ARBS Annual Review of Biomedical Sciences

Theme Topic on "Cell Receptors and Signaling"

pdf freely available at http://arbs.biblioteca.unesp.br

2009;11:T51-85

Family A G-Protein Coupled Receptors: an Overview of Structure-Function Relationships

Laerte Oliveira[*,†]

Department of Biophysics, Universidade Federal de São Paulo, São Paulo, SP, BRAZIL

Received: 14 August 2009; accepted 17 January 2010
Online on 21 February 2010

Abstract

Oliveira L. Family A G-Protein Coupled Receptors: an overview of structure-function relationships. Annu Rev Biomed Sci 2009;11:T51-85. Family A G-protein coupled receptors (AGPCRs) form the largest group of correlate receptors whose structure, a bundle of seven-trans-membrane (7 TM) helices, may be activated thus becoming able to transduce a signal from the extracellular medium to the cytosol. This activation may be constitutional, for instance due to permanent structural modifications, or be physiologically triggered by agonist binding at an external and accessible specific site. Based on the structures of agonists, AGPCRs may be divided according to pharmacological assays into many classes of receptors, each one comprising many types or sub-types of proteins, as differentiated by specific binding of inhibitors, all of them performing a multitude of functions. It is noteworthy that AGPCRs have been more recently cloned and their sequences of amino acids determined in a large scale, a condition that has allowed these receptors to be sorted by a new criterium. Sequence analyses have consistently matched functional assays for classification of AGPCRs except for a certain number of functionally unknown receptors which have been cataloged as orphan receptors. A colossal number of AGPCRs, more than 10,000 sequences belonging to more than 1,000 different types of receptors, may nowadays be multiply-aligned what has been enabling the determination of parameters of residue conservation and characterization of special motifs along the structure of these proteins. There are at the present time, high-resolution 3D structures for the following AGPCRs: inactive rhodopsin, retinal-free opsin, β-adrenoceptor and adenosine receptors. Among them, hodopsin structures are reliable enough to be used as prototypes for analyses of residue conservation and mechanisms of activation of receptors, specially at the level of the more conserved structure in the cytosolic half of their 7TM bundle.

Keywords: seven-transmembrane receptors, G-protein, rhodopsin, residue conservation, activation

Table of Contents

[*] Retired Adjunct Professor.

[†]**Correspondence**
Laerte Oliveira. Rua Domingos de Morais, 1338, apto 112C, Vila Mariana, 04010-200, São Paulo, SP, Brazil. E-mail: laerteoliveira7@gmail.com

1. Introduction

G-protein coupled receptors (GPCRs) form a big superfamily of membrane-spanning proteins linked to transduction of an extracellular signal across plasma membranes towards intracellular proteins such as those of the G-protein system. Often found constitutionally activated and mostly being activated by binding of a first-messenger (agonist), the receptors can also be triggered by light (opsins), odors (odorant receptors), flavors (gustative receptors) and other types of stimuli.

GPCRs regulate the majority of intracellular processes, being their importance measured by the number of medicines available (about 50% of the total) for treatment of multiple receptor-mediated pathological occurrences in living beings. These receptors are ubiquitous since they are found in numerous living species of all kingdoms, namely of animals, plants, fungi, prokaryotae (bacteria) and protoctista (algae, molds, and protozoans). Precise relative numbers of receptors distributed in all these kingdoms are unavailable yet, considering that studies about this subject are still in progress.

GPCRs are divided into many families according to functional properties and more recently according to their amino acid sequences[1]. Family A-, B- and CGPCRs are more numerous but minor

[1] along the text and in figures, amino acid residues shall be identified by three-letter and single-letter codes, respectively.

groups as DGPCRs (fungal pheromone receptors) and EGPCRs (cAMP receptors) are also known. Here, we are going to address only AGPCRs for which a greater deal of information is available. So far, AGPCRs or rhodopsin-like receptors constitute the most numerous of GPCR families comprising the most known receptors as bioamine, prostanoid, glyco-protein hormone, peptide (chemokine, angiotensin II, bradykinin, endothelin, neuropeptide and others), opsins and most of the odorant receptors. It is aimed at this review to address some relationships between structure and function of these receptors. Besides, it is recommended to the readers the access to GPCRDB website (www.gpcr.org/7tm) where data about GPCRS are more extensively presented (Horn *et al.*, 2003).

2. Making GPCRs Known: First Studies

Long ago before cloning procedures started, structures of GPCRs had already been studied. In fact, the oldest references found about this subject are from the late 70s and early 80s, and practically deal with rhodopsin (and less often with β-adrenoceptors) which is the receptor present at high density in the retina rod cells and, for this reason, more likely to be purified at higher concentrations. In the 80s, complete amino acid sequences of rhodopsins had already been determined by chemical procedures (Hargrave *et al.*, 1983; Ovchinnikov *et al.*, 1983) as well as from cDNA (Ovchinnikov *et al.*, 1988). As a consequence, hypothetical 3D structures had been brought up for discussion thus enabling the formulation of the first models about functioning of this protein (Ovchinnikov, 1982; Findlay & Pappin, 1986). Earlier reports about β-adrenoceptors had already emerged (Yarden *et al.*, 1986).

However, only in the 90s, when the full development of gene recombination techniques had been accomplished, GPCRs were revealed in a very large scale. The process was first slow but shortly later it attained an astonishing pace so that a big superfamily of GPCRs right came into existence, consisting of many families (A,B, C, etc), each one with many types of receptors and sequences defined by specific amino acid profiles, all of them with a multitude of functions.

When enough numbers of GPCR sequences were available, the first procedure largely used to study them was the multiple sequence alignment. Immediately, some basic features of these receptors were revealed, as follows:

(1) The GPCR families could be recognized by a few aspects of amino acid profiles. Each family of receptors has very conserved amino acid residues at well-defined positions of structures, a pattern that was so systematically repetitive that a simple observation of sequences turned out to be a quick and very reliable procedure to identify these proteins (Oliveira *et al.*, 1993).

(2) Hydropathic analyses (Kyte & Doolittle, 1982) of GPCR aligned amino acid sequences revealed seven stretches of 20 to 30 amino acids, mostly aliphatic residues intercalated with a few polar or even charged molecules. As the hydrophobic stretches were long enough to span a membrane bilayer, it was concluded that all these proteins were folded as a seven-trans-membrane-helix (7TM I-VII or A-G) with polar and charged sequences forming terminal and intermediate loop domains in contact with the extracellular and cytosolic aqueous media.

(3) Early analyses could define the location of glycosylation (Fukuda *et al.*, 1979) and phosphorylation sites (Wilden & Kühn, 1982) at the N- and C-terminal segments of the GPCR sequences, respectively, thus leading to the conclusion that the receptors have their primary structures beginning at the extracellular side and ending at the cytosolic sides of the membranes.

(4) A significant event in GPCR research was that the sequence of rhodopsin and opsins in general could be matched to the consensual sequence of AGPCRs. For this reason, this receptor family was called rhodopsin-like receptors.

(5) At this time, bacteriorhodopsins were already known displaying several features found in rhodopsin. First, they were activated by light absorption through a retinal moiety. Second, their high-resolution structures (Henderson *et al.*, 1990) showed a 7TM bundle surrounding a central cavity, with the extracellular half presenting a retinal moiety forming a Schiff base (SB) bond with a helix VII lysine (Lys^{1}) residue. Both extracellular and cytosolic sides of the structure central cavity were open to extra-membrane spaces.

(6) As rhodopsin and all opsins also had a Lys residue at the same relative position of the bacteriorhodopsin helix VII, it was concluded that all these opsins could make a similar Lys-linked SB and, as all AGPCRs, share the same 3D structure of bacteriorhodopsin.

(7) Thus, during the 90's decade, AGPCRs were approached as if they were structurally and functionally similar to bacteriorhodopsin (Khorana, 1993). Mechanisms of proton transport via hydrogen bond network found in this protein were applied in order to interpret functional features of receptors (Maeda *et al.*, 1997). Models of AGPCR structures were systematically built by homology to bacteriorhodopsin structure (Trumpp-Kallmeyer *et al.*, 1992; Oliveira *et al.*, 1993). All this happened despite the fact that, as early as 1993 and shortly later, low-resolution structures of rhodopsin had already been determined showing a differently-folded 7TM bundle (Schertler *et al.*, 1993; Unger *et al.*, 1997).

(8) Delimiting the beginning of a new era for GPCR researches, high-resolution structure of rhodopsin (2.8 A) was finally determined (Palczewski *et al.*, 2000) bearing a bacteriorhodopsin-like docking of retinal through a SB at the protein helix VII Lys residue. On the other hand, the new structure showed some striking differences in relation to bacteriorhodopsins such as: (1) the pattern of 7TM bundle folding; (2) the unexpected structure of the the second extracellular loop (EC2) arranged as a distorted hairpin placed inside the 7TM bundle and closing any free communication between the structure central cavity and the extracellular medium; (3) the existence of a 90^{o} bend in the protein main chain right after helix VII followed by a cytosolic helix VIII disposed parallelly to the membrane interface, which confirmed previous prediction of structure arisen from comparative sequence-structure studies of all AGPCRs (Oliveira *et al.*, 1999).

(9) From this time on, studies of AGPCRs have been addressed using structures of rhodopsins as templates. However, the determination of other receptor structures along the last two years, opens the possibility that this picture is likely to change soon.

(10) While this does not happen, a rhodopsin-like design has been used to mimic the general fold of AGPCRs which consists of a bundle of seven membrane-spanning (7TM) helices (I-VII) surrounding an ellipsoidal central cavity (Fig. 1A-1C). The structure starts as an extracellular domain (N_t) and ends as a cytosolic domain (C_t), both of variable lengths, with the helices linked by three extracellular loops EC1, EC2 and EC3 (in order, between II-III, IV-V and VI-VII pairs of helices) and three cytosolic loops IC1, IC2 and IC3 (in order, between I-II, III- IV and V-VI pairs of helices).

