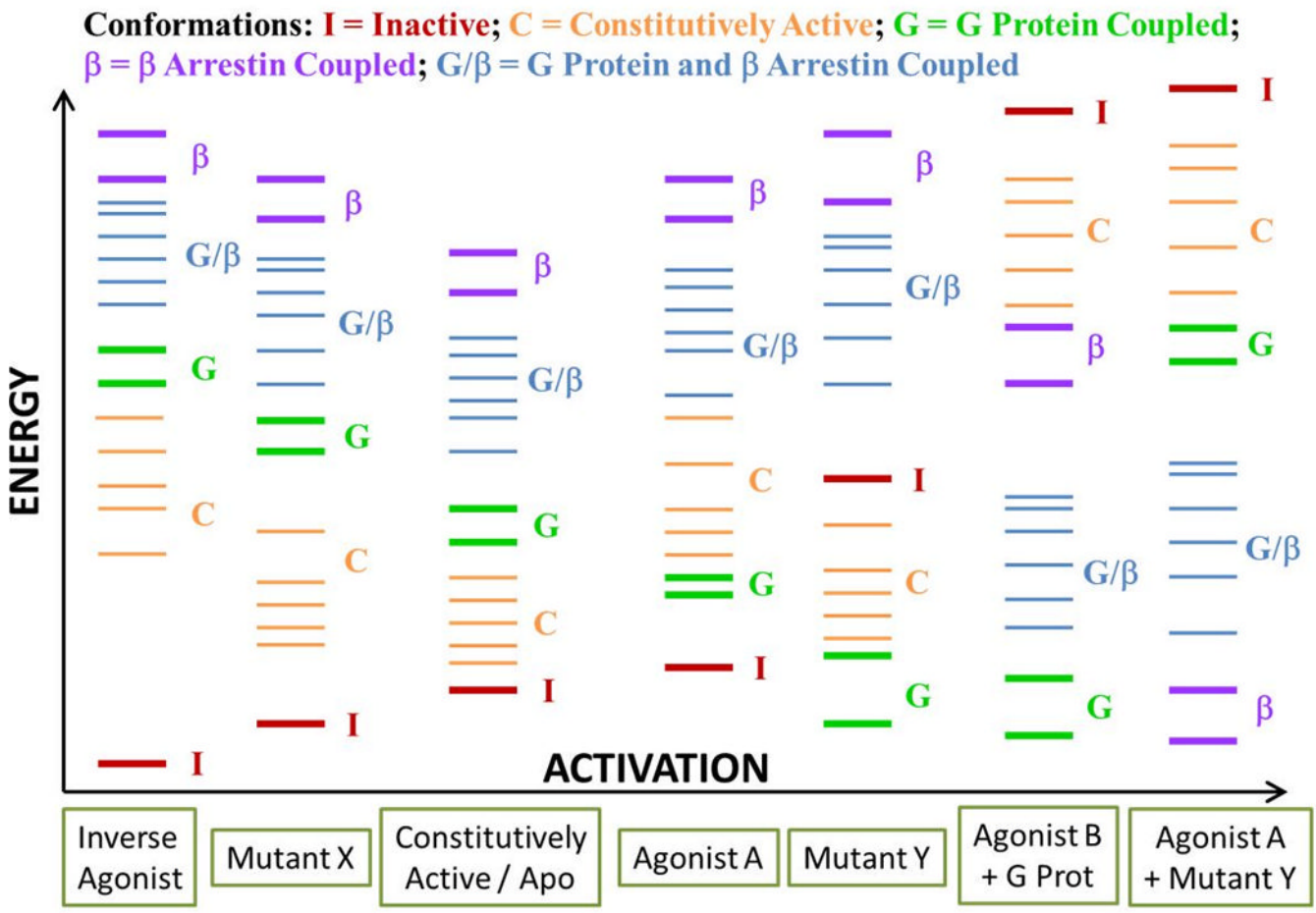

Figure 1.
Functional and thermodynamic view of 7TMR conformational ensembles.

