Structural Models of Class A G Protein-Coupled Receptors as a Tool for Drug Design: Insights on Transmembrane Bundle Plasticity

Xavier Deupi[†], Nicole Dölker[†], María Luz López-Rodríguez[‡], Mercedes Campillo[†], Juan A. Ballesteros[†] and Leonardo Pardo[†,*]

[†]*Laboratori de Medicina Computacional, Unitat de Bioestadística, Facultat de Medicina, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain*; [‡]*Departamento de Química Orgánica I, Facultad de Ciencias Químicas, Universidad Complutense, E-28040 Madrid, Spain*

Abstract: G protein-coupled receptors (GPCRs) interact with an extraordinary diversity of ligands by means of their extracellular domains and/or the extracellular part of the transmembrane (TM) segments. Each receptor subfamily has developed specific sequence motifs to adjust the structural characteristics of its cognate ligands to a common set of conformational rearrangements of the TM segments near the G protein binding domains during the activation process. Thus, GPCRs have fulfilled this adaptation during their evolution by customizing a preserved 7TM scaffold through *conformational plasticity*. We use this term to describe the structural differences near the binding site crevices among different receptor subfamilies, responsible for the selective recognition of diverse ligands among different receptor subfamilies. By comparing the sequence of rhodopsin at specific key regions of the TM bundle with the sequences of other GPCRs we have found that the extracellular region of TMs 2 and 3 provides a remarkable example of *conformational plasticity* within Class A GPCRs. Thus, rhodopsin-based molecular models need to include the plasticity of the binding sites among GPCR families, since the "quality" of these homology models is intimately linked with the success in the processes of rational drug-design or virtual screening of chemical databases.

1. INTRODUCTION

G protein-coupled receptors (GPCRs) are one of the largest protein families in vertebrates [1], and they are able to recognize and respond to an extraordinary chemical and structural diversity of extracellular signals, from odors and tastes to neurotransmitters, from peptides to ions, from hormones to photons [2]. In addition, they are the target of about 40% of the prescribed drugs [3, 4] and of around 25% of the top-selling drugs [5]. GPCRs, thus, constitute one of the most attractive pharmaceutical targets for designing new ligands with therapeutic benefits. Phylogenetic analyses show that human GPCRs form five main families: rhodopsin (Class A or family 1), secretin (Class B or family 2), glutamate (Class C or family 3), adhesion, and frizzled/taste2 [1]. Specialized databases of GPCRs can be found at http://www.gpcr.org/7tm [6] and http://www.iuphar-db.org.

To date, the only crystal structure available is that of the inactive state of rhodopsin [7-11]. Rhodopsin is formed by an extracellular N–terminus of four β-strands, seven transmembrane helices (TM1 to TM7) connected by alternating intracellular (I1 to I3) and extracellular (E1 to E3) hydrophilic loops, a disulfide bridge between E2 and TM3, and a cytoplasmic C–terminus containing an α-helix (HX8) parallel to the cell membrane. Statistical analysis of the residues forming the TM helices of the rhodopsin family of GPCRs shows a large number of conserved sequence patterns [12], which suggests a common transmembrane structure. In contrast, the extracellular N-terminus, cytoplasmic C-terminus and loop fragments are highly variable in length, aminoacid content, and presumably in structure among Class A receptors. Thus, the putative structural homology between rhodopsin and other GPCRs probably does not extend to the extracellular domain, which is highly structured in rhodopsin, blocking the access of the extracellular ligand to the core of the receptor [13]. In addition, the TM segments of rhodopsin, and probably of other Class A GPCRs, are far from being ideal α-helices [14]. The deformations in the helical structure are due to 'unusual' proline-induced kinks, resulting in local openings of specific helical turns, which relocate certain residues of structural and/or functional relevance. Complex networks of polar interactions, including intramolecular water molecules, stabilize these local distortions [9, 10, 15-19].

*Address correspondence to this author at the Unitat de Bioestadistica, Edifici M, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain; Tel: +34 93 581 2797; Fax: +34 93 581 2344;
E-mail: Leonardo.Pardo@uab.es

The sequence conservation in the TM domain of Class A GPCRs has been used by Ballesteros & Weinstein [20] to define a general numbering scheme, where the position of each residue is described by two numbers: the first (1 through 7) corresponds to the helix in which the residue is located; the second indicates its position relative to the most conserved residue in that helix, arbitrarily assigned to 50. These conserved residues are Asn55$^{1.50}$ (100%) (residue number in the bovine rhodopsin sequence and the general number in superscript), Asp83$^{2.50}$ (94%), Arg135$^{3.50}$ (96%), Trp161$^{4.50}$ (96%), Pro215$^{5.50}$ (77%), Pro267$^{6.50}$ (100%), and Pro303$^{7.50}$ (96%). These patterns are easily identifiable on a multiple sequence alignment and allow easy comparison among residues in the 7TM segments of different receptors. This generic numbering scheme is employed all through the manuscript.

Drug discovery has traditionally made progress by a combination of random screening and rational design [3, 21]. In practice, the latter approach has often been frustrated by the scarcity of structural experimental data that define the properties of the biological target. Nowadays, this situation is starting to change due to the significant increase of detailed 3D structural information deposited in the Protein Data Bank [22]. In the field of GPCRs, the publication of the structure of rhodopsin clearly opened a new era [23]. This structure allows the use of homology modeling techniques for building three-dimensional models of other homologous GPCRs [24, 25]. We have previously proposed that the different Class A receptor families have developed a remarkable degree of structural plasticity in order to adapt their recognition properties to the chemical diversity of the various ligands [14, 26]. In this review we are going to expand these results, describing in detail the differences between the transmembrane bundle of bovine rhodopsin and other Class A GPCRs. This plasticity of the binding sites among GPCR families is ultimately responsible for the selective affinity of a drug for a given receptor, and, thus, it is relevant for structure-based drug design, as the "quality" of the homology models is intimately linked with the success of the design approach.

2. MOLECULAR DYNAMICS SIMULATIONS AS A TOOL TO STUDY SPECIFIC STRUCTURAL FEATURES OF TRANSMEMBRANE HELICES

The effect of specific sequence motifs on the structure of TM helices can be effectively studied using molecular dynamics simulations of poly-alanine helices containing the motif under study

1568-0266/07 $50.00+.00

in an explicit hydrophobic environment [27]. We have shown that methane molecules provide an environment able to reproduce the structural characteristics of membrane-embedded proteins [28]. The ensemble of TM structures obtained in the simulations can be clustered into conformationally related subfamilies using, for instance, the program NMRCLUST [29], to select representative structures. These representative structures can be further analyzed with the program HELANAL [30]. Bend and twist angles are two relevant parameters to define the distortion of a helix. The bend angle is defined as the angle between the axes of the cylinders formed by the residues preceding and following the motif that induces the distortion in the helix, and measures large-scale structural distortions. The twist angle analyzes helical uniformity, and is related to local changes in the geometry [30]. This parameter is interpreted as follows: an ideal α-helix, with approximately 3.6 residues per turn, has a twist of approximately 100º (360º/3.6); a closed helical segment has less than 3.6 residues per turn, and a twist>100º; finally an open helical segment has more than 3.6 residues per turn, i.e. a twist<100º. This local effect is ultimately translated in a change in the orientation of the aminoacid sidechains, which can be of structural or functional importance in the context of the protein. This technique has been used to study the structural features of TM2 and TM3 in chemokine receptors [31, 32] or TM3 in amine receptors [33].