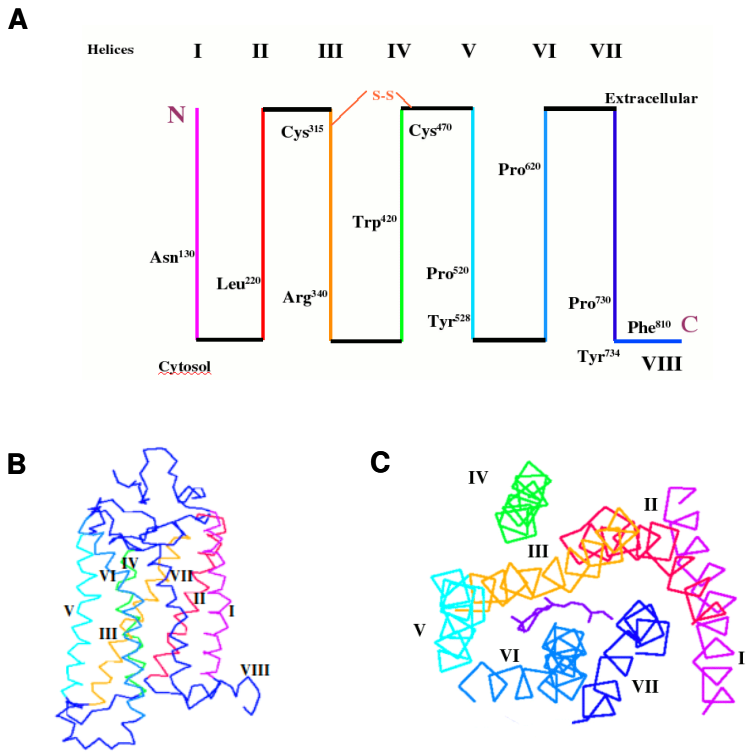

Figure 1. 7TM bundle of AGPCRs (Vriend, 1990). (A) I-VII stands for the 7TM helices and VIII is an additional cytosolic helix; N and C are the N- and C-terminal ends of the structure in the extracellular and cytosolic media, respectively. Positions of conserved residues used as references for alignment are shown for all helices with the normalized numbers. A conserved disulfide bond linking the top of helix III to EC2 loop, is also shown. Longitudinal view (B) and cross-section (C) of the 7TM bundle of inactive rhodopsin structure showing the relative positions of helices with the retinal inside of the central cavity.

3. Structure of AGPCRs

3.1. Normalized numbering system

The numbering system for residue positions in proteins follows the general sequence of the amino acids without considering regional distribution of specific protein domains. Thus, if two proteins of the same class have different total numbers of residues due to insertions or deletions (indels), comparative analysis of their sequences faces a very hard difficulty which can only be circumvented by the adoption of a normalized numbering system based on structural patterns. In the case of AGPCRs, the usual procedure for normalization of numbering is to define from known 3D structures, segments of the sequences that form stable secondary structures (helices and turns) and to select as references for each of these segments, a conserved residue position designated by a strategic number. The positions of each segment are then numbered according to the respective reference. Among some available procedures for numbering AGPCRs (Ballesteros *et al.*, 2001), we shall adopt that used in GPCRDB (Horn *et al.*, 2003).

Positions of GPCRs are identified by three-digit numbers: the first digit (1-8) stands for the helix or other adjoining domains of the sequences; the other two digits stand for the location of the position in the respective segment in relation to the reference. As an illustration, Fig. 2 shows the regular and normalized numbers for the positions in AGPCR (rhodopsin) sequence and the limits of the helices and the loops. In the text, positions of rhodopsin and AGPCRs shall be identified by the three-letter code name of the residue followed by two superscripted numbers, the normalized and the natural

ones, being the second written within brackets. For instance, the position 55 in the rhodopsin sequence (Fig. 2) is identified as Asn$^{130(55)}$.

The system applied for numbering the rhodopsin sequence (designated as procedure 1) is used to number the 7TM structure in the majority of AGPCRs which have referential conserved residues identified. For the highly-variable extra-membrane segments of these receptors, a complementary numbering scheme is adopted. Extra-membrane sequences of AGPCRs may suggest the presence of special motifs (identifiable from inserted stretches of conserved amino acids) in some classes of AGPCRs. In this case, provided a reference and a strategic number are selected, the procedure 1 is applied for numbering the sequences. If no special motif is found, the segment is then numbered according to a procedure 2 as follows. The segment of sequence in question is divided into two halves which are numbered following the referential number of the corresponding adjoining helix. For N_t and C_t segments, only the references of helices I and VIII are used, respectively. Unless when located up to about five positions from boundaries of helices, the residue positions numbered according to the procedure 2 are not suitable for comparative analyses since they could be part of quite different structures. The numbers then used to identify them are useful for considerations pertaining a same class of proteins.

A special care in these numbering procedures is that whatever a motif is present in a sequence segment of some receptor classes and a referential position is selected for procedure 1, the number of this reference should not coincide with the numbers used in the numbering procedure 2 applied to the same segment of other receptors without the motif.

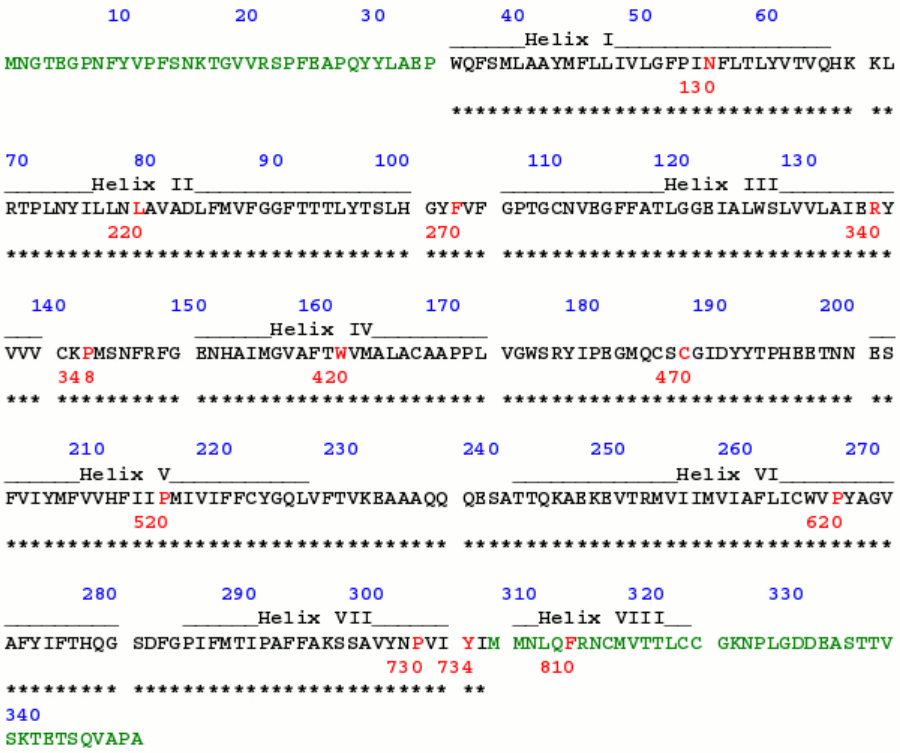

Figure 2. Sequence of bovine rhodopsin. Amino acid residues are shown in one letter code: natural sequence numbers (blue) and normalized numbers (red); residues of sequences included in multiple alignment of all receptors are in black; opsin-specific sequences are in green; referential residues used for alignment and normalized numbering are in red; the extensions of helices I-VIII and mostly frequently aligned sequences are labeled by _s and *s,respectively.

3.2. GPCR fold based on 3D structures

3D structures have been reported for the following AGPCRs: inactive (dark) rhodopsin (Palczewski *et al.*, 2000; Li *et al.*, 2004; Okada *et al.*, 2004), retinal-free opsin (Park *et al.*, 2008) and retinal-free opsin coupled to C-terminal peptide of Gát chain (Scheerer *et al.*, 2008), squid rhodopsin (Murakami & Kouyama, 2007), β_1-adrenoceptor (Warne *et al.*, 2008), β_2-adrenoceptor (Cherezov *et al.*, 2007; Hanson *et al.*2008) and adenosine A2a receptor (Jaakola *et al.*, 2008). All of them have shown the general fold described initially for rhodopsin (Fig. 1A and 1B). A few reports about structure determinations of BGPCRS, CGPCRS and others, are for extra-membrane domains (Muto *et al.*, 2007; Pioszak *et al.*, 2008).

The structure of inactive rhodopsin (Li *et al.*, 2004; Okada *et al.*, 2004) is herein used as a basis for future analyses regarding AGPCRs. In a cross section of this structure, the broader sides of the central cavity are flanked by helices II-III and VI-VII and the narrower sides are closed by helices I and V (Fig. 1C). The helix IV is not part of the bundle since it is not making contact with the structure central cavity. This cavity is accessible from the intracellular side despite the existence of interactions linking the cytosolic ends of helices I, II, III, V, VI and VII. The extracellular side is closed by the EC2 loop hairpin which lies inside the central cavity parallel to the membrane surface making interactions with side-chains of helices. The most remarkable of these interactions is a disulfide bridge ($Cys^{315(110)}$-$Cys^{470(187)}$) between EC2 loop and the top of helix III which is found in a large majority of AGPCRs (Fig. 3A). After helix VII, the main chain of rhodopsin structure is bent and continues as a cytosolic helix (helix VIII) parallel to the membrane. Helix VIII is linked to helix VII, to the VII-VIII bend and to helices I and II by side-chain interactions such as that involving $Tyr^{734(306)}$ and $Phe^{810(313)}$. (Fritze *et al.*, 2003). This interaction together with other bond between helix III $Arg^{340(135)}$ and helix VI $Glu^{600(247)}$ (Fig. 3A) can contribute to keep the integrity of the 7TM bundle or the inactive state of AGPCRs (Ballesteros *et al.*, 2001).

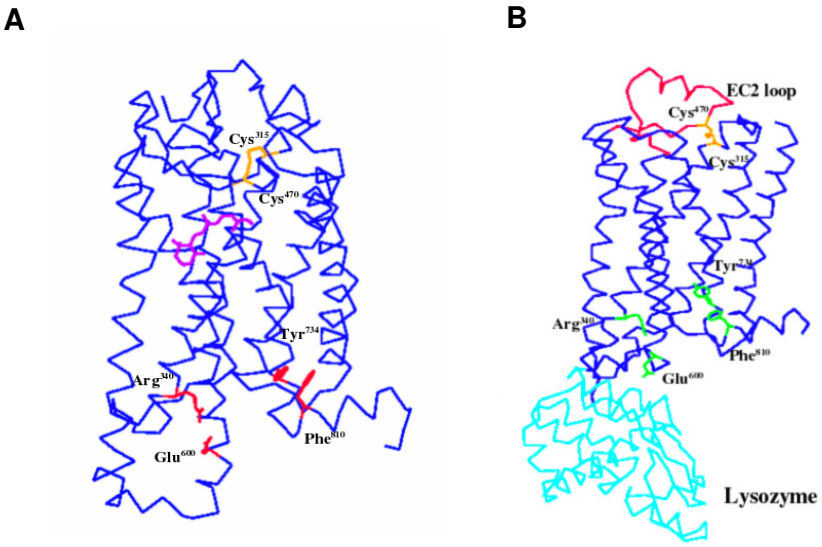

Figure 3. (A) The structure of inactive rhodopsin shows in a longitudinal view the retinal (magenta), the conserved disulfide bond (gold) and two basic interactions between cytosolic ends of helices (red) which are known to stabilize the inactive receptor. (B) Structure of â$_2$-adrenoceptor with a molecule of lysozyme inserted in the IC3 loop. Note the EC2 loop (red) displaced from the location observed in rhodopsin but keeping the conserved disulfide bond (gold) with helix III, the N_t segment truncated and free, and the cytosolic bonds (green) between the pairs of residues $Arg^{340(135)}$-$Glu^{600(247)}$ and $Tyr^{734(306)}$-$Phe^{810(313)}$ disrupted indicating that the receptor might be under a constitutively activated form.