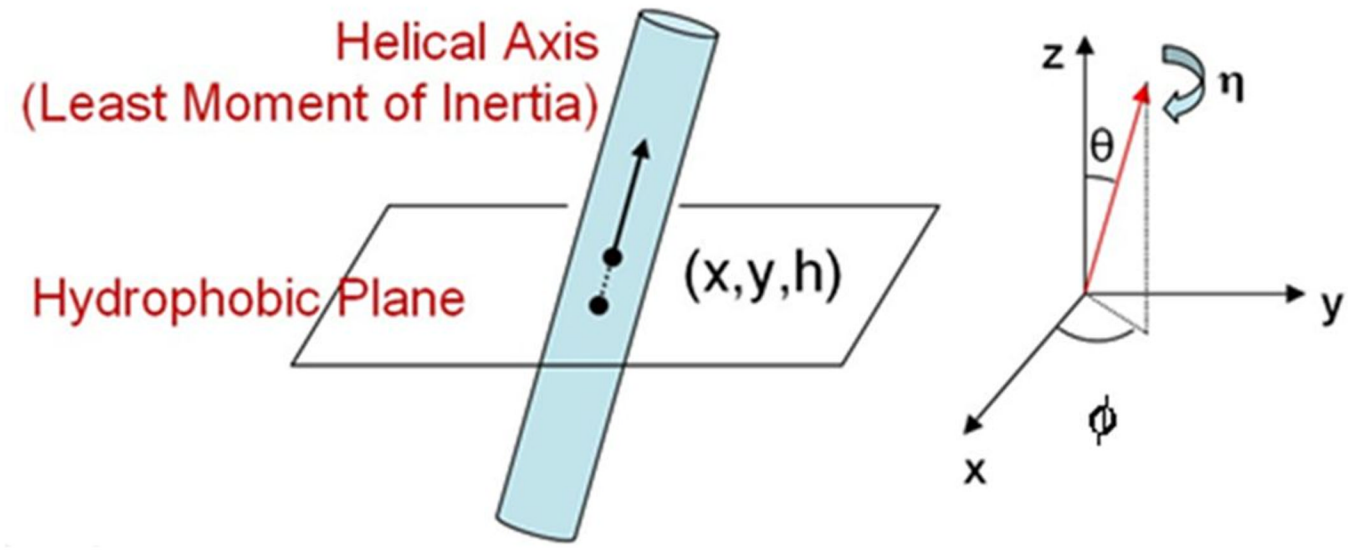

Figure 2.
Coordinates specifying the orientation of a TM helix in a membrane.