3. DIFFERENCES AND SIMILARITIES IN THE TRANSMEMBRANE BUNDLE OF GPCRS

GPCRs interact with an extraordinary diversity of ligands by means of their extracellular domains and/or the extracellular sides of the transmembrane segments. As the process of ligand recognition will largely depend on the specific receptor subtype, there is a low degree of sequence conservation among GPCR families at these extracellular domains [12]. Conversely, the processes of recognition and activation of a small number of G proteins involve the cytoplasmic ends of the transmembrane bundle and the cytoplasmic C-terminal tail of the receptor. As a result, the most conserved residues in the rhodopsin-like family of GPCRs are clustered in the central and intracellular regions of the receptor [12]. This sequence conservation pattern suggests that each receptor subfamily has developed specific sequence motifs to adjust the structural characteristics of its cognate ligands to a common set of conformational rearrangements of the helices near the G protein binding domains. We have previously suggested that GPCRs have fulfilled this adaptation during their evolution by customizing a relatively preserved scaffold through *conformational plasticity* [26]. We use this term to describe the structural differences among different receptor subfamilies within the extracellular side of the transmembrane bundle, near the binding site crevices, responsible for recognition of diverse ligands among different receptor subfamilies.

We want to describe in this section how different sequence patterns are translated into receptor plasticity. Therefore, we are going to compare the sequence at specific key regions of the transmembrane bundle of bovine rhodopsin with the sequences of other GPCR subfamilies.

3.1. Transmembrane Helix 2 is Key in Determining the Conformation of the Extracellular Part of the Receptor

The currently available crystal structures of rhodopsin [7-11] show that TM2 runs parallel to TM3 from the cytoplasmic side of the transmembrane bundle to approximately three turns below the beginning of the first extracellular loop, at position 2.57. At this point TM2 bends 33º [10] towards TM1, and leans away from TM3 (Fig. (**1**)). Two successive Gly residues at positions 2.56 and 2.57, present in 90% of the rhodopsin vertebrate type 1 receptors but absent in other rhodopsin-like GPCRs, cause this distortion of TM2. In particular, the C_α-H group of $Gly^{2.57}$ forms a hydrogen bond with the backbone carbonyl of the residue at position 3.27 in TM3 (Fig. (**2**); red line). This type of C_α-H•••O=C hydrogen bond has been show to be an important determinant of stability in membrane proteins [34]. Interestingly, in addition to the strong kink, TM2 opens at the 2.55-2.58 turn, accomodating more than 3.6 residues per turn (twist < 100º as calculated by the HELANAL program [30]). This opening of the helix induces an unusual intramolecular α–helical hydrogen bond network. Ideal α-helices are stabilized by hydrogen bonds between the carbonyl oxygen of residue at position i and the N–H amide group at position i+4, in the following turn of the helix. However, in rhodopsin, the carbonyl group at position 2.53 interacts with the N-H amide at positions 2.57 and 2.58 (i.e. both i+4 and i+5), the 2.54 carbonyl with the 2.59 N-H (i+5), and the 2.57 carbonyl with the 2.61 N-H (i+4) (Fig. (**2**); yellow lines). Thus, carbonyl groups at positions 2.55 and 2.56 lack their N-H

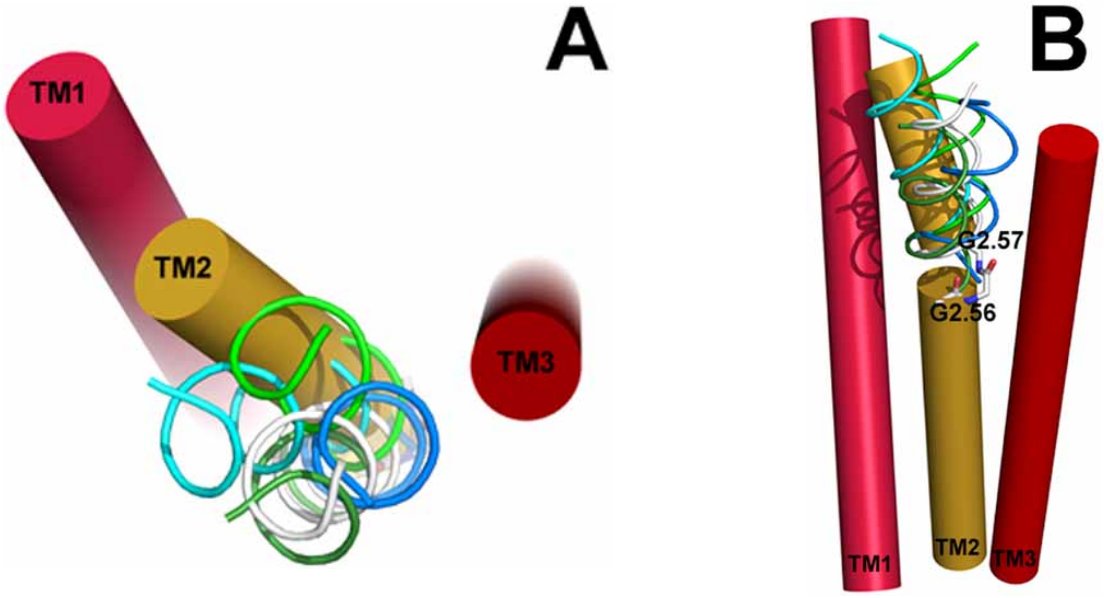

Fig. (1). Detailed view of the TM1-TM3 region, shown as cylinders, in the structure of bovine rhodopsin (PDB id: 1GZM) (Panel A, view from the extracellular side; Panel B, lateral view). Computer simulated TM helices (tube ribbon) containing polyAla (control, white), Pro at positions 2.57 (dark green), 2.59 (light green), and 2.60 (light blue), and the $Thr^{2.56}$-X-$Pro^{2.58}$ motif (dark blue), superimposed on the cytoplasmic end of TM2. Pictures were created with PyMOL [57].