Besides a 7TM bundle, the structures of β-adrenoceptors (Warne *et al.*, 2008; Cherezov *et al.*, 2008; Hanson *et al.*, 2008; Bokoch et al., 2010) and adenosine (Jaakola *et al.*, 2008) receptors show N_t segments partially unfolded but keeping intact the conserved disulfide bridge between helix III $Cys^{315(110)}$ and EC2 loop $Cys^{470(187)}$ (see Fig. 3B for β-adrenoceptor). They also have the most N-terminal segment (extracellular side) of helix I dissociated, a feature which, differently from the pattern observed in rhodopsins, made the structure central cavity accessible from the external medium. Furthermore, also the cytosolic mouth of the 7TM central cavity of β-adrenoceptors and adenosine receptors are also open to the cytosolic medium, what is attested by disruptions of interactions between helix III $Arg^{340(135)}$ and helix VI $Glu^{600(247)}$ or VII-VIII bend $Tyr^{734(306)}$ and helixVIII $Phe^{810(247)}$.

These characteristics might a priori be associated with a condition of activated receptor since they are observed in retinal-free opsin which, as it will be discussed below, is a constitutively activated protein (Park *et al.*, 2008). Another possibility is to assume that β-adrenoceptors exist under two conformations, open or closed, in consonance with states in which the bond between helix III $Arg^{340(135)}$ and helix VI $Glu^{600(247)}$ is present or not (Dror *et al.*, 2009).

3.3. How receptors may be identified

AGPCRs may be identified according to the types of conserved residues and their positions in the protein structure (see residues and numbers in red in Fig. 2):

(1) Asn^{130} at the C-terminal third of helix I;
(2) Leu^{220} at the N-terminal third of helix II, more often at 24 positions after the helix I conserved Asn^{130} residue; four positions after the helix II Leu^{220} position, there is a conserved Asp^{224};
(3) Trp^{270} (or Phe^{270}) at the middle of EC1 loop, more often at 23 positions after the helix II Leu^{220} position;
(4) Arg^{340} residue at the C-terminal third of helix III, more often at 25 positions after a conserved helix III Cys^{315} residue;
(5) Pro^{348} at the middle of IC2 loop, more often at 7 residues after the helix III Arg^{340} position:
(6) Trp^{420} at the middle of helix IV, more often at 20 positions after the IC2 loop Pro^{348} position;
(7) Cys^{470} at the middle of EC2 loop, residue which makes disulfide bond with helix III Cys^{315} residue;
(8) Pro^{520} at the middle of helix V;
(9) Pro^{620} at the C-terminal third of helix VI;
(10) Pro^{730} at the C-terminal third of helix VII;
(11) Tyr^{734}(Phe^{734}) at the bend between helices VII and VIII, more often at 3 positions after the helix VII Pro^{730} residue;
(12) Phe^{810}(or aliphatic side-chains) at the middle of helix VIII, more often at 10 positions after helix VII Pro^{730} residue.

3.4. Functional Classification

Procedures aiming at the identification and characterization of AGPCRs based on functional criteria have been carried out since the beginnings of the pharmacological sciences. In a first stage, substances were characterized as agonists and their specificity and efficiency were used to define classes of receptors as adrenergic, muscarinic, dopaminergic and so on. Afterwards, a more refined phase was reached in which receptors were discriminated at the level of types and subtypes with the use of receptor inhibitors, most of them synthetic compounds. The extensive development of cloning procedures along the 90s decade led AGPCRs to have their amino acid sequences determined and studied thus complementing the receptor classification task. Nowadays, with a large numbers of receptors identified, it is then possible to compare results coming from both experimental approaches and show the level of correspondence between them.

The process of AGPCR identification has progressed over at least two different phases. After an initial slow development probably by the lack of an adequate technology, a first boom of receptor

discovery started about some decades ago with the identification of multiple categories of AGPCRs but without a large increase of the number of receptors itself. In a second phase, coinciding with the improvement of recombinant gene technology, there was an expressive impulse for genome projects leading to an astonishing increase of the total number of receptors, however without increasing the number of different classes of these proteins.

The explanation for these disparities is that in the first phase only few mammal species were available for research, but despite of that, investigations could already identify most of the existing classes of receptors. It was found in these cases that a same type of receptors taken from different mammal species had sequences of amino acids with high identity chiefly in the 7TM bundle helices and a very similar total length indicating absence of indels. Further, in the second phase, material from multiple species of living beings, including vertebrates, invertebrates and even unicellular organisms, were available for study thus allowing to show that the same receptor classes were present in organisms of different taxonomic classes but with the difference that the sequences of these proteins were very frequently longer than the corresponding sequences of mammal species due to many indels observed in the N_t and C_t segments and EC1, EC3 and IC3 loops.

The classes (including types and sub-types) of AGPCRs obtained from a classification based on sequences of amino acids are shown in the Table 1 [Appendix 1][a]. Previously tested by pharmacological assays, most of these receptors are identified by the functional names. Table 2 [Appendix 2][b] shows sequences of most of AGPCRs, but with residues aligned only at the 7TM helix segments. It is remarkable the size of missing sequences shown within brackets, denoting the extension with which, along the evolution, insertions were made in the extra-membrane regions of the receptors.

3.5. Multiple sequence alignment

Two decades ago, when studies on AGPCRs were still beginning, a supposedly conserved feature of these receptors was that a long segment of their sequences, consisting of a full stretch of residues from helix I to helix IV and separate sequences of helices V, VI, VII and VIII, were quite prone to be easily aligned. In fact, in all receptors then available, this segment had the same length and the same referential conserved residues placed at the same relative positions. However, following the booming development of cloning studies on various living species over the last two decades, this framework has considerably changed since new classes of AGPCRs, or previously undetected variations in the sequences of old receptors, have been brought up for analysis at a consistently high pace. Altogether, as a consequence of these data, a new general consensus emerged that there are no regions in AGPCR sequences, conserved enough to be easily aligned, specially when a quite heterogeneous collection of receptors is analyzed.

Alignment of AGPCR sequences has been performed according to sequence-structure profiles of maximal homology (MaxHom) (Sander & Schneider, 1991) built for different types and sometimes sub-types of receptors. A MaxHom profile is a file in which each line contains information about alignment for each position of the sequence. The first columns of the data correspond to the alignment of secondary structures. It has been assumed that whatever helices were formed in a certain segment of the sequence, no deletion or insertion of residues is admitted in the alignment. In order to drive the alignment according to this rule, large numbers, meaning heavy penalties for accepting residue insertion or deletion, are written in the first columns of the profile lines corresponding to positions where changes in the sequences are forbidden. The same principle is used to select the type of residue side-chain. For each position of the sequence, numbers proportional to fractions of the twenty amino acids found are added to the subsequent columns of the profile corresponding line.

Proper algorithms can read these so built profiles and thus drive the process throughout the less penalized pathway to attain a reliable, significant and representative alignments of sequences. The criteria by which alignments are so defined are (1) absence of insertions or deletion of residues; (2) match of fully conserved residues all over the alignment; (3) presence of correlated residue mutation patterns characterized by conservation of residues in tandem along receptors of a same class; and mainly (4) the matching between positions of the sequence and positions of the 3D structure.

[a]Appendix 1: http://arbs.biblioteca.unesp.br/appendixes/appendix_1.pdf
[b]Appendix 2: http://arbs.biblioteca.unesp.br/appendixes/appendix_2.pdf

Since the knowledge about AGPCR structures is limited yet, the alignment of quite heterogeneous collections of sequences, namely those of all receptors or belonging to many different classes of these molecules, can only be significantly aligned for short segments corresponding to helices and adjoining loops (see rhodopsin positions identified by *s in Fig. 2). However, when more homogeneous series of AGPCRs are set for analysis, as those consisting of types and subtypes of receptors, the alignment can be extended to other regions of the sequences, sometimes attaining almost the fullness of them. In these cases, the procedure turns out to be more a match by residue conservation than a match by structure.

The only way to fulfill the criteria of significance for the alignment of all AGPCRs, is to perform this procedure through at least two steps: (1) Full sequences of types (sub-types) of receptors intended to be analyzed, are separately aligned according to the corresponding profiles; (2) the separate aligned sequences of the different types are joined and the segments of them whose alignments are likely to be significant, are matched according to the references of conserved residues. These references are shown in Fig. 2 for rhodopsin or AGPCR sequences and are the same selected for numbering procedures as it will be discussed in next sections.

More than 10,000 sequences comprising helices and adjoining loops of all AGPCR classes described (see above) were aligned by the two-step procedure (Table 3) [Appendix 3][c]. In order to make the alignments fully representative, the C-terminal three residue positions of helix II and the positions of the entire C_t segment were deleted since they were not found or their residues were not well-characterized in some classes of receptors, such as C-X-C chemokine, chemokine-like, vasopressin, endothelial differentiation (sphingolipid) GPCRs (EDGs) and many other (specially orphan) receptors.

3.6. Aligment and numbering of special motifs

AGPCRs contain many probably-inserted sequence motifs in their extra-membrane segments which are found in different and apparently-unrelated classes of receptors and require specific referential positions for numbering. Profiles might be built with these references but proper weights must be attributed to them in order to allow precise alignment of residues. Features of these special motifs which should be presented soon in the new version of GPCRDB (Horn *et al.*, 2003), are shown here in a resumed form.

3.6.1. Nt segment and EC3 loop

Two short insertions of 5 to 10 residues may be found at the sequences of angiotensin II, bradykinin, endothelin, chemokine, Cys-leukotriene, purinergic, some neuropeptide and other AGPCRs. The former insertion is in the N_t segment and the latter in the middle of EC3 loop, both containing a conserved Cys residue supposedly forming a second extracellular disulfide bond (Fig. 4) [Appendix 4][d]. In these AGPCR classes, the numbering procedure 1 described above should be applied being the numbers 100 and 650 designated for the N_t and EC3 loop Cys residues.

Numbering N_t segments of AGPCRs, others than those possessing a second disulfide bond, should be applied with numbers below 100, as shown for rhodopsin and TSHR (Fig. 4 A) [Appendix 4][d], in order to avoid overlapping with the numbers attributed to the insertion bearing Cys^{100}. For receptors containing no EC3 loop insertion, the numbering is carried out by means of the procedure 2 described above as shown for rhodopsin and TSHR (Fig. 4B) [Appendix 4][d].

3.6.2. Cytosolic helix VIII and helix VII-helix VIII bend

The characterization of a cytosolic helicoidal structure (helix VIII) together with the finding of an EC2 loop hairpin, was an astonishing surprise of the rhodopsin structure. Despite being at first considered an artifact due to crystal contact, the helix VIII was later equally found in other structures of rhodopsin and other AGPCRs and thus, its existence was no longer matter of dispute.