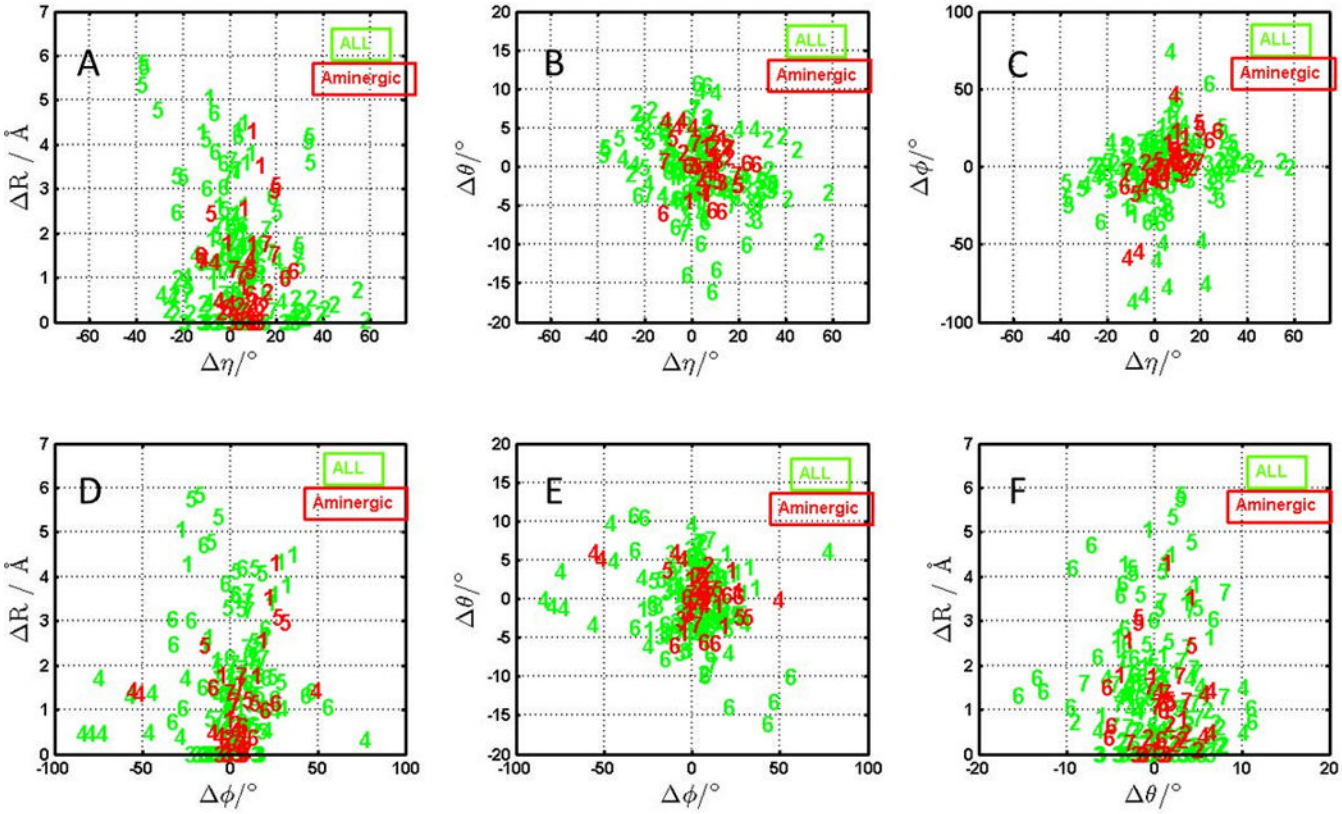

Figure 3.

Deviations in TM helix topology parameters for inactive conformations of all (green) and only aminergic (red) crystallized 7TMRs. **A.** Deviation in R vs η; **B.** Deviation in θ vs η; **C.** Deviation in ϕ vs η; **D.** Deviation of R vs ϕ; **E.** Deviation of θ vs ϕ; **F.** Deviation of R vs θ.