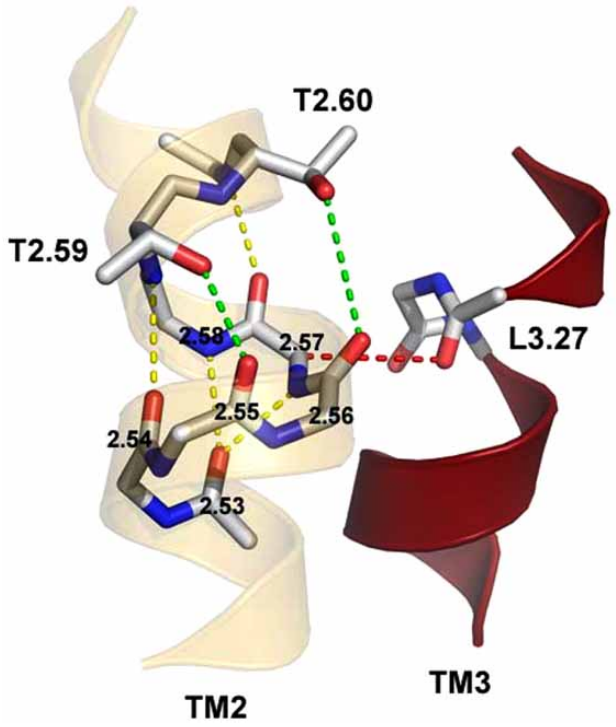

Fig. (2). Detailed view of the rhodopsin structure at the TM2-TM3 interface. The C_α-H group of $Gly^{2.57}$ forms a hydrogen bond with the backbone carbonyl of the residue at position 3.27 in TM3 (red line). TM2 contains, in this region, an unusual intramolecular α–helical hydrogen bond network: the carbonyl group at position 2.53 interacts with the N-H amide at positions 2.57 and 2.58 (i.e. both i+4 and i+5), the 2.54 carbonyl with the 2.59 N-H (i+5), and the 2.57 carbonyl with the 2.61 N-H (i+4) (yellow lines). The local opening of TM2 is stabilized by the interaction between $Thr^{2.59}$ and $Thr^{2.60}$ with the backbone carbonyls at position 2.55 and 2.56, respectively (green lines).

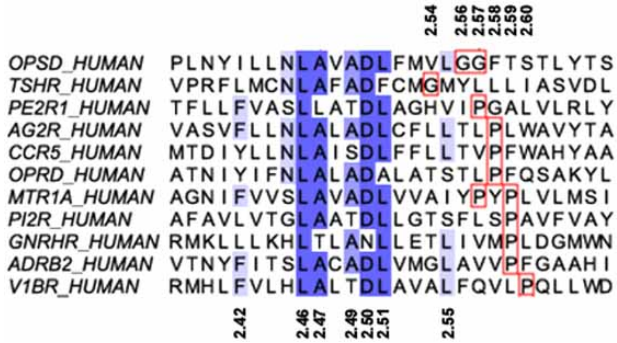

Fig. (3). Multiple sequence alignment of TM2 of various human Class A GPCRs [6]: *biogenic amine* (ADRB2_HUMAN), *peptide* (angiotensin, AG2R_HUMAN; chemokine, CCR5_HUMAN; opioid, OPRD_HUMAN; vasopresin, V1BR_HUMAN), *glycoprotein hormone* (TSHR_HUMAN), *rhodopsin* (OPSD_HUMAN), *prostanoid* (prostaglandin, PE2R1_HUMAN; prostacyclin, PI2R_HUMAN), *gonadotropin-releasing hormone* (GNRHR_HUMAN) and *melatonin* (MTR1A_HUMAN) receptors. The highly conserved residues in the cytoplasmic region of TM2 are shadowed in blue, with their position in the sequence shown at the bottom of the alignment. Clearly, the extracellular region is less conserved, and features different family-specific motifs (boxed in red) with their positions shown at the top of the alignment: the $Gly^{2.56}Gly^{2.57}$ motif of the rhodopsin family; $Pro^{2.57}$ in prostaglandin receptors; $Pro^{2.58}$ (forming part of a Thr-X-Pro motif) in peptide receptors; $Pro^{2.59}$ in prostanoid, gonadotropin-releasing hormone and amine receptors; $Pro^{2.60}$ in vasopressin-like receptors; and the $Pro^{2.57}$-X-$Pro^{2.59}$ motif in melatonin receptors. The putative structural effects of these sequence motifs are shown in Fig. (**1**). The figure was created with Jalview [58].

counterpart. Rhodopsin vertebrate type 1 receptors also contain two highly conserved (83%) consecutive Thr residues at positions 2.59 and 2.60. The short side chain of Thr is capable to hydrogen bond the backbone carbonyl in the previous turn of the helix, inducing and/or stabilizing distortions in TMs [35]. Specifically, $Thr^{2.59}$ and $Thr^{2.60}$ interact with the backbone carbonyls at position 2.55 and 2.56, respectively (Fig. (**2**); green lines). The interactions of both Thr residues with these carbonyls stabilize this extreme conformation of TM2 in rhodopsin. In summary, rhodopsin contains a GGxTT motif that distorts TM2 and induces a family-specific TM2-TM3 interface (see below). The absence of this motif in other rhodopsin-like GPCRs reveals the extracellular region of TM2 and TM3 as an example of local structural variability in the rhodopsin family (*conformational plasticity*).

Fig. (**3**) shows a multiple sequence alignment of TM2 in various Class A GPCRs, including rhodopsin. The primary structure of TM2 is highly conserved in its cytoplasmic side at positions 2.40 (N:40%; D:10% of the sequences), 2.42 (F:39%; Y:28%), 2.45 (N:51%; S:29%), 2.46 (L:91%), 2.47 (A:74%), 2.49 (A:58%), 2.50 (D:94%), 2.51 (L:60%), and 2.52 (L:60%) [12]. This conservation pattern suggests a common structural and functional role of the intracellular domain of TM2 in the rhodopsin-family of GPCRs. In contrast, the aminoacid sequence is strongly divergent at the extracellular side. For instance, there are several Pro residues at the 2.57 (2%), 2.58 (41%), 2.59 (37%), and 2.60 (4%) positions [12] capable of introducing structural changes. Pro residues are normally observed in TM helices [36] where they induce a significant distortion named Pro-kink [37]. The steric clash between the pyrrolidine ring of Pro and the carbonyl oxygen of the residue in the preceding turn [38] induces a bend angle of approximately 20º in the helical structure [27]. Specifically, prostaglandin E2 subtype EP1 receptors contain Pro at position 2.57 (see Table **1**). To determine the consequence that $Pro^{2.57}$ might have on the structure of these receptors we superimposed a computer simulated TM helix [27, 28] containing Pro at position 2.57 to the highly conserved intracellular domain of TM2 (Fig. (**1**)). It can bee seen how the kink induced by $Pro^{2.57}$ (dark green) orients the extracellular moiety of TM2 away from the helical bundle. On the other hand, 41% of Class A GPCRs have Pro at position 2.58. Statistical analysis of the aminoacids present in the vicinity of $Pro^{2.58}$ shows that almost in all receptors there is a Ser or Thr side chain in the i-3/i+3 range (see Table **2** in [27]). In particular the $Thr^{2.56}$-X-$Pro^{2.58}$ (X being a non-conserved residue) sequence motif is the most abundant, present in 71% of the $Pro^{2.58}$-containing receptors. For instance, it is highly conserved in angiotensin, chemokine, and opioid receptors (Fig. (**3**)). The sole presence of $Pro^{2.58}$ would orient the extracellular end of TM2 towards TM3 and away from TM1 (results not shown). However, the presence of Thr in the $Thr^{2.56}$-X-$Pro^{2.58}$ motif increases the helical bend angle by about 7-10° [27], causing the extracellular side of TM2 to lean even more toward TM3 and slightly toward the center of the bundle [31] (dark blue helix in Fig. (**1**)). In contrast, 37% of the receptors (biogenic amine, olfactory, prostanoid and gonadotropin-releasing hormone receptors, see Table **1** and Fig. (**3**)) contain a Pro residue at position 2.59, which induces a bend helix in the direction of TM1 (light green helix in Fig. (**1**)). Finally, $Pro^{2.60}$-containing receptors (as beta adrenergic types 3 and 4, vasopressin-like and thyrotropin-releasing hormone receptors, see Table **1** and Fig. (**3**)) would bend TM2 away from the helical bundle (light blue helix in Fig. (**1**)). It is worth to note that melatonin receptors possess both $Pro^{2.57}$ and $Pro^{2.59}$ (Table **1** and