From the observation of many thousands of AGPCR sequences in comparison with pertinent features of rhodopsin structure, it has been suggested that the cytosolic segment of the sequence following helix VII has some special characteristics as: (1) the main chain bend before the start of helix VIII; (2) a very conserved residue of Tyr at two or three residues after the referential position of helix VII Pro^{730};

[c]Appendix 3: http://arbs.biblioteca.unesp.br/appendixes/appendix_3.pdf
[d]Appendix 4: http://arbs.biblioteca.unesp.br/appendixes/appendix_4.pdf

and (3) the helix VIII (Fig. 5) [Appendix 5][e]. Thus, for numbering the region accordingly: (1) the end of helix VII is set as the position 732; (2) the position 733 is set as a first insertion; (3) the conserved Tyr residue is set as the referential position 734 for numbering the bend between helices VII and VIII; (4) the positions 736, 737, etc, are set as other insertions before the starting of the cytosolic helix VIII (Fig. 5) [Appendix 5][e].

The segment corresponding to the helix VIII of rhodopsin is variable along the sequences of all AGPCRs. It usually has a pair of conserved residues, FR, scarcely mutating to Y (or aliphatic residues) and to Lys, respectively, but there are still a large number of the receptors whose sequences after Tyr^{734} are not indicating the existence of a regular helix VIII at all (Fig. 5) [Appendix 5][e]. Numbering of AGPCR helix VIII, within the length shown in Fig. 2 for rhodopsin, was performed by means of the procedure 1 described above for receptors containing the FR pair (or mutant variants)with the number 810 assigned as reference for the aromatic residue and the number 806 for the starting position of the helix.

Several classes of AGPCRs as opsins, adrenoceptors, EDG and other receptors display a pair of Cys residues or a single Cys residue at the end of helix VIII (positions 819 and 820 in rhodopsin) (Fig. 5) [Appendix 5][e]. Once palmitoylated these residues can anchor at the membrane lipid moiety thus forming a fourth cytosolic loop and keeping all extension of helix VIII at short distance from the membrane bilayer. As discussed above for N_t segment, numbering of C_t segment of AGPCRs when helix VIII and the Cys residues at the end of this helix are not detected, should follow the procedure 2 described above. In these cases, numbers above 820 should be used thus avoiding the overlapping with numbers used in the procedure 1 applied to the helix VIII.

3.6.3. EC1 loop

Displaying a long history of residue indels, the EC1 loops of AGPCRs should be an accessible and flexible part of the receptor structure (Fig. 6) [Appendix 6][f]. In most of non-olfactory AGPCRs, EC1 loop has a referential conserved residue, Trp^{270} (less often of Phe and Tyr), in a β-turn as that seen in the rhodopsin structure (Fig. 2). Also, at the top of helix III, there are two referential conserved residues, Gly^{311}, the first position of Helix III placed at three positions after Trp^{270} residue and Cys^{315}, a partner of a conserved disulfide bond of GPCRs placed at four positions from Gly^{311}. Thus, receptors with all these features should be numbered through the procedure 1 described above, first for EC1 loop using the reference 270 and further following the regular numbering of helix III using as reference Cys^{315} or even Arg^{340} (Fig. 6A and 6B) [Appendix 6][f].

Some classes of receptors with the referential Cys^{315}, and mostly with Gly^{311}, are target for several residue indels of variable lengths in positions before and after the reference Trp^{270} (Fig. 6A) [Appendix 6][f]. Glyco-protein hormone receptors (FSHR, LSHR, TSHR, GSH1 and GSH2) have an insertion of 8 residues before the reference 270, forming a little domain which interact with the N_t domain of these receptors, being by this way involved in the mechanism of activation. These receptors also have a deletion of one position after the reference 270. The Leu-rich-repeat (LGRs) and relaxin receptors also present a glycoprotein-hormone-receptor-like insertion and deletion whereas endothelin ET1 (or ETA) receptors have an insertion of five residues but after the Trp^{270} reference. Within the same category of hormone receptors and LGRs, some prostanoid receptors (EP3, prostaglandins F, F2 and F2a, and thromboxane A2) have insertions in the EC1 loop before and after the Trp^{270} reference.

Also, a long list of other receptors [neurotensins, orexins, ecdysis triggering hormone, melatonin, prolactin releasing hormone, neuromedin U, PRX (pyrokinin) amide peptides, prokinetocinin, SREB (super conserved *r*eceptor expressed in brain), serotonin 1c, 2a, 2b, 5a, 5b and 7, dopamine D3 and D4] have an insertion but only before Trp^{270} reference. Other receptors such as histamine H4, Ile8, chemokine C-C1, C-C2, C-C4, C-C5, C-C7, C-C8, C-C9 and C-C11, have a deletion of a single position before Trp^{270} (Fig. 6A and 6B) [Appendix 6][f].

On the other hand, the remaining groups of AGPCRs. including olfactory, some prostanoid (EP1, EP2, Prostaglandin I and Prostacyclin I2), adenosine and viral receptors, have the conserved Cys^{315} position, scarcely the conserved Gly^{311} position, but not the referential position of Trp^{270}. Thus, they

[e]Appendix 5: http://arbs.biblioteca.unesp.br/appendixes/appendix_5.pdf
[f]Appendix 6: http://arbs.biblioteca.unesp.br/appendixes/appendix_6.pdf

should have their EC1 loop numbered according to the procedure 2 described above (Fig. 7) [Appendix 7][g].

Melanocortin, mas, cannabinoid, EDG, growth factor and tubulin receptors are devoid of the Gly^{311}, Cys^{315} and Trp^{270} residues and have the space of EC1 loop occupied by variable number of residue positions which are accordingly also numbered by the procedure 2 (Fig. 8). [Appendix 8][h].

3.6.4. EC_2 loop

After the end of helix IV at the position 431, the sequence of opsins, muscarinic, α1-adrenoceptor, serotonin 1A and 7, histamine H1-3, octopamine, adenosine, cannabinoid, EDG1-8 and other receptors show an inserted motif of three or four residues ending with a conserved Trp. For numbering purposes, Trp was used as reference 435 and the segment was identified with numbers varying from 432 to 435. Following this short insert, the variable EC2 loop itself was then numbered by the procedure 1 described above using the reference Cys^{470} (Fig. 9A and 9B) [Appendix 9][i].

The EC_2 loop of olfactory receptors seems to be an expanded form of the loop observed in most AGPCRs. It has more often three Cys residues, intercalated with sequences of non-conserved 9 residues (Fig. 10) [Appendix 10][j]. Only the intermediate position is fully conserved and thus it has been identified as the reference Cys^{470}, a residue that makes a disulfide bond with helix III Cys^{115}. The other two Cys positions are designated by numbers 460 and 480.

For other AGPCRs (mas, melanocortin and tubulin) containing no conserved Cys residues at the EC2 loop, the procedure 2 described above is adopted for numbering this extracellular region of receptors (Fig. 11) [Appendix 11][k]. The same procedure is applied for numbering IC1 (Fig. 12) [Appendix 12][l] and IC3 (Fig. 13) [Appendix 13][m] AGPCR loops.

3.6.5. IC2 loop

The length of AGPCR IC2 loop is reasonably conserved presenting a very conserved residue of Pro designated by the number 348 as referential position for numbering the segment of sequence in question. This position is at 20 positions from the helix IV referential position Trp^{420} but variations may occur as shown in chemokine receptors (Fig. 14A, 14B and 14C) [Appendix 14][n]. Some sequences present an insertion between the helix III referential position Arg^{340} and Pro^{348}. According to rhodopsin structure, all along the N-terminal portion of IC2 loop, only at the end of helix III (corresponding to rhodopsin Val^{344} in Fig. 14A) [Appendix 14][n] is that residues might be added without producing sterical hindrance. Thus an insertion was exactly set at the position 345 being the next position (346) the start of the common IC2 loop sequence. Fig. 14A [Appendix 14][n] shows the position 345 occupied by an additional Cys residue in serotonin 4 receptors.

IC2 loop was early recognized in rhodopsin as target for extensive conformational changes upon light activation. As this region is conserved in AGPCRs, it is presumed that the same mechanism should be operative in most receptors of this family.

3.6.6. Glycosylation and phosphorylation sites

Numerous experimental findings have been showing that there is no consensual sequences in phosphorylation sites of AGPCRs and that receptors can be multiply-phosphorylated at sites distributed all over the intracellular regions of the receptors. It seems that the role of phosphorylation is to add a general negative charge to the cytosolic loop end thus favoring the binding of adaptors (Tobin, 2008). However, on the other hand, 3D structures of arrestin-phosphorylated-receptor complexes indicate that the adaptor binding is preferentially directed to the most extreme C-terminal end of the receptor chain.

[g]Appendix 7: http://arbs.biblioteca.unesp.br/appendixes/appendix_7.pdf
[h]Appendix 8: http://arbs.biblioteca.unesp.br/appendixes/appendix_8.pdf
[i]Appendix 9: http://arbs.biblioteca.unesp.br/appendixes/appendix_9.pdf
[j]Appendix 10: http://arbs.biblioteca.unesp.br/appendixes/appendix_10.pdf
[k]Appendix 11: http://arbs.biblioteca.unesp.br/appendixes/appendix_11.pdf
[l]Appendix 12: http://arbs.biblioteca.unesp.br/appendixes/appendix_12.pdf
[m]Appendix 13: http://arbs.biblioteca.unesp.br/appendixes/appendix_13.pdf
[n]Appendix 14: http://arbs.biblioteca.unesp.br/appendixes/appendix_14.pdf

It is advisable that residues of Ser and Thr, involved in the phosphorylation process, should be identified by special numbers (between 950 and 999) according the different classes of receptors. The same reasoning may be indicated for numbering positions linked to glycosylation (numbers between 1-10).

4. Function in AGPCRs

4.1. Residue conservation patterns

Residue conservation patterns were characterized from aligned sequences by entropy-variability analysis as described elsewhere (Oliveira *et al.*, 2003a). A procedure aiming at eliminating redundancies in sequence sampling, was used as summarized below. Aligned sequences presenting 95% or more of residue identity all over the alignment positions were grouped and mean values of parameters for the groups of sequences, instead of individual values for each sequence, were used in further calculations. Weights whose values were proportional to the total number of sequences and inversely proportional to residue variability, were assigned for each group.