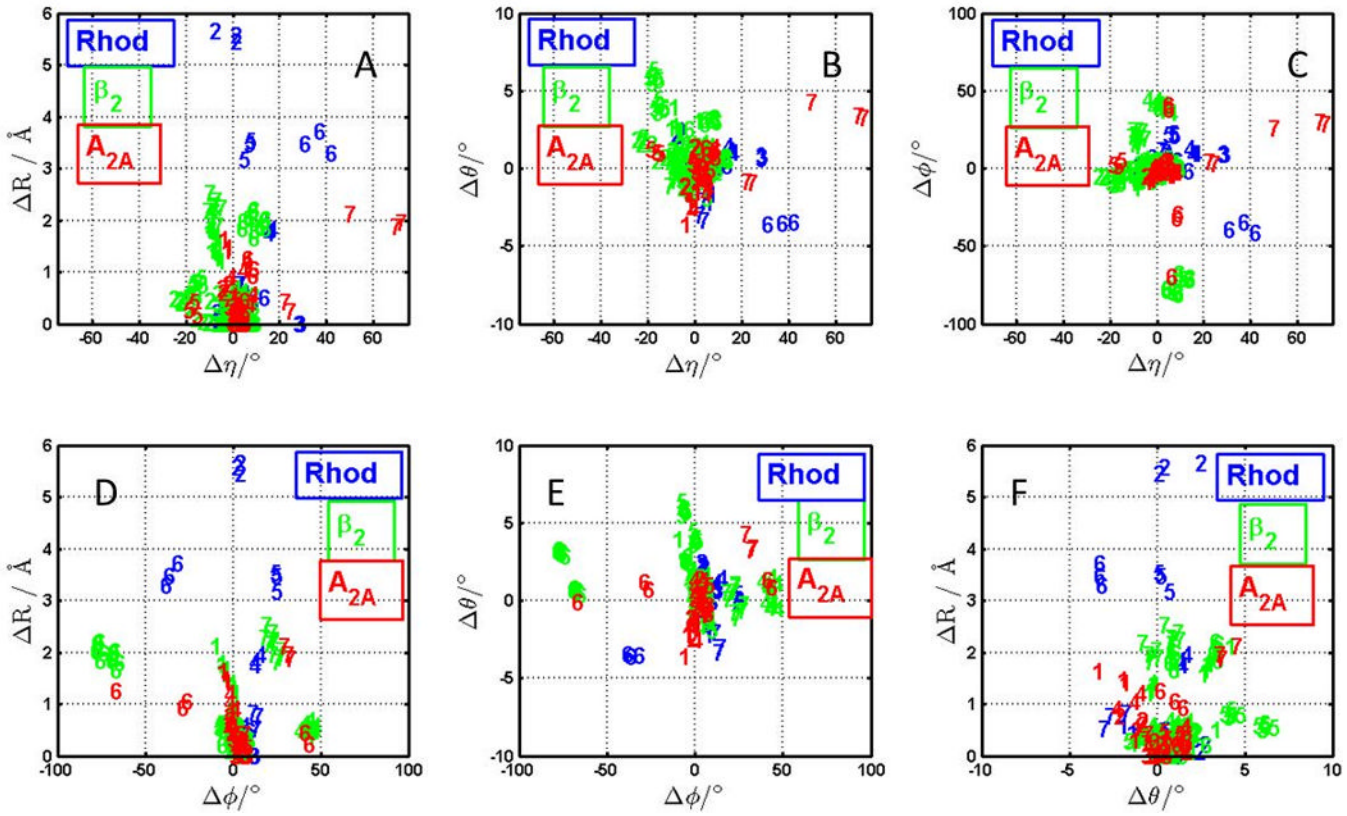

Figure 4.

Deviations in TM helix topology parameters for active conformations of rhodopsin (green), β_2 (green) and A_{2A} (red) receptors. Description of panels A through F same as in Figure 3.