Table 1. Sequence Motifs in TM2 of Class A GPCR Subfamilies (as in GPCRdb [6]), with a Subfamily Conservation Greater than ~50%

	$Gly^{2.54}$	$Gly^{2.56}/Gly^{2.57}$	$Pro^{2.57}$	$Pro^{2.58}$	$Pro^{2.59}$	$Pro^{2.60}$
Amine	44[(1)]	0	0	0	**88**	6[(5)]
Peptide	1	0	0	63	21	5[(6)]
Hormone protein	**100**	0	0	0	0	0
(Rhod)opsin	0	30[(3)]	0	7	15	11
Olfactory	2	0	0	0	**98**	1
Prostanoid	0	0	11[(4)]	0	62	0
Nucleotide-like	45[(2)]	0	0	55	47	0
Cannabis	0	0	0	0	0	0
Platelet activating factor	0	0	0	**100**	0	0
Gonadotropin-releasing hormone	0	0	0	0	**100**	0
Thyrotropin-releasing hormone & Secretagogue	0	0	0	33	0	**67**[(7)]
Melatonin	0	0	**100**	0	**100**	0
Viral	0	0	0	**81**	0	0
Lysosphingolipid & LPA (EDG)	**97**	0	0	0	0	0
Leukotriene B4 receptor	0	0	0	**100**	0	0
Class A Orphan/other	18	0	0	43	20	2

[(1)] Acetylcholine receptors lack $Pro^{2.59}$ but feature $Gly^{2.54}$. Histamine receptors have these two residues simultaneously.
[(2)] 100% conserved in adenosine receptors, in addition to $Pro^{2.59}$.
[(3)] 100% conserved in vertebrate type 1 opsins. The rest of vertebrate opsin family has either Gly or Ser residues at position 2.56 or 2.57.
[(4)] 100% conserved in prostaglandin E2 subtype EP1 receptors.
[(5)] 100% conserved in beta adrenoreceptors types 3 and 4.
[(6)] 100% conserved in vasopressin-like receptors.
[(7)] 100% conserved in thyrotropin-releasing hormone receptors.

Fig. (**3**)); thus, TM2 in this subfamily will probably adopt a different conformation that those proposed in Fig. (**1**).

In conclusion, the extracellular part of TM2 provides a remarkable example of *conformational plasticity* within Class A GPCRs, where each family has developed a particular sequence motif to adapt this region to its structural and functional requirements. Fig. (**1**) exemplifies this effect by showing representative structures of helices without Pro aminoacids (white), Pro at positions 2.57 (dark green), 2.59 (light green), and 2.60 (light blue), and the $Thr^{2.56}$-X-$Pro^{2.58}$ motif (dark blue), superimposed on the cytoplasmic end of TM2 of the rhodopsin structure. Clearly, the different aminoacid sequence in the extracellular side of TM2 in rhodopsin and other homologous GPCRs is translated into structural divergences in this region of TM2 and, consequently, the nearby TM3 (see below).

3.2. The Interface Between Transmembrane Helices 2 and 3

TM3 mediates a helix-helix interaction with TM2 through the backbone carbonyl of the residue at position 3.27 and the C_α-H group of $Gly^{2.57}$ (Fig. (**2**)), specific of the opsin family. Most of the other members of the rhodopsin family of GPCRs possess at this 2.57 position a bulky β-branched or γ-branched aminoacid (Leu, 39%; Val, 14%; Ile, 3%) [12] that prevents the interaction with the 3.27 carbonyl. Therefore, it seems reasonable to suggest an alternative TM2-TM3 interface for GPCRs containing this bulky side chain at position 2.57, or containing Pro residues in TM2 (see above). We propose that TM2 in other GPCRs possesses a different conformation compared to rhodopsin, resulting in the subsequent relocation of TM3 (Fig. (**4**)). Fig. (**5**) shows the content in Pro, Gly, Ser, or Thr residues at the 3.22-3.52 positions of TM3 in Class-A GPCRs. The virtual absence of Pro residues in this helix suggests that the conformational plasticity might be dominated by the action of other aminoacids. Notably, Gly and Ser or Thr residues are

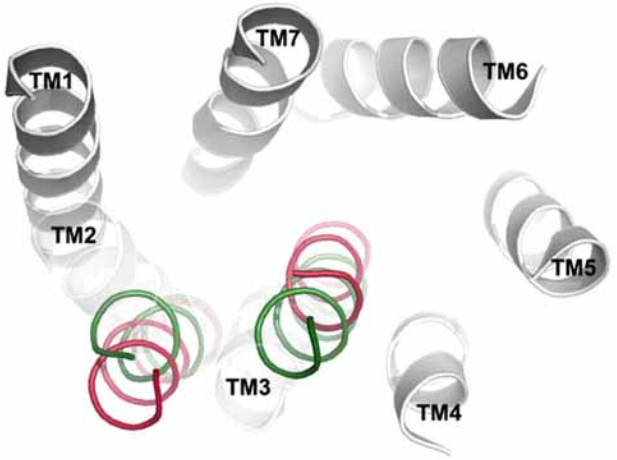

Fig. (4). Structural plasticity of TM2 and TM3 in Class A GPCRs. The TM helices of rhodopsin are shown in white. Alternative geometries for TM2 and TM3 in amine receptors (green), due to the presence of $Pro^{2.59}$ in TM2 and Ser/Thr residues in TM3 [26]; and in peptide receptors (red), due to the presence of the $Thr^{2.56}$-X-$Pro^{2.58}$ motif in TM2 and the subsequent relocation of TM3 [32].