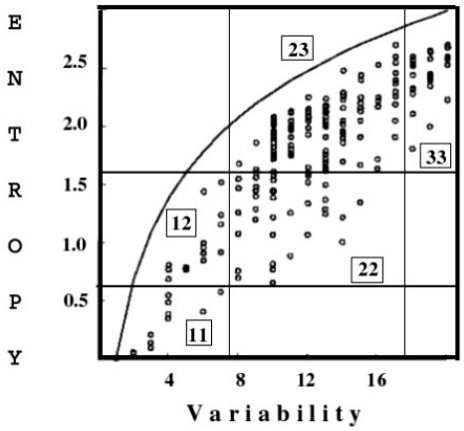

Figure 15. Entropy/Variability plot calculated for residue positions in multiple sequence alignment of about 10,000 AGPCRs. Variability is the number of different residues (1-20) found at each position. Entropy S (0.0-3.0) is given by the equation $S= -\sum (i=1\text{-}20) f_i(\ln f_i)$ where f_i is the fraction of each amino acid found at each position. In the numbers designating boxes, 11, 12, 22, 23 and 33, the first and the second digits express increasing ranges of V and E values, respectively. Thus, 11 and 33 are boxes containing highly conserved and highly variable residue positions, respectively.

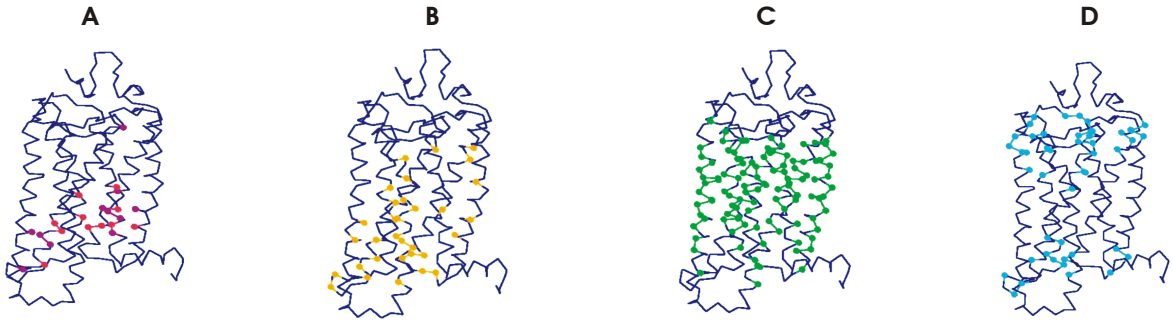

Figure 16. Bovine inactive rhodopsin structure with a mapping of residue positions taken from multiple sequence alignment of all AGPCRs and sorted in different boxes according with residue heterogeneity profiles obtained from entropy/variability (EV) analysis (Fig. 15): (A) positions of EV boxes 11 (magenta) and 12 (red); (B) positions of EV boxes 22 (gold); (C) positions of EV boxes 23 (green); (D) positions of EV boxes 33 (light blue).

This procedure is indicated for very heterogeneous series of sequences as those formed by all AGPCRs. More than 10,000 receptor sequences, the half of which consisting of olfactory receptors, were aligned and the values of variability V and entropy E for each position of the alignment were calculated and plotted against each one (Fig. 15: see legend for more technical details). In a whole, the plot shows that the positions of the AGPCR alignment could be grouped by levels of residue heterogeneity, according to different ranges of E and V values, into EV boxes 11, 12, 22, 23 and 33. Mapped in the inactive rhodopsin structure (Li *et al.*, 2004), these positions showed not a random but an oriented distribution (Fig. 16A-16D):

(1) the most conserved positions of EV boxes 11 and 12 (magenta and red in Fig. 16A, respectively) are internally located flanking the central cavity of the 7TM cytosolic half;
(2) the positions of EV box 22 (gold in Fig. 16B) with intermediate E and V values are located at internal positions all along the 7TM bundle and specially in the middle of this structure;
(3) variable positions of EV box 23 (green in Fig. 16C) are also internally located flanking the central cavity in the extracellular half of the 7TM bundle (at the retinal or agonist site); part of these positions are on the external wall of the 7TM bundle forming an interface with the membrane bilayer;
(4) the most variable positions of EV box 33 (light blue in Fig. 16D) are in both internal and external locations of the extracellular and cytosolic domains adjacent to the 7TM bundle ends. Positions mainly located at these terminal regions of the structure which failed to fulfill the rule for significance and representativity in the alignment, are shown in blue in Fig. 16A-16D. Table 4 [Appendix 15][o] shows for each of all AGPCR alignment positions, a normalized position number, V and E values, the location in the rhodopsin structure and the 5 most prevalent amino acids found.

The 7TM bundle central cavity of AGPCRs is mainly flanked by internal faces of all helices excepting helix IV. There is an internal distribution of residue positions along helices I, II, III, V and VII according to patterns of residue conservation: (a) some of the most conserved residues (EV box 11) and the most variable residues (EV boxes 23 and 33) are mainly in positions of the cytosolic and extracellular halves of these helices, respectively. Positions with intermediate residue conservation (EV boxes 12 and 22) are dispersed along the central region of helices I, III and V. The internal face of the helix VI is an exception to this rule since it contains a few conserved residues which are in the middle of its structure. Separated from the bundle itself, helix IV has most of its positions with variable residues of EV boxes 23 and 33.

4.2. Binding sites

When well-known folds such as those of globin, serine-proteases and Gα chains were analyzed by EV plots (Oliveira *et al.*, 2003a), it was found that a functional unit could be identified along the structure of these proteins with the conserved and variable positions forming, respectively, the main site (heme site for globins and catalytic sites for serine-proteases or Gα GTPases) and modulator sites (Ca^{2+} site in serine proteases, effector sites in Gα chains and sites for interfacing homologous chains in hemoglobin dimers or tetramers).

In this scenario, the main site acts as an input gate through it binding of an external ligand can start the activation of a protein functional unit. The protein class- or even type-specific modulator sites are output gates by which messages may be transmitted to or received from surrounding functional units. Various positions with intermediate E and V values are in the core of the structures enabling mechanistic pathways so that through them, in the course of activation, signals might be bi-directionally transmitted between the main site and modulator sites.

A straightforward way to understand these mechanisms is to assume that in most globular proteins the incorporation of the functional sites and of the correlate amino acids followed a divergent scheme (Fig. 17). The contact of an ancestral form of a protein fold with a ligand or with a ligand or a certain environment can induce the formation of a main binding site thus defining a class (family) of the

[o]Appendix 15: http://arbs.biblioteca.unesp.br/appendixes/appendix_15.pdf

protein. Later, along the evolution, samples of these ancestors are transported, perhaps during the phase of cell differentiation, to different compartments of the biological system where by contact with different environments and related ligands, they can enable the formation of secondary or modulator sites, each one with a certain specificity and defining a sub-class (sub-families or types) of the protein.

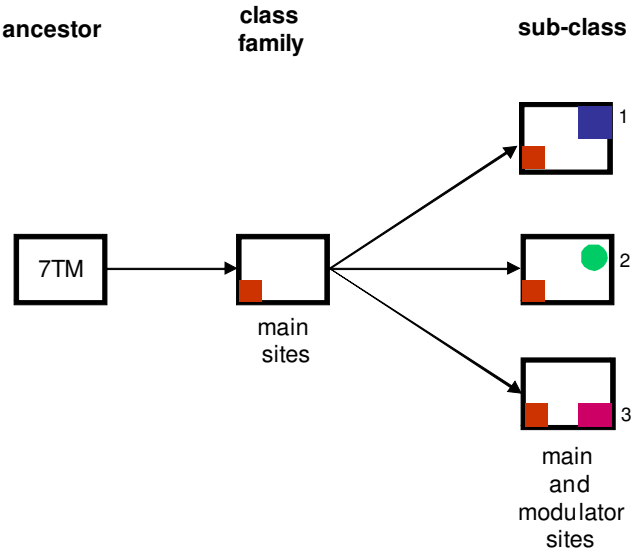

Figure 17. Divergent scheme for evolution of AGPCRs: 1st phase: incorporation of a main site into an ancestral form of these proteins thus giving rise to a family (AGPCR); 2nd phase: incorporation of different modulator sites 1-3 into initial AGPCR forms, after these molecules are transported to environments 1-3, thus giving rise to sub-classes 1-3 of these receptors.

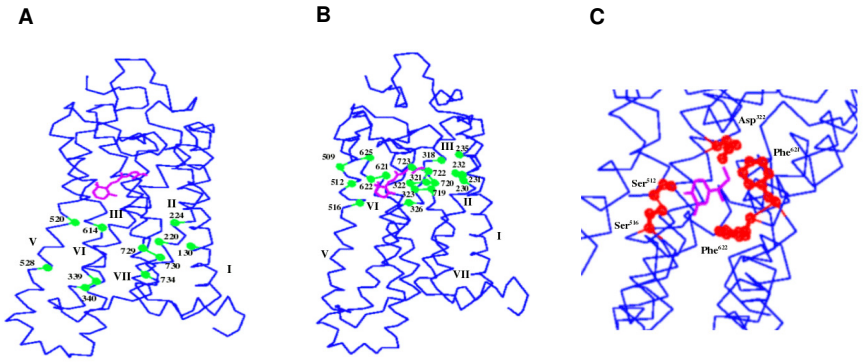

Figure 18. AGPCR binding sites shown in the structure of inactive rhodopsin (Li et al., 2004). (A) The main site comprised of EV 11, 12 and 22 box positions (green) and located at the cytosolic half of the 7TM bundle. (B) The agonist site comprised of EV 23 and 33 box positions (green) and located at the extracellular half of the 7TM bundle (retinal site: see retinal in magenta). (C) Site for binding of nor-adrenaline (magenta) in β-adrenoceptors comprising residues (red) in helices III (Asp^{322} for binding ligand amino group), V (Ser^{512} and Ser^{516} for making hydrogen bonds with ligand phenoxyl groups) and VI (aromatic rings of Phe^{621} and Phe^{622} to stabilize the ligand aromatic ring).

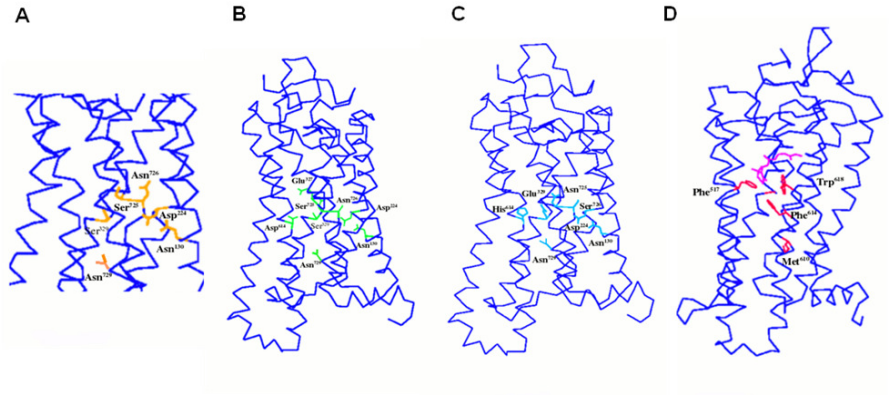

Figure 19. Structure of inactive rhodopsin (Li et al., 2004). Pathways mostly comprised of AGPCR positions occupied by residues with intermediate level of conservation (EV boxes 12, 22 and 23) and some key very conserved residues (EV box 11), all of them located in the central part of the 7TM bundle where they supposedly can transport a signal between receptor agonist and main sites. (A) Common pathway for most receptors; (B) pathway for glyco-protein hormone receptors; (C) pathway for olfactory receptors: (D) pathway for rhodopsin (retinal in magenta). See Table 4 [Appendix 1][a] for identification of positions within the EV boxes.