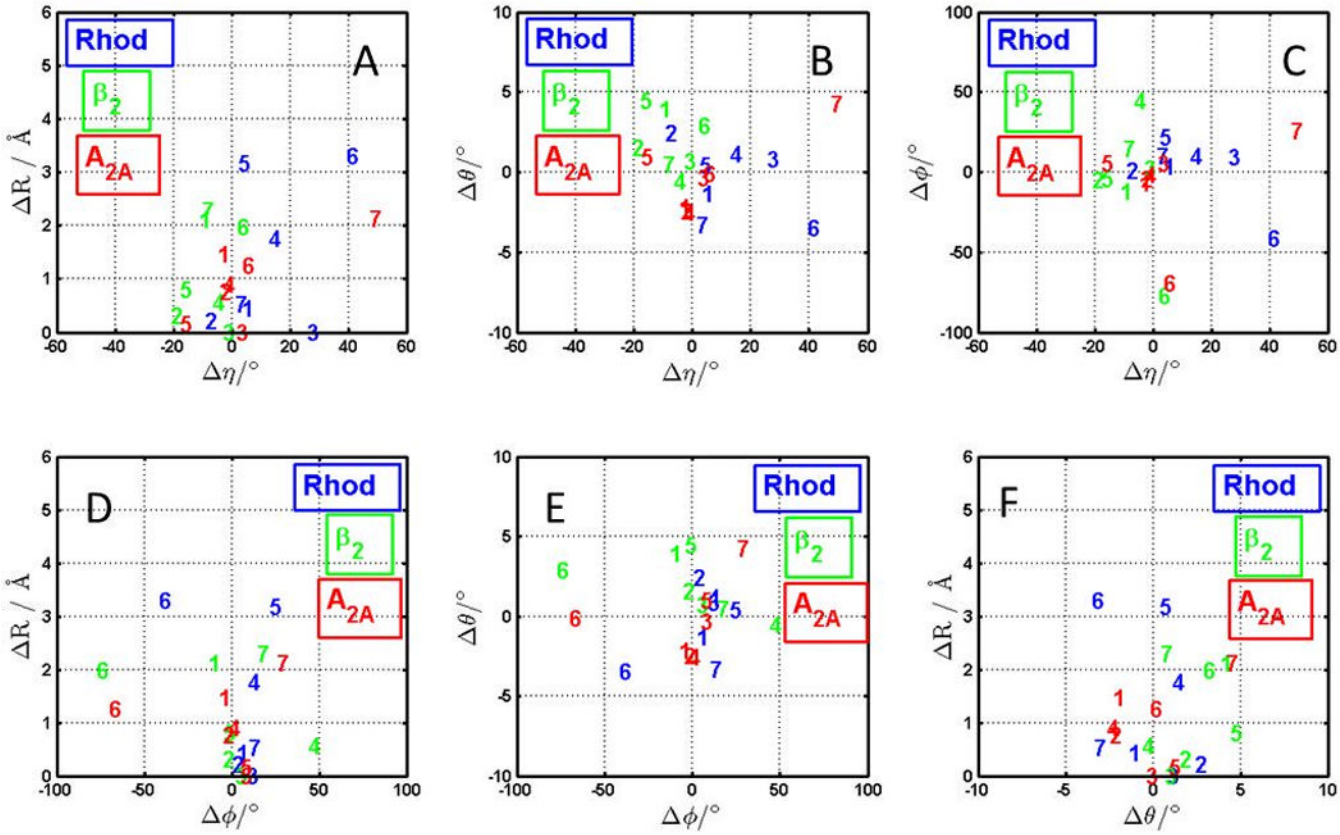

Figure 5.

Deviations in TM helix topology parameters for dark state vs metaII state of rhodopsin (green), inactive β_2 conformation vs Gs bound β_2 conformation (green), and inactive A_{2A} conformation vs agonist bound active A_{2A} conformation (red). Description of panels A through F same as in Figure 3.

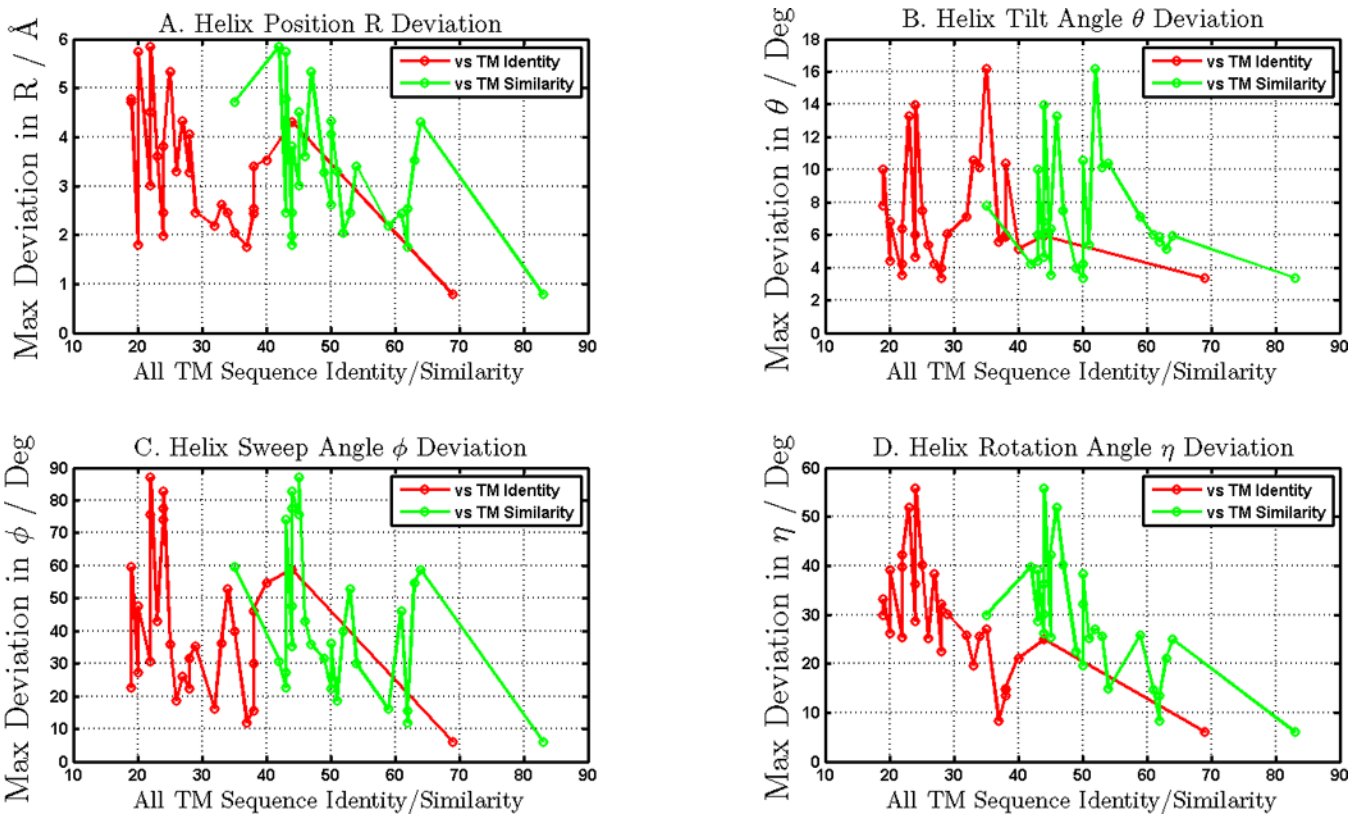

Figure 6.