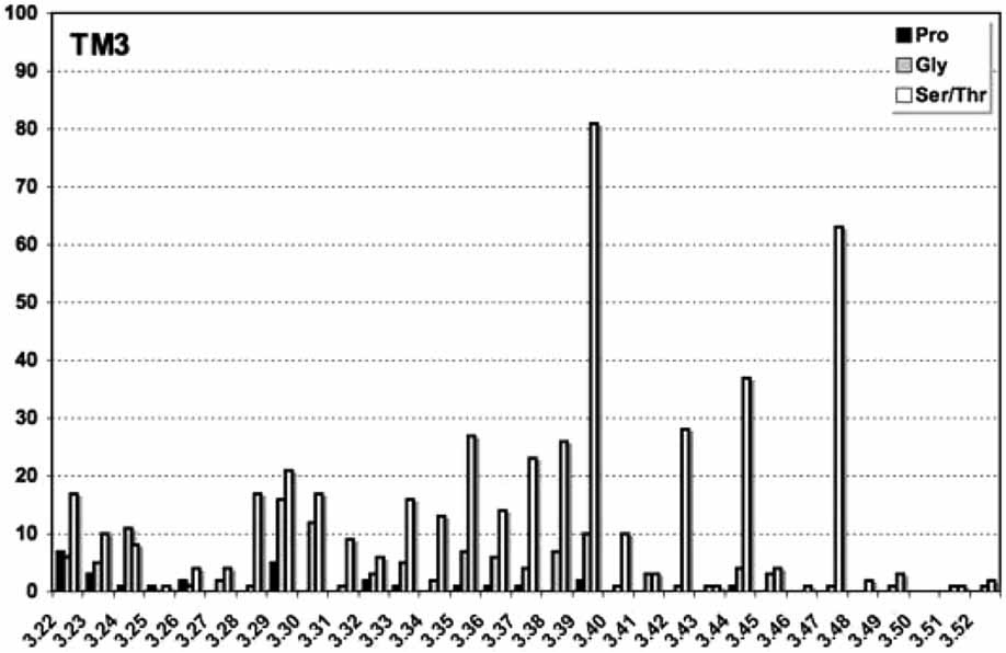

Fig. (5). Content (in %) of Pro (black), Gly (gray), or Ser/Thr (white) residues in TM3 of Class A GPCRs [12].

relatively prevalent at different parts of TM3. Ser and Thr residues have been shown to induce or stabilize distortions in transmembrane segments due to the additional hydrogen bond formed between the Oγ-H moiety of the side chain and the peptide carbonyl oxygen of the preceding helical turn [35]. Moreover, the additional flexibility in the helix due to the presence of Gly residues (because of the lack of the side chain) might reinforce this effect. The fact that Ser/Thr and Gly are spread over TM3 rather than strongly conserved at a certain positions suggests structural differences among different Class A GPCR families in TM3.

Within this family, TM3 is structurally conserved at the cytoplasmic end of the helix. In this region, TM3 contains the highly conserved $(D/E)^{3.49}R^{3.50}(Y/W)$ motif [39], bulky hydrophobic residues at positions 3.43 (L:74%; I:10%; V:6%) and 3.46 (I:58%; L:15%; M:15%; V:8%), and a hydrophilic residue at position 3.39 (S:75%; T:6%) [12]. Thus, as for TM2, these structural and possibly functional similarities suggest that, while the cytoplasmic side is conserved, the extracellular side will present family-dependent conformations. This concept is of special relevance, since TM3 is located at the center of the transmembrane bundle and forms, together with TMs 2, 5, 6, and 7, the binding-site crevice for extracellular diffusible ligands [40].

More than 90% of GPCRs contain a disulfide bridge between TM3 and the E2 loop connecting TM4 and TM5. The high variability of the E2 loop with respect to length (from 4 to more than 50 residues [12]) and aminoacid composition suggests a non-conserved structure of the loop, which would result in a different interface at the extracellular side of TM3-TM5 in different GPCR families. It has been suggested for the dopamine D2 receptor that TM3 is bent towards TM5 to account for the spatial constraint imposed by the short loop of four residues between the Cys engaged in the disulfide bridge with TM3 and TM5 [41]. In addition, we have suggested [33] that $Thr^{3.37}$, 85% conserved in the neurotransmitter family of GPCRs, is responsible for the relocation of TM3 towards TM5 to facilitate the known interaction of aminergic ligands with $Asp^{3.32}$ in TM3 and a series of Ser/Thr residues at positions 5.42 and 5.43 in TM5 [16] (Fig. (**4**)). Similarly, we have shown for chemokine receptors that the bending of TM2 toward TM3, in its outer half, due to the $Thr^{2.56}$-X-$Pro^{2.58}$ motif, is tolerated in the context of the CCR5 helical bundle as the result of the relocation of TM3 toward TM5 [32] (Fig. (**4**)).

3.3. Local Opening of Transmembrane Helix 5

Despite the presence of the highly conserved $Pro^{5.50}$ (77% conserved in Class A GPCRs), TM5 has a bend angle of only 13º [10], significantly lower than the average Pro-kink induced angle of 20º [27]. In order to remove the steric clash between the pyrrolidine ring of $Pro^{5.50}$ and the carbonyl oxygen at position 5.46, TM5 is opened at the 5.45-5.48 turn (>3.6 residues/turn, twist < 100º). The 5.46 carbonyl oxygen adopts an unusual conformation that is stabilized by the interactions with $Glu^{3.37}$ (present in only 4% of the sequences) and $Leu^{3.40}$ (L, 9%; V, 25%; I, 42%) (Fig. (**6A**)). Fig. (**6B**) compares TM5 of rhodopsin (red) with a helix featuring a standard Pro-kink (gray). Importantly, the opening of the helix at the 5.45-5.48 turn modifies the orientation of the aminoacid side chains at the extracellular domain, which point towards a different position. This distortion is of key structural and functional importance. For instance, the residue at position 5.42 (part of the binding-site in the N-formyl, C3a, and C5a peptide receptors [42, 43] or amine receptors [44], among others [40] would be incorrectly oriented towards the lipid environment if TM5 was modeled as a regular Pro-kinked helix (Fig. (**6B**), gray). In contrast, the opening of the helix properly positions this side chain towards the binding site crevice (Fig. (**6B**), red). These data suggest that TM5 shares a common conformation in $Pro^{5.50}$-containing GPCRs. The question arises whether TM5 of other GPCRs lacking this Pro, for instance glycoprotein hormone and cannabinoid receptors, have evolved an alternative structural mechanism to adopt a conformation similar to rhodopsin.

3.4. The Interface Between Transmembrane Helices 6 and 7

TM6 of rhodopsin presents the most pronounced kink in the TM bundle (~35º [10]) and, moreover, it is opened at the 6.46-6.49 turn (>3.6 residues/turn, twist < 100º) (Fig. (**7**)). This severe distortion is energetically stabilized through two structural and functional elements involved in GPCR activation. First, $Pro^{6.50}$ of the highly conserved CWxPx (Y/F) motif introduces a flexible point in TM6 facilitating this extreme conformation. The Pro-kink