A similar interpretation may be formulated for results taken from EV analysis of AGPCRs (Fig. 15) (Oliveira *et al.*, 2003b) allowing to conclude that there are also a main (family-identifier) site and modulator sites in the structure of these receptors, common to all proteins and specific for their classes, respectively. Signal elicited at the main site may be carried to the modulator site and vice-versa via residues with intermediate EV values placed in between these sites. It should be noted that, in the context of receptors, the agonist-mediated activation of receptors is determined by a signal initially formed at the extracellular modulator site and further transmitted to the cytosolic main site.

The main site of AGPCRs corresponds to a common cluster of conserved polar residues such as Asn^{130}, $Asp^{224,}$ Asp^{339}, Arg^{340}, Asn^{729} and Tyr^{734} (Fig. 18A) including a part of a sodium allosteric site for some receptor classes (Neve *et al.*, 2001). Located at the cytosolic side of the 7TM bundle structure, this site would transport the signal coming from the 7TM extracellular site to cell cytosolic factors such as G-proteins, protein kinases, adaptors and others, and vice versa.

The most important modulator site, the agonist site, is at the extracellular side of the 7TM bundle central cavity and corresponds to the retinal site in rhodopsin structure (Fig. 18B and 18C). Interestingly, for AGPCRs with small agonist molecules (with the same dimensions and similar configuration of retinal), the agonist site seems to overlap entirely the rhodopsin retinal site. The same positions, or neighbor positions, in sites for different agonists seem to be functionally important for binding with a difference that they are occupied by different residues thus giving rise to the different specificities of multiple receptor classes (Fig. 18B) (Bywater, 2005). Data demonstrating this correspondence are presented (Table 5) [Appendix 16][p] whereas a scheme for a minimal site for noradrenaline in adrenoceptors, built with part of these data, is shown in Figure 18C (Trumpp-Kallmeyer *et al.*, 1992). For larger AGPCR agonists such as peptides and proteins, it has been shown that besides occupying the retinal site-like region of receptors, the agonist sites are spread to extracellular loops of the structure (Hjorth *et al.*, 1994; Lu *et al.*,1998; Gether, 2000; Fan & Hendrickson, 2005; Kristiansen, 2004; Costa Neto *et al.*, 2000; Gimpl *et al.*, 2008; Miura & Karnik, 1999; Santos *et al.*, 2008; Leeb *et al.*, 1997; Leeb *et al.*, 1997; Fathy *et al.*, 2000; Kyle, 2000; Meng *et al.*, 2000; Greenfeder *et al.*, 1999; Hoffmann *et al.*, 2000; Bhertold *et al.*, 1997; Blanpain *et al.*, 2003; Gerber *et al.*, 2001; Higginbottom *et al.*, 2005).

[p]Appendix 16: http://arbs.biblioteca.unesp.br/appendixes/appendix_16.pdf

4.3. Intermediate pathways

Analyses of EV boxes 12, 22 and even 23 (Fig. 16A-16C) allowed us to delineate the residue positions of AGPCRs which due to their location in the 7TM bundle, are involved in the formation of possible intermediate pathways for signaling between the main and the modulator sites of these receptors. As these positions have intermediate value of variability (number of different residues), the number of different signaling pathways formed by them would be limited and non-specific, that is, a same pathway would be found in different classes of receptors. Mutational studies have been able to confirm this hypothesis by identifying a few residues forming intermediate clusters in the 7TM receptor bundle of AGPCRs (Fig. 19A-19D and Table 6) [Appendix 17][q] whose mutations may impair the signaling process in different classes of these receptors as glyco-protein hormone receptors (Min *et al.*, 2000; Angelova *et al.*, 2000; 2002; Urizar *et al.*, 2005), opsins (Shieh *et al.*, 1997; Han *et al.*,1998), angiotensin II AT1 receptor (Balmforth *et al.*, 1997; Neve *et al.*, 2001; Feng *et al.*, 2005).

These results have been supporting a general hypothesis that activation of AGPCRs could be carried out by means of a few different pathways each one specifically operating in groups of receptors, irrespective of their specific agonists and cytosolic events (G-protein activation or phosphorylation) they give rise to. Despite being attractive, this idea has not been kept consistent with new knowledge about AGPCR activation mechanisms, recently learned from studies with rhodopsin, as it will be discussed below.

4.4. Emergence of modulatory site and evolution

Aiming at knowing how modulator sites were incorporated along a divergent scheme (Fig. 17), a comparative analysis of amino acid sequences in AGPCR classes was carried out. Groups of class-specific sequences with increasing levels of residue heterogeneity were multiply-aligned and submitted to EV plot analysis. As an illustration, maps of the positions occupied by the more conserved residues are shown for the following series of receptors: all-AGPCRs (family level), AGPCR_subgroup 1 (family-level with all AGPCRs except olfactory receptors), AGPCR_subgroup 2 (family level with all AGPCRs except opsins and olfactory receptors), bioamine (sub-family level), adrenoceptor (class level), α-adrenoceptor (sub-class level), α_1-adrenoceptor (type level) and α_{1c}-adrenoceptor (sub-type level) (Fig. 20A-20H).

In alignments of receptor sequences with high residue heterogeneity (all-AGPCRs and AGPCR subgroups 1 and 2, in Fig. 20A-20C, respectively), few positions occupied by highly-conserved residues are in the cytosolic halves of 7TM helices. In alignments of sequences belonging to receptors with intermediate levels of residue heterogeneity or in the center of classification tree (for instance, sub-family bioamine, class level of adrenoceptor and sub-class level of α-adrenoceptors in Fig. 20D-20F, respectively), the highly-conserved residues are more often found at the cytosolic side and started being also seen at the extracellular side of the structure. In alignments of receptor sequences with higher residue heterogeneity such as type level of α_1-adrenoceptor and chiefly sub-type level of α_{1c}-adrenoceptors (in Fig. 20G and 20H, respectively) the highly-conserved residues are found spread all over the 7TM structure.

These patterns confirm previous conclusions (Fig. 16A-16D) that there is a concerted amino acid incorporation into the 7TM structure of AGPCRs as long as the sequences of receptors are being adapted to the environment along a divergent scheme (Fig. 17). The more the receptors evolved along this adaptation, the more their positions are occupied by conserved residues, firstly at the cytosolic half of the 7TM central cavity, later at the extracellular half of this cavity and finally all over the structure. In this context, the evolution may be directly monitored when two steps corresponding to the appearance of two natural functions of receptors, agonist binding and Gα chain coupling, are considered (Watson & Arkinstall, 1994). It is known that AGPCRs gain a site for specifically binding noradrenaline (or adrenaline) at the level of adrenoceptors (Fig. 20E) which accordingly show a large density of conserved residues at the extracellular side of their 7TM bundle (region of retinal site in rhodopsin). Also, the ability to discriminate the types of Gα chains is acquired by receptors at the level of α_1-adrenoceptor (Fig. 20G) at which most of the cytosolic positions are fully conserved.

[q]Appendix 17: http://arbs.biblioteca.unesp.br/appendixes/appendix_17.pdf

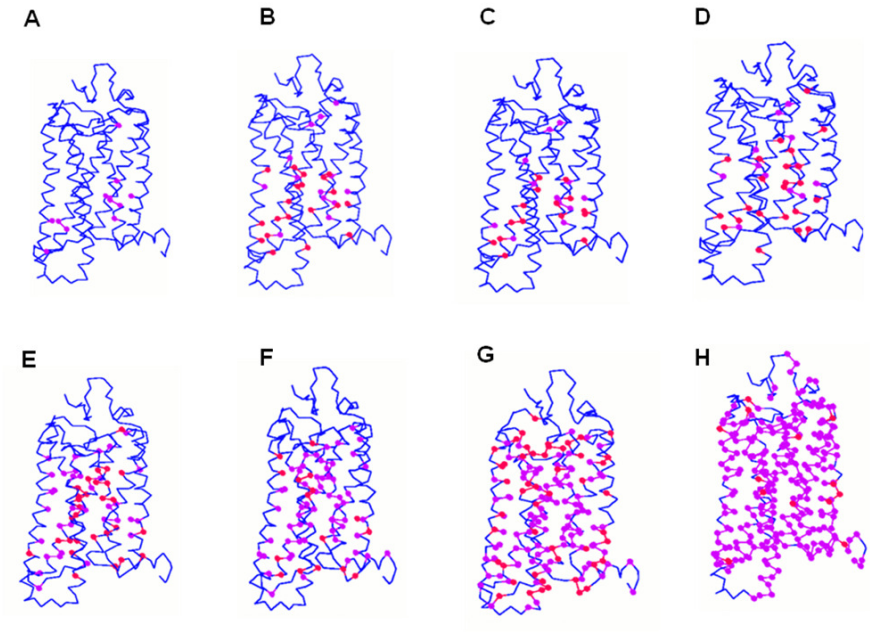

Figure 20. Mapping in the inactive structure of rhodopsin of positions bearing very conserved residues (EV boxes 11 and 12, in magenta and red, respectively) obtained from multiple alignment of sequences of (A) all-AGPCRs, (B) AGPCR_subgroup 1 (without olfactory receptors), (C) GPCR_subgroup 2 (without opsins and olfactory receptors), (D) bioamine receptors, (E) adrenoceptors, (F) α-adrenoceptors, (G) α_1-adrenoceptors and (H) α_{1c}-adrenoceptors.

However, an enormous discrepancy is apparent from these considerations. The numbers of conserved residues found in the extracellular half of adrenoceptor structure and in the cytosolic half of α_1-adrenoceptor structure are much larger than the numbers of residues experimentally identified to be required for agonist binding and Gα chain coupling functions. For instance, for the former function, this conclusion is quite clear if it is considered that a few residues are found constituting the agonist site (Fig. 18C) (Trumpp-Kallmeyer *et al.*, 1992), compared with a great deal of conserved residues revealed at the extracellular side of adrenoceptor structure (Fig. 20E). For G-protein association, there are properly no well-defined sites for binding but the same reasoning used for agonist site is plausible since only receptor limited regions (cytosolic ends) are shown to play the coupling function (Ernst *et al.*, 2000).