Correlation of maximum deviation (over all TMs) in orientation parameters with sequence identity (red) and similarity (green): **A.** Helix position R deviation; **B.** Helix tilt angle θ deviation; **C.** Helix sweep angle ϕ deviation; **D.** Helix rotation angle η deviation.

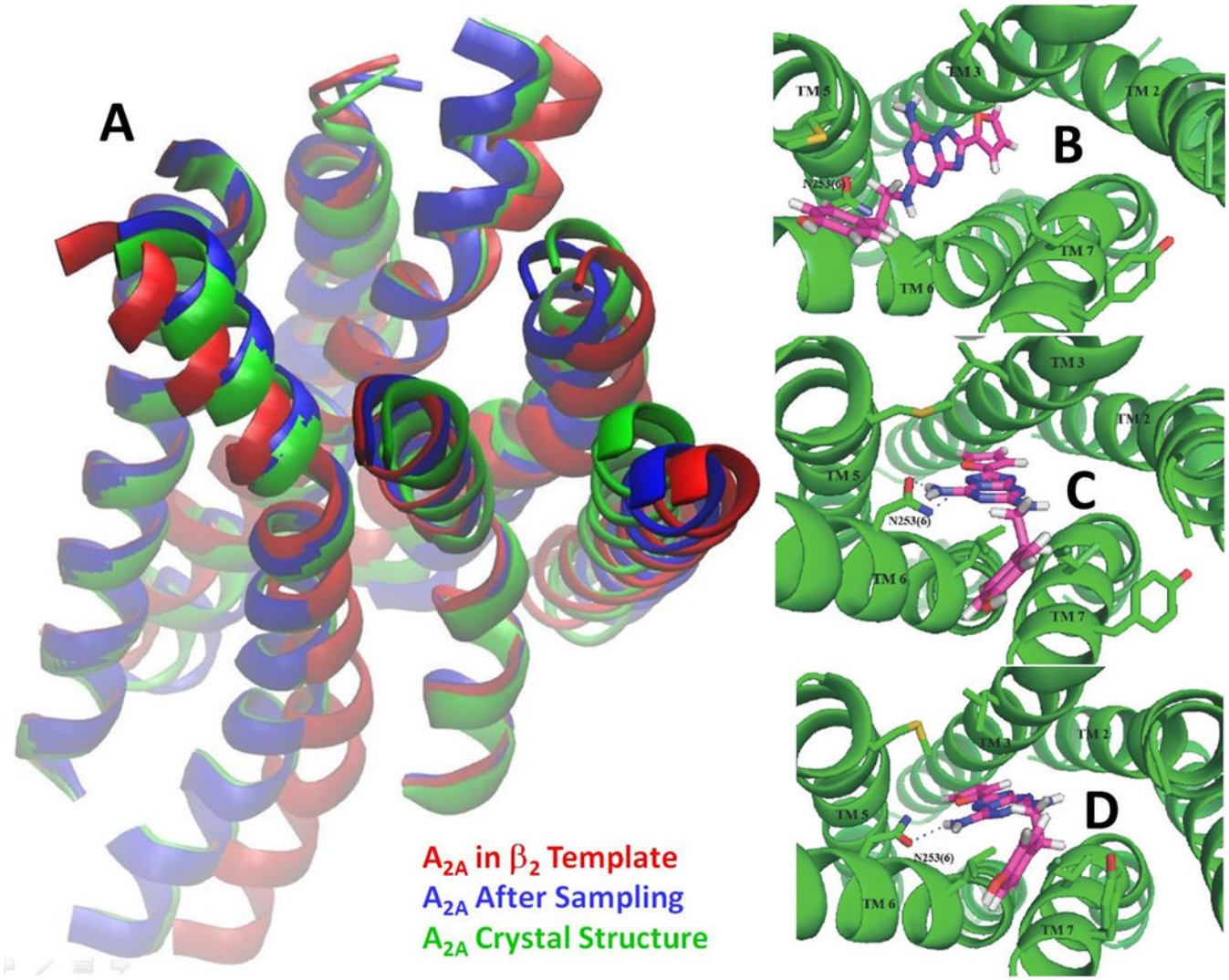

Figure 7.