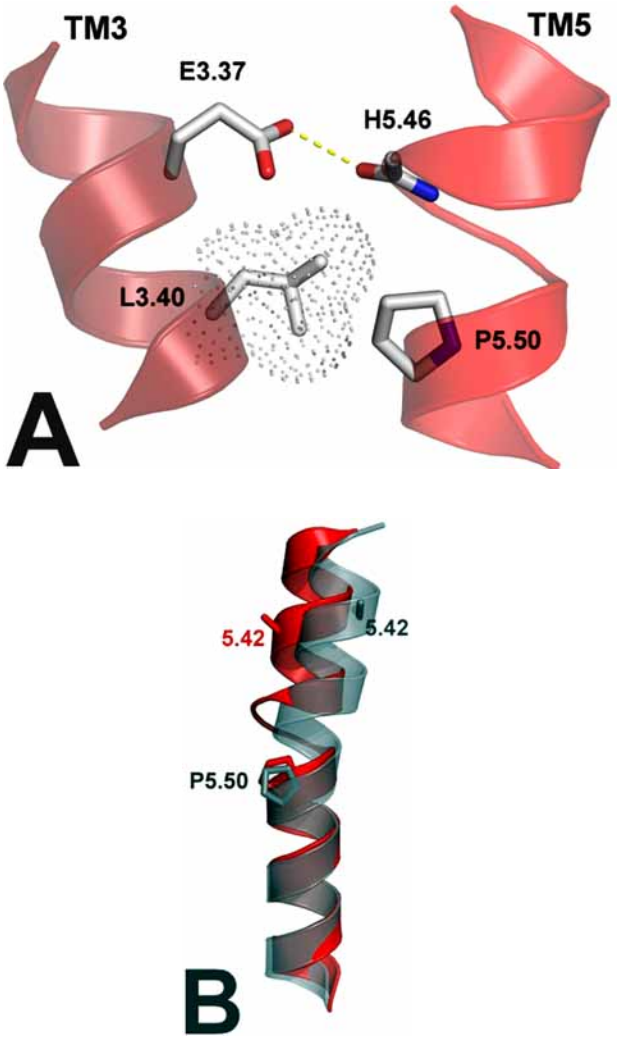

Fig. (6). A. Detailed view of the interface between TM3 and TM5 in rhodopsin. The opening of TM5 at the 5.45-5.48 turn is depicted as a coil. The unusual conformation of the backbone carbonyl at position 5.46 is stabilized by an interaction with $Glu^{3.37}$ (yellow line) and a steric hindrance with $Leu^{3.40}$ (white dots). B. Comparison between the structures of TM5 in rhodopsin (red) and a TM helix featuring a standard Pro-kink (gray). The helices are superimposed at the intracellular part. The opening at the 5.45-5.48 turn (shown as a coil) properly positions the side chain at position 5.42, part of the binding site crevice in several families, towards the TM bundle.

distortion of TM6 disrupts the intramolecular hydrogen bond between the carbonyl group at position 6.47 and the N-H amide at positions 6.51 (Fig. (**7**)).

The structure of rhodopsin [10] reveals a water molecule located between TM6 and TM7 (Wat#1), which acts as a counterpart of these groups, linking the backbone carbonyl at position 6.47 with the backbone N–H amide at position 6.51. In addition, this water molecule interacts with the backbone carbonyl at position 7.38 (Fig. (**7**)). Importantly, 60% of Class A GPCRs contain a non-bulky aminoacid at position 7.42 (A:40%; G:20%), creating a small cavity between TM6 and TM7 that allows the accommodation this water molecule. An additional 24% of GPCRs (as cannabinoids or acetylcholine receptors) contain a small and polar side chain (S:13%; T:7%; C:4%) at this position. We hypothesize that this polar side chain further stabilizes the carbonyl oxygen at position 6.47 through an extra hydrogen bond interaction.

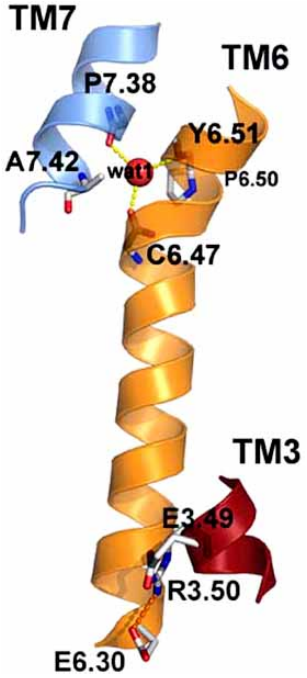

Fig. (7). Detailed view of TM3 (dark red), TM6 (orange), and TM 7 (blue) in rhodopsin. The strong distortion at $Pro^{6.50}$ is stabilized by a water molecule (wat1), which links the backbone carbonyl at position 6.47 with the backbone N–H amide at position 6.51. In addition, this water molecule links TM6 and TM7 through the interaction with the carbonyl at position 7.38. The small side chain of $Ala^{7.42}$ creates a small cavity between TM6 and TM7 that allows the accommodation this water molecule. The distortion of TM6 is further stabilized by the ionic interaction between $Glu^{6.30}$ in TM6 and the $Glu^{3.49}$-$Arg^{3.50}$ pair in TM3.

Therefore, we suggest that this complex, highly packed, and probably functional interface between TM6 and TM7 is conserved in most Class A GPCRs. However, a bulky β-branched or γ-branched aminoacid (Leu, 4%; Val, 1%; Ile, 3%) [12] is present in, for instance, in the melanocortin and orexin families of GPCRs, suggesting a structural difference in this region.

The second element that stabilizes the strong distortion of TM6 is the ionic interaction ("ionic lock") between $Asp/Glu^{6.30}$ in TM6 and the $Asp/Glu^{3.49}$-$Arg^{3.50}$ pair of the (D/E)R(Y/W) motif in TM3, which links the cytoplasmic ends of these two helices in the inactive state of several GPCR families [15, 45, 46] (Fig. (**7**)). This ionic lock provides the energy required to maintain the extreme conformation of TM6 (see below). Disruption of the ionic lock [15, 47, 48], aided by the protonation of $Asp/Glu^{3.49}$ [49], induces conformational changes at the cytoplasmic side of TM3 and TM6 [50], considered to be an essential step in the process of GPCR activation.

3.5. The Interface Between Transmembrane Helices 3 and 6

While the (D/E)R(Y/W) motif in TM3 is highly conserved in Class A GPCRs, the acidic residue at position 6.30 is only present in 32% of the sequences (D, 7%; E, 25%). Other receptors may have evolved a comparable mechanism using different residues. For instance, opioid receptors feature a Leu in 6.30, so the role of $Glu^{6.30}$ is likely to be played by $Thr^{6.34}$ through a similar although specialized set of intramolecular interactions with $Arg^{3.50}$ [51]. Notably, many GPCRs contain a basic residue at this 6.30 position (34%; K, 18%; R, 16%), which precludes a direct interaction with $Arg^{3.50}$. These receptors will probably possess a totally different network of interhelical interactions at the intracellular side that remains to be identified.

4. CONCLUSIONS

Recent studies in GPCR pharmacology have lead to the development of new concepts in the field of drug discovery [52]. Specifically, collateral efficacy describes the ability of some ligands to selectively induce some of the receptor functions (activation of G proteins, interactions with other proteins, dimerization, phosphorylation, desensitization, or internalization) but not others. In addition, permissive antagonism is the property of an antagonist to block some but not all receptor-mediated signals. These ideas can be linked to GPCR structure through the concept of receptor plasticity. A number of studies have provided convincing evidence that GPCRs co-exist in different conformations, strongly suggesting an inherent flexibility [53-56]. Specific sequence motifs provide this degree of flexibility to TM helices, resulting in a range of receptor conformations [26]. Thus, ligands can selectively recognize particular conformations of the TM bundle triggering specific properties.