Thus, it has to be assumed in AGPCRs that an expressive fraction of residue positions found in the surroundings of agonist and G-protein coupling sites have no function at all or are involved in other functions than those related to agonist and G-protein coupling. This supposition would more likely be true at the last of the transitions considered in the series of receptor sequences of Fig. 20A-20H, the transition to the subtype level (α_{1c}-adrenoceptors in Fig. 20H) at which the large majority of positions of the AGPCR structure becomes fully-conserved. In this case, residue conservation might in part be related to receptor discriminative binding to non-natural antagonists which turned out to be the base for their characterization and classification into types and sub-types.

4.5. Functional unit

In general, in a mingled system of biological molecules, a functional unit of proteins is in contact with other functional units forming together a succession of messengers and effectors. For instance, in a well-known chain of events, the membrane-spanning 7TM bundle structure of AGPCRs

may be activated by a ligand coming from the external medium (first messenger) what would give rise to a signal transmitted to cytosolic G-protein system. By this way, first effectors such as phospholipase C, phospho-diesterases and adenylate- or guanylate-cyclases may be activated leading to an ultimate regulation of cytosolic concentrations of inositol triphosphate (IP3) or cyclic nucleotides (second-messengers) and consequently of many intracellular events.

As a functional unit, the AGPCR molecule consists as described above, of a main site and modulator sites possibly connected by intermediate pathways. Schemes showing binding of many ligands at AGPCR unit sites may be drawn (Fig. 21A-21C). The agonist A can bind at its specific (modulator) site thereby triggering signaling which is efficiently transferred to the main site at the other side of the structure and thus to the cytosolic second messenger level. If the signal so produced is not efficiently transduced, the ligands (I instead of A) are non-productive even displaying high-affinity binding and act as competitive antagonists by blocking the binding of true agonists to the respective site (state ^{I}R instead of ^{A}R in Fig. 21B).

Ligands D may at distance (allosterically) regulate agonist A or competitive inhibitor I bindings to the modulator agonist site, by interacting for instance with the main site surroundings at the cytosolic half of the AGPCR structure, or with other modulator sites at the extracellular loops. As the signal tranduction between the main site and the modulator sites are bi-directional, some ligands as Gα chains can couple AGPCRs thus becoming able to remotely stabilize the complex between agonist but not antagonist, and receptor agonist site ($_{D}{}^{A}R$ but not $_{D}{}^{I}R$ in Fig. 21B). As illustration, a quite instructive example of this mechanism may be demonstrated in rhodopsin system. Once the signal produced by light activation of this opsin arrives at its 7TM cytosolic ends, G proteins are coupled to them and the signal bounces back to the extracellular side of its structure. This two-way transmission is interrupted by Gα uncoupling, as it occurs upon GTP binding, thus allowing the photo-regeneration of the receptor (Ernst *et al.*, 2000). External ligands may bind modulator sites, others than the classic agonist site, in the extracellular loops of muscarinic receptors (Avlani *et al.*, 2007) and AGPCRs in general (De Amici *et al.*, 2009).

Constitutive activation of AGPCRs (R to R^{C} transition in Fig. 21B) occurs in the absence of agonist and drives a portion of the receptors to an active state R^{c}. If this state of the receptor is in the presence of agonist A or of a competitive antagonist I, the complexes $^{A}R^{C}$ or $^{I}R^{C}$ would be formed , respectively, remaining to evaluate if the complex $^{A}R^{C}$ is as active as or more active than the complex ^{A}R and the complex $^{I}R^{C}$ is as inactive as the complex ^{I}R.

Binding of inverse agonist V at a modulator site which might be the agonist site itself, can revert constitutive activation but only in complexes R^{C} and $_{D}R^{C}$ leading to the formation of $^{V}R^{C}$ and $_{D}{}^{V}R^{C}$, respectively (Fig. 21C). This hypothesis is suggested by structures of β-adrenoceptor (Cherezov *et al.*, 2007; Hanson *et al.*, 2008) and adenosine (Jaakola *et al.*, 2008) receptors, in which sites for inverse agonists seems to overlap the true agonist site. This means that receptor forms, basically constitutively activated and made inactive by means of inverse agonist binding, could not be activated again by agonist binding site would be blocked. Definition and elucidation of this condition and other parallel conditions require careful experimental analysis as it has been performed for AT1 receptors (Miura *et al*., 2006).

A plausible hypothesis about evolution of AGPCR structure formulated in order to explain all eventualities shown in Fig. 21B and 21C is that, before incorporation of any modulator site including the agonist site, with the main site probably already installed, ancestral forms of AGPCRs would exist with constitutive activation performed through the main site in contact with cell cytosol. The constitution of agonist sites would occur later when the ligands were put in contact with the receptors, leading to extinction of constitutive activity but under a special and likely advantageous condition. The receptor would be now triggered towards activation but only in the presence of the specific agonist.

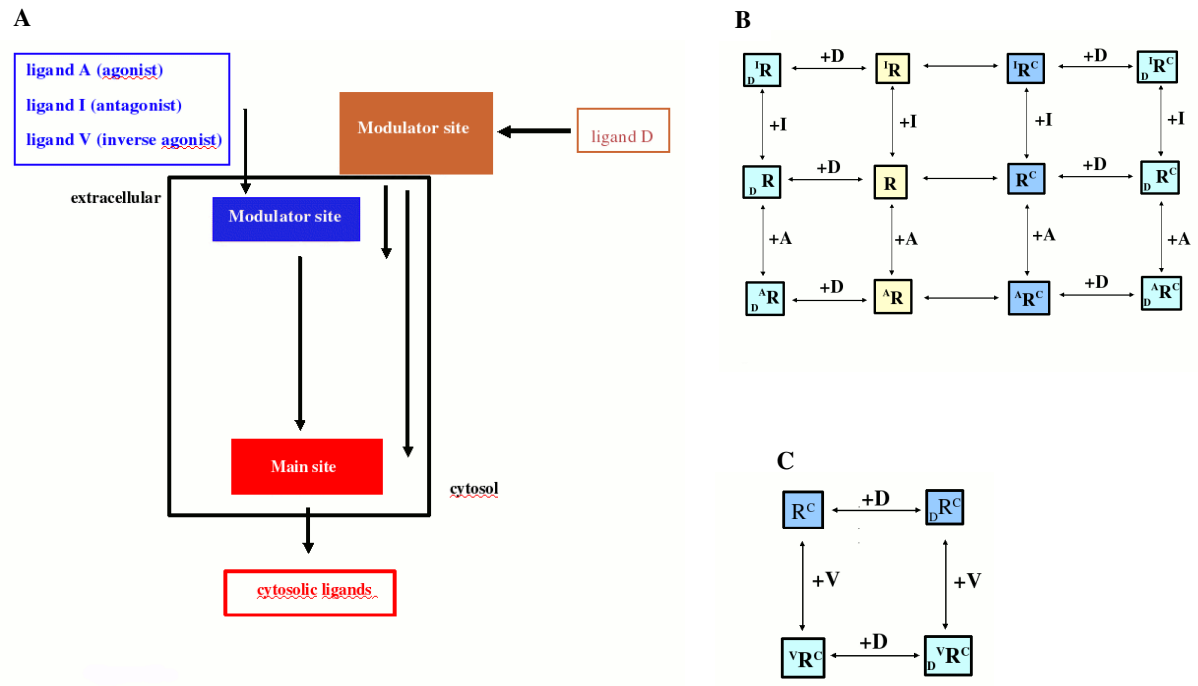

Figure 21. Multiple interaction in AGPCRs. (A) The receptor is assumed to have three binding sites, the main site (red box) with ability to interact with cytosolic factors, and two kinds of modulator sites, an agonist site (blue box) with ability to interact with an antagonist I. an agonist A or an inverse agonist V, and a putative site supposedly located at the extracellular medium (brown box) with binding specificity to a ligand D. Regulated by binding of these ligands, these sites can interfere with one another. (B) Diagram for multiple interaction in AGPCRs. The interactions of ligands with the agonist site are shown in columns: the first and third lines of the diagram contain all receptor states with agonist site bound to antagonist I and agonist A, respectively. The first two columns contain inactive receptor states in the absence of agonist (R states) whereas in the third and fourth columns the receptors are constitutively activated (R^c states). The second column states are in equilibrium with the first column D-bound states. The third column states are in equilibrium with fourth column D-bound states. (C) Constitutively activated states with free agonist sites (R^c and $_DR^c$ states) may have these sites bound to ligand V, a supposed inverse agonist, and so their activity may be eliminated.

Thus, the agonist site is a special domain specifically created at a certain moment in the receptor structure by inductive fit to the agonist molecule, leading ultimately to control of the mechanisms of activation. Interestingly, action of inverse agonists deactivating constitutive activation in AGPCRs by probably binding at the agonist site or surroundings (as suggested in receptor structures: Cherezov *et al.*, 2007; Hanson *et al.*, 2008; Jaakola *et al.*, 2008), might be similar to the effects due to the appearance of this modulator site but with a major difference. The inverse agonist binding leads to elimination of constitutive activation but without recovering the site capability to bind its specific ligand thus regaining control of its own function of activation. This is an additional obscure point in the scheme of receptor activation mechanisms that requires further experimentation for elucidation.

4.6. Conservation patterns and 3D structure

The expressive number of different amino acid sequences and 3D structures of proteins available have allowed along the last decade a great deal of more extensive and comprehensive studies about these macromolecules (Caetano-Anollés *et al.*, 2009). The most important consequence of this new era is that more realistic conclusions have been emerging about sequence-structure relationships. Intriguing observations showing that two different sequences may share a same fold, started to be found so often that old definitions stating that amino acid sequence dictates the structure (Anfinsen, 1973), had to be totally reformulated. In fact, the proteins are now characterized as homologs when they have a same

fold and belong to a same family (contain very conserved residues at certain positions) or analogs when they have a same fold but different sequences thus belonging to different families (Russell *et al.*, 1998a). In consonance with these findings, there is nowadays the consensus that 3D structures of proteins are more conserved than sequence of amino acids and thereby new interpretations for these relationships have to be found (Holm & Sander, 1996; 1997; Grishin, 1997, 2001).

It would be too drastic to consider that amino acid sequences have no relation at all with the type of fold which a protein acquires for instance when left in aqueous environment. Under this condition, the protein would be folded upon an ideal energetic minimum with most hydrophobic residue side-chains placed in the more internal positions of the structure and polar and charged residue side-chains in contact with the solvent. In case the solvent is changed in the external medium, it is expected that the distribution of residues would be changed accordingly because this is a trend which is followed by all molecules upon evolution.

However, there are two conditions in which amino acid sequences and structure in proteins may certainly be related.