A. Comparison of A2A homology structure (red) based on humBeta2 template and de novo predicted structure (blue) with the crystal structure (green); **B.** Ligand binding site for homology structure; **C.** Ligand binding site for de novo predicted structure; **D.** Ligand binding site in the co-crystal [17].

Table 1

TM Helix orientation parameters and η residue for the first human β2 adrenergic receptor structure [16].

hBeta2AR (PDBid: 2rh1) + partial inverse agonist carazolol (2.4 A)								
TM	**x**	**y**	**HPC(h)**	**θ**	**ϕ**	**η**	**ηRES**	
1	19.80	−3.98	43.61	36.23	132.58	196.34	N	51
2	9.86	0.00	81.45	28.60	340.28	275.87	D	79
3	0.00	0.00	117.52	28.80	217.12	310.30	D	113
4	−5.03	6.30	160.22	7.43	267.40	347.71	W	158
5	−8.54	−9.09	209.51	15.70	275.27	85.35	P	211
6	−0.38	−13.82	285.70	13.15	208.53	318.93	P	288
7	9.82	−10.55	316.83	15.31	57.00	68.18	P	323

Table 2

Sequence comparison across 5 7TMR sequences, with first number for the whole sequence and the second number for all TMs: **A.** Sequence Identity; **B.** Sequence Similarity.

A	**% Sequence Similarity (Full_Seq / Only_TMs)**							
	bovRhod	**humBeta2**	**turBeta1**	**humA2A**	**squRhod**	**humD3**	**humCXCR4**	**humH1**
bovRhod	100/100	14/20	14/22	14/25	21/32	16/28	15/19	10/19
humBeta2	14/20	100/100	43/69	21/33	17/28	22/40	17/24	18/37
turBeta1	14/22	43/69	100/100	22/38	16/27	21/44	15/22	17/38
humA2A	14/25	21/33	22/38	100/100	18/23	17/34	13/24	16/35
squRhod	21/32	17/28	16/27	18/23	100/100	12/20	13/22	11/26
humD3	16/28	22/40	21/44	17/34	12/20	100/100	16/29	22/38
humCXCR4	15/19	17/24	15/22	13/24	13/22	16/29	100/100	13/24
humH1	10/19	18/37	17/38	16/35	11/26	22/38	13/24	100/100
B	% Sequence Identity (Full_Seq / Only_TMs)							
	bovRhod	humBeta2	turBeta1	humA2A	squRhod	humD3	humCXCR4	humH1
bovRhod	100/100	31/43	26/42	27/47	38/59	30/49	31/35	24/43
humBeta2	31/43	100/100	55/83	33/50	33/50	35/63	31/44	31/62
turBeta1	26/42	55/83	100/100	34/54	31/50	31/64	28/45	29/62
humA2A	27/47	33/50	34/54	100/100	32/46	27/53	26/44	24/52
squRhod	38/59	33/50	31/50	32/46	100/100	25/44	27/45	21/51
humD3	30/49	35/63	31/64	27/53	25/44	100/100	27/44	36/61
humCXCR4	31/35	31/44	28/45	26/44	27/45	27/44	100/100	23/43
humH1	24/43	31/62	29/62	24/52	21/51	36/61	23/43	100/100

Table 3

Pearson and Spearman Correlation Coefficients for maximum deviation in TMHOP variables and TM Identity or TM Similarity across all pairs of 7TMRs considered in this work.

Correlation Coefficients	Pearson		Spearman	
	TM Identity	TM Similarity	TM Identity	TM Similarity
max[ΔR]	−0.53	−0.51	−0.46	−0.43
max[Δθ]	−0.11	−0.19	−0.07	−0.11
max[Δϕ]	−0.37	−0.44	−0.31	−0.37
max[Δη]	−0.69	−0.69	−0.72	−0.72