Rhodopsin crystal structures have shown that the helices comprising the TM bundle are far from being ideal α-helices. TM helices containing Pro are also distant from standard Pro-kinked helices. These distortions are energetically stabilized through complementary intra- and inter-helical interactions involving polar side chains, backbone carbonyls and, in some cases, specific structural and functional water molecules embedded in the TM bundle.

Clearly, homology models of GPCRs developed for rational drug-design or for virtual screening of chemical databases, need to include the *conformational plasticity* of the receptor. On one hand, some of the structural peculiarities of rhodopsin will be present in the receptor under study, either because the specific sequence motif is conserved in the receptor being modeled, or because its family has evolved a compatible mechanism to stabilize a similar structure. Otherwise the putative family-specific structural dissimilarities relative to rhodopsin, must be explicitly modeled. Without any doubt, inclusion of experimental results improves the reliability of the computer models. These tailor-made models are conceptually opposed to automated model building techniques, which do not take into account family-specific structural peculiarities. Thus, the available crystal structures of rhodopsin are a double-edged weapon: while they provide a realistic template for homology model building of Class A GPCRs, they can also lead to an over-simplification of GPCR structure. Clearly, one single template cannot explain the rich diversity of GPCR function.

ACKNOWLEDGMENTS

This work was supported by the European Community (LSHB-CT-2003-503337), MEC (SAF2004-07103, SAF 2006-04966), and AGAUR (SGR2005-00390). Xavier Deupi is supported by a Juan de la Cierva contract from the Ministerio de Educacion y Ciencia (Spain).

REFERENCES

[1] Fredriksson, R.; Lagerstrom, M. C.; Lundin, L. G.; Schioth, H. B. The g-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.*, **2003**, *63* (6), 1256-72.

[2] Kristiansen, K. Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function. *Pharmacol. Ther.*, **2004**, *103* (1), 21-80.

[3] Drews, J. Drug discovery: a historical perspective. *Science*, **2000**, *287* (5460), 1960-4.

[4] Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.*, **2002**, *1* (9), 727-30.

[5] Klabunde, T.; Hessler, G. Drug design strategies for targeting G-protein-coupled receptors. *Chembiochem*, **2002**, *3* (10), 928-44.

[6] Horn, F.; Bettler, E.; Oliveira, L.; Campagne, F.; Cohen, F. E.; Vriend, G. GPCRDB information system for G protein-coupled receptors. *Nucleic Acids Res.*, **2003**, *31* (1), 294-7.

[7] Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, **2000**, *289* (5480), 739-45.

[8] 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* (26), 7761-72.

[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* (9), 5982-7.

[10] 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* (5), 1409-38.

[11] 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 A crystal structure. *J. Mol. Biol.*, **2004**, *342* (2), 571-83.

[12] Mirzadegan, T.; Benko, G.; Filipek, S.; Palczewski, K. Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry*, **2003**, *42* (10), 2759-67.

[13] Bourne, H. R.; Meng, E. C. Rhodopsin sees the light. *Science*, **2000**, *289*, 733-734.

[14] Ballesteros, J. A.; Shi, L.; Javitch, J. A. Structural mimicry in G protein-coupled receptors: implications of the high-resolution structure of rhodopsin for structure-function analysis of rhodopsin-like receptors. *Mol. Pharmacol.*, **2001**, *60* (1), 1-19.

[15] Ballesteros, J. A.; Jensen, A. D.; Liapakis, G.; Rasmussen, S. G.; Shi, L.; Gether, U.; Javitch, J. A. Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem.*, **2001**, *276* (31), 29171-7.

[16] Shi, L.; Javitch, J. A. The binding site of aminergic G protein-coupled receptors: the transmembrane segments and second extracellular loop. *Ann. Rev. Pharmacol. Toxicol.*, **2002**, *42*, 437-67.

[17] Jongejan, A.; Bruysters, M.; Ballesteros, J. A.; Haaksma, E.; Bakker, R. A.; Pardo, L.; Leurs, R. Linking ligand binding to histamine H_1 receptor activation. *Nat Chem. Biol.*, **2005**, *1*, 98-103.

[18] Urizar, E.; Claeysen, S.; Deupi, X.; Govaerts, C.; Costagliola, S.; Vassart, G.; Pardo, L. An activation switch in the rhodopsin family of G protein coupled receptors: The thyrotropin receptor. *J. Biol. Chem.*, **2005**, *280* (17), 17135-41.

[19] Xu, W.; Campillo, M.; Pardo, L.; de Riel, J. K.; Liu-Chen, L.-Y. The seventh transmembrane domains of the delta and kappa opioid receptors have different accessibility patterns and interhelical interactions. *Biochemistry*, **2005**, *44* (49), 16014-25.

[20] 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. *Methods Neurosci.*, **1995**, *25*, 366-428.

[21] Waszkowycz, B. Structure-based approaches to drug design and virtual screening. *Curr. Opin. Drug Discov. Devel.*, **2002**, *5* (3), 407-13.

[22] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.*, **2000**, *28* (1), 235-42.

[23] Kirkpatrick, P. Structure-based drug design: Can GPCR models work? *Nat. Rev. Drug Discov.*, **2003**, *2*, 171-171.

[24] Bissantz, C. Conformational changes of G protein-coupled receptors during their activation by agonist binding. *J. Recept. Signal. Transduct. Res.*, **2003**, *23* (2-3), 123-53.

[25] Fanelli, F.; De Benedetti, P. G. Computational modeling approaches to structure-function analysis of g protein-coupled receptors. *Chem. Rev.*, **2005**, *105* (9), 3297-351.

[26] Deupi, X.; Govaerts, C.; Shi, L.; Javitch, J. A.; Pardo, L.; Ballesteros, J. Conformational Plasticity of GPCR Binding sites. In *The G Protein-Coupled Receptors Handbook*. Devi L.A.; Ed; Humana Press: Totowa, NJ, **2005**; pp. 363-388.

[27] 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*, 105-15.

[28] Olivella, M.; Deupi, X.; Govaerts, C.; Pardo, L. Influence of the environment in the conformation of alpha-helices studied by protein database search and molecular dynamics simulations. *Biophys. J.*, **2002**, *82* (6), 3207-13.

[29] Kelley, L. A.; Gardner, S. P.; Sutcliffe, M. J. An Automated Approach For Clustering An Ensemble Of NMR-Derived Protein Structures Into Conformationally-Related Subfamilies. *Protein Eng.*, **1996**, *9*, 1063-1065.

[30] Bansal, M.; Kumar, S.; Velavan, R. HELANAL: a program to characterize helix geometry in proteins. *J. Biomol. Struct. Dyn.*, **2000**, *17* (5), 811-9.

[31] Govaerts, C.; Blanpain, C.; Deupi, X.; Ballet, S.; Ballesteros, J. A.; Wodak, S. J.; Vassart, G.; Pardo, L.; Parmentier, M. The TXP motif in the second transmembrane helix of CCR5. A structural determinant of chemokine-induced activation. *J. Biol. Chem.*, **2001**, *276* (16), 13217-25.

[32] Govaerts, C.; Bondue, A.; Springael, J. Y.; Olivella, M.; Deupi, X.; Le Poul, E.; Wodak, S. J.; Parmentier, M.; Pardo, L.; Blanpain, C. Activation of CCR5 by chemokines involves an aromatic cluster between transmembrane helices 2 and 3. *J. Biol. Chem.*, **2003**, *278* (3), 1892-903.

[33] Lopez-Rodriguez, M. L.; Vicente, B.; Deupi, X.; Barrondo, S.; Olivella, M.; Morcillo, M. J.; Behamu, B.; Ballesteros, J. A.; Salles, J.; Pardo, L. Design, synthesis and pharmacological evaluation of 5-hydroxytryptamine(1a) receptor ligands to explore the three-dimensional structure of the receptor. *Mol. Pharmacol.*, **2002**, *62* (1), 15-21.

[34] Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M. The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. USA*, **2001**, *98* (16), 9056-61.

[35] Ballesteros, J. A.; Deupi, X.; Olivella, M.; Haaksma, E. E.; Pardo, L. Serine and threonine residues bend alpha-helices in the chi(1) = g(-) conformation. *Biophys. J.*, **2000**, *79* (5), 2754-60.

[36] Senes, A.; Gerstein, M.; Engelman, D. M. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.*, **2000**, *296* (3), 921-36.

[37] von Heijne, G. Proline kinks in transmembrane alpha-helices. *J. Mol. Biol.*, **1991**, *218* (3), 499-503.

[38] Chakrabarti, P.; Chakrabarti, S. C--H...O hydrogen bond involving proline residues in alpha-helices. *J. Mol. Biol.*, **1998**, *284* (4), 867-73.

[39] Flanagan, C. A GPCR that is not "DRY". *Mol. Pharmacol.*, **2005**, *68* (1), 1-3.

[40] Surgand, J. S.; Rodrigo, J.; Kellenberger, E.; Rognan, D. A chemogenomic analysis of the transmembrane binding cavity of human G-protein-coupled receptors. *Proteins*, **2006**, *62* (2), 509-538.

[41] Shi, L.; Javitch, J. A. The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. USA*, **2004**, *101* (2), 440-5.

[42] Mills, J. S.; Miettinen, H. M.; Cummings, D.; Jesaitis, A. J. Characterization of the binding site on the formyl peptide receptor using three receptor mutants and analogs of Met-Leu-Phe and Met-Met-Trp-Leu-Leu. *J. Biol. Chem.*, **2000**, *275* (50), 39012-7.

[43] DeMartino, J. A.; Konteatis, Z. D.; Siciliano, S. J.; Van Riper, G.; Underwood, D. J.; Fischer, P. A.; Springer, M. S. Arginine 206 of the C5a receptor is critical for ligand recognition and receptor activation by C-terminal hexapeptide analogs. *J. Biol. Chem.*, **1995**, *270* (27), 15966-9.

[44] Liapakis, G.; Ballesteros, J. A.; Papachristou, S.; Chan, W. C.; Chen, X.; Javitch, J. A. The forgotten serine. A critical role for Ser-203(5.42) in ligand binding to and activation of the beta 2-adrenergic receptor. *J. Biol. Chem.*, **2000**, *275* (48), 37779-88.

[45] Greasley, P. J.; Fanelli, F.; Rossier, O.; Abuin, L.; Cotecchia, S. Mutagenesis and modelling of the alpha(1b)-adrenergic receptor highlight the role of the helix 3/helix 6 interface in receptor activation. *Mol. Pharmacol.*, **2002**, *61* (5), 1025-32.

[46] Shapiro, D. A.; Kristiansen, K.; Weiner, D. M.; Kroeze, W. K.; Roth, B. L. Evidence for a model of agonist-induced activation of 5-hydroxytryptamine 2A serotonin receptors that involves the disruption of a strong ionic interaction between helices 3 and 6. *J. Biol. Chem.*, **2002**, *277* (13), 11441-9.

[47] Montanelli, L.; Van Durme, J. J.; Smits, G.; Bonomi, M.; Rodien, P.; Devor, E. J.; Moffat-Wilson, K.; Pardo, L.; Vassart, G.; Costagliola, S. Modulation of ligand selectivity associated with activation of the transmembrane region of the human follitropin receptor. *Mol. Endocrinol.*, **2004**, *18* (8), 2061-73.

[48] Yao, X.; Parnot, C.; Deupi, X.; Ratnala, V. R.; Swaminath, G.; Farrens, D. L.; Kobilka, B. Coupling ligand structure to specific conformational switches in the beta-2 adrenoreceptor. *Nat. Chem. Biol.*, in press.

[49] Scheer, A.; Fanelli, F.; Costa, T.; De Benedetti, P. G.; Cotecchia, S. Constitutively active mutants of the alpha 1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. *EMBO J.*, **1996**, *15* (14), 3566-78.

[50] 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* (5288), 768-770.

[51] Huang, P.; Visiers, I.; Weinstein, H.; Liu-Chen, L. Y. The local environment at the cytoplasmic end of TM6 of the mu opioid receptor differs from those of rhodopsin and monoamine receptors: introduction of an ionic lock between the cytoplasmic ends of helices 3 and 6 by a L6.30(275)E mutation inactivates the mu opioid receptor and reduces the constitutive activity of its T6.34(279)K mutant. *Biochemistry*, **2002**, *41* (40), 11972-80.

[52] Kenakin, T. New concepts in drug discovery: collateral efficacy and permissive antagonism. *Nat. Rev. Drug Discov.*, **2005**, *4* (11), 919-27.

[53] Swaminath, G.; Xiang, Y.; Lee, T. W.; Steenhuis, J.; Parnot, C.; Kobilka, B. K. Sequential binding of agonists to the beta2 adrenoceptor. Kinetic evidence for intermediate conformational states. *J. Biol. Chem.*, **2004**, *279* (1), 686-91.

[54] Peleg, G.; Ghanouni, P.; Kobilka, B. K.; Zare, R. N. Single-molecule spectroscopy of the beta(2) adrenergic receptor: observation of conformational substates in a membrane protein. *Proc. Natl. Acad. Sci. USA*, **2001**, *98* (15), 8469-74.

[55] Ghanouni, P.; Gryczynski, Z.; Steenhuis, J. J.; Lee, T. W.; Farrens, D. L.; Lakowicz, J. R.; Kobilka, B. K. Functionally different agonists induce distinct conformations in the G protein coupling domain of the beta 2 adrenergic receptor. *J. Biol. Chem.*, **2001**, *276* (27), 24433-6.

[56] Gether, U.; Ballesteros, J. A.; Seifert, R.; Sanders-Bush, E.; Weinstein, H.; Kobilka, B. K. Structural instability of a constitutively active G protein-coupled receptor: Agonist-independent activation due to conformational flexibility. *J. Biol. Chem.*, **1997**, *272*, 2587-2590.

[57] DeLano, W. L. The PyMOL Molecular Graphics System. *On World Wide Web at http://www.pymol.org* **2002**.

[58] Clamp, M.; Cuff, J.; Searle, S.; Barton, G. The Jalview Java alignment editor. *Bioinformatics*, **2004**, *20* (3), 426-7.

Received: June 19, 2006