(1) The relative position of less voluminous residue side-chains in the amino acid sequences is a factor that would specify the type of fold which a certain sequence may acquire. This factor is based on the principle that the small residues are required in the core of the structure in order to sterically allow the fold itself. But bulky residue side-chains are also required for folding of proteins at the boundaries of the structure core to block the entry of solvent thus implying that the small residue side-chains cannot be prevalent. Thus, the number of different types of folds which depends on the number of small chains and on their positions in the sequences, would have to be small compared to the huge numbers of sequences (Russell *et al.*, 1998a; 1998b)

(2) The presence of conserved (or functionally-conserved) positions in certain positions of the fold is also a vital factor to determine the structure. By aligning amino acid sequences of homologs, according to their positions with conserved residues, and superposing their respective 3D structures, one may observe that perfect correspondence of structure is more systematically found at domains containing one or more conserved residues. Also, no insertion of residues seem to be accepted at these regions (Russell & Ponting, 1998). As reason for these relationships, it has been postulated that conserved residues may be usually performing side-chain interactions thereby stabilizing the local secondary and tertiary structures. In analogs the sequences change but the structures are kept, a condition that seems to be linked to the fact that different but functionally-conserved residues (all with ability to perform some kind of side-chain interactions) are found in the different domains of the proteins all of them with ability to equally stabilize the structures (Russell *et al.*, 1998a; Mirny & Shakhnovich, 1999).

As referred above, there are nowadays some high-resolution 3D structures of AGPCRs, including rhodopsins (inactive and retinal free opsin),β-adrenoceptor and adenosine receptors, but two years ago, when only the structure of inactive (dark) rhodopsin was known, debates were common aiming at knowing if the general structure of AGPCRs followed the opsin fold with a complete 7TM bundle configuration.

Today it is known that α-adrenoceptor and adenosine receptors form an opsin-like 7TM bundle but displaying some structural disagreements with rhodopsin in the extracellular halves of some helices and loops. Taking advantage of this previous knowledge, it is now opportune to make predictions of how the still-undetermined structure of AGPCRs would be as a function of the patterns of residue conservation obtained from alignment of all receptors. Fig. 22A and 22B show that the positions with high residue conservation (EV boxes 11, 12 and 22) are in the cytosolic half of the 7TM bundle, mostly at the internal faces of helices in contact with the central cavity.

Considering the relationships between structure and residue conservation discussed in previous paragraphs, this finding strongly suggests that the 7TM configuration of the cytosolic half of the 7TM bundle central cavity should be kept for most receptors. On the other hand, the extracellular half of the bundle (region of agonist site) and the extra-membrane loops are likely to have their structure changed according to the agonist structure and environmental factors. In the case of helices, it is important to emphasize that in the all-AGPCR alignment the helices IV and VI also showed variable sequences along the full extension of their structures thus announcing beforehand that they might also have variable structures in many receptor classes. In this case, variation of structures would mostly mean interruption of the helix by formation of bulges, and change of its relative location in the 7TM bundle.

In the case of extracellular loops, there would be restrictions to changes in the structure of EC-2 loop of many AGPCRs, due to the presence of a very conserved disulfide bond linking the top of helix III to the loop (Fig. 1A and 3A). Sequence insertions are the most notable variations found in the loops of AGPCRs that are consequence of inductive fit mechanisms driven by ligands and other environmental factors on the structure of these receptors.

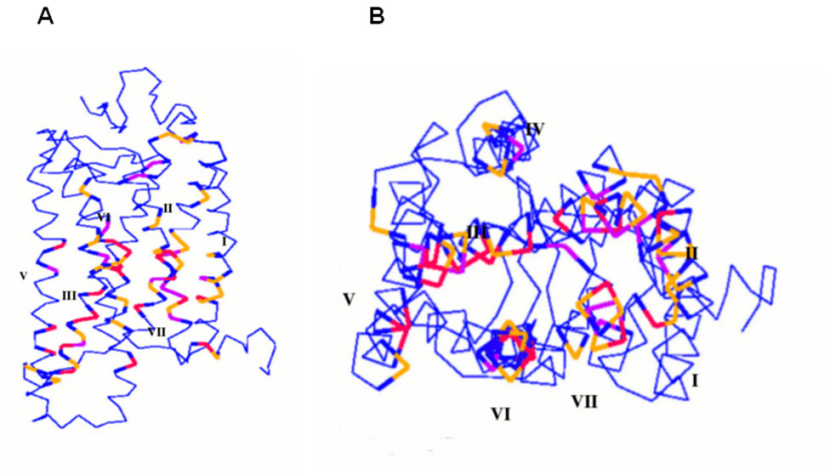

Figure 22. Inactive structure of AGPCRs. Mapping of positions occupied by very conserved residues of EV box 11 (magenta), 12 (red) and 22 (gold): (A) longitudinal view and (B) cross-section of the same structure.

5. Activation of AGPCRs

5.1. General concepts

Activation of AGPCRs is promoted by modifications on the structure of receptors starting at the extracellular half, propagating across the 7TM bundle and ending at the cytosolic ends. For most receptors this physiological process is transitory since it is elicited by agonist binding at the specific modulator site in the extracellular half of the 7TM bundle but can be started by light stimulation as in opsins. These AGPCRs have a special prosthetic group, the retinal, previously docked at the agonist site under an isomeric cis form which can undergo a light-modulated cis-trans-isomerization and thus start the activation mechanism.

Activation of AGPCRs is followed by the transference of a message to cytosolic effectors what triggers a chain of events leading to a ultimate effect. The two most studied pathways of events driven by AGPCRs are activation of G protein chains of different types and all effectors linked to them, and receptor phosphorylation by protein kinases, followed by binding of adaptors as arrestins and others, and cellular internalization. These pathways seem to be interdependent and somewhat interconvertible (Maudsley *et al.*, 2005; Perez & Karnik, 2005; Wei *et al.*, 2003) sometimes requiring the simultaneous participation of all receptor cytosolic ends (Yu *et al.*, 2002; Natochin *et al.*, 2003). They may still be specifically selected by factors inherent to the structure of agonists (Miura *et al.*, 1999; Bartl *et al.*,

2005; Le *et al.*, 2002) and non-specifically by environmental factors (Maudsley *et al.*, 2005).

The agonist-dependent selectivity of effects derived from AGPCR activation has nowadays been regularly studied in opsins and a few other systems of receptors. Indeed, syntheses of retinal derivatives obtained by chemical modifications at different positions of the original moiety have been succeeded in order to produce agonists with ability to change the equilibrium between the states that form the pathways of opsin activation (Bartl *et al.*, 2005; Lüdeke *et al.*, 2005; Vogel *et al.*, 2005; Bhattacharya *et al.*, 2008). In the system of AT1 receptor and the octapeptide-agonist angiotensin II (Asp^{1}-Arg^{2}-Val^{3}-Tyr^{4}-Ile^{5}-His^{6}-Pro^{7}-Phe^{8}), activation by native peptide binding can trigger activation of G_q machinery, a process that is regulated by agonist residues Tyr^{4} and Phe^{8} (Miura *et al.*, 1999). On the other hand, binding of Ala4- or Ala8-angiotensin II can lead receptor to successive events such as phosphorylation by protein kinases, binding to arrestin and internalization (Feng *et al.*, 2005; Kalatskaya *et al.*, 2004; Perez and Karnik, 2005; Maudsley *et al.*, 2005; Wei *et al.*, 2003).

AGPCRs may also be constitutively activated (absence of agonist) by permanent modifications of their own structures or other receptor-coupled proteins, usually produced by residue mutations. Constitutive activation usually denotes a condition in which the receptor is able to activate Gα chains but it is also seen driving the phosphorylation pathway. It can be elicited by changes in the receptor structure or surrounding environment along the entire pathway of activation (Pauwels & Wurch, 1998).

For instance in a pathway leading to phosphorylation, mutations of residues in the receptor cytosolic ends or in the protein kinase structure leading to reduction of the catalysis rate, may increase the levels of receptors available thus producing relative increase of constitutive Gα chain activity by a mechanism at distance.

In the region between the agonist site and cytosolic ends, signal transduction is performed leading receptors to be stabilized under an activated state. If mutations are made in this region and higher stabilization of this state is reached, a condition of constitutive activation may be installed now through a mechanism directed to the signaling mechanism itself.

Analysis of the mechanisms leading to activation of AGPCRs is nowadays extremely facilitated due to the great deal of kinetic and structural information about rhodopsin activation obtained from spectroscopic assays, crystallization and determination of 3D structures (Schertler, 2005). The extension with which this knowledge can be applied to other receptor classes is measured in maps of residue conservation built from multiple alignment of all-receptor sequences on the structure of rhodopsin. These maps show that residue conservation is in the internal positions of the structure cytosolic half and heterogeneity is seen in the extracellular side. Thus, activation occurring at the extracellular half of the structure 7TM bundle would have to be mediated by receptor-class-specific mechanisms whereas, on the other hand, it would have a same mechanism for all receptors when occurring at the cytosolic half of the structure.

5.2. Activation of rhodopsin

5.2.1. Activation may be monitored from rhodopsin and retinal-free opsin structures

The first determined structure of rhodopsin (dark opsin) is an inactive form of AGPCRs (Palczewski *et al.*, 2000; Li *et al.*, 2004) but the retinal-free opsin (Park *et al.*, 2008) is constitutively activated when helix VII $Lys^{723(296)}$ (Fig. 23A-23C; see Fig. 2 for numbers of positions) is mutated to negatively-charged or neutral residues (Glu or Gly) (Zukhovsky *et al.*, 1991; Li *et al.*, 1995; Rim & Oprian, 1995; Shi *et al.*, 1998]. Thus, it might a priori mimic an active structure of these receptors.

Retinal-free opsin with free $Lys^{723(296)}$ (Fig. 23A) is inactive but can recover activity when incubated and regenerated with retinal (Bhattacharya *et al.*, 1992). Lys723(296)Glu(Gly) mutant of opsin is constitutively activated and may acquire physiological activity in the presence of light when regenerated with retinal (Zhukovsky *et al.*, 1991). The existence of constitutive activity in mutated but not in wild forms of retinal-free opsin, might be associated with a simple mechanism. A free Lys at the helix VII 723(296) position of wild retinal-free opsin would simply make a salt bond with the γ-carboxyl of helix III $Glu^{318(113)}$ (Fig. 23A) or EC2 loop $Glu^{464(181)}$ (Fig. 23B) whereas upon Lys723(296)Glu mutation

(Fig. 23C) this occurrence is ruled out.

Interestingly, in agreement with this hypothesis, Gly231(90)Asp or Ala719(292)Glu mutations drive rhodopsin to constitutive activity (Rao *et al.*, 1994; Rim & Oprian, 1995) perhaps because the introduction of negative side-chains at these positions leads to an effect similar to that due to $Glu^{723(296)}$ (Fig. 23D). The bond between $Lys^{723(296)}$ and $Glu^{318(113)}$ or $Glu^{464(181)}$ would link the extracellular half of helix VII to the top of helix III or to the EC2 loop thus keeping compact these domains of the opsin which otherwise would be loose. It seems that a somewhat disordered structure in the extracellular half of 7TM bundle of opsins favors the activation of these proteins. Owing to this fact, the retinal-free opsin structure (Park *et al.*, 2008) bearing a Lys723(296)Glu mutation was used as a template for studying activation of rhodopsin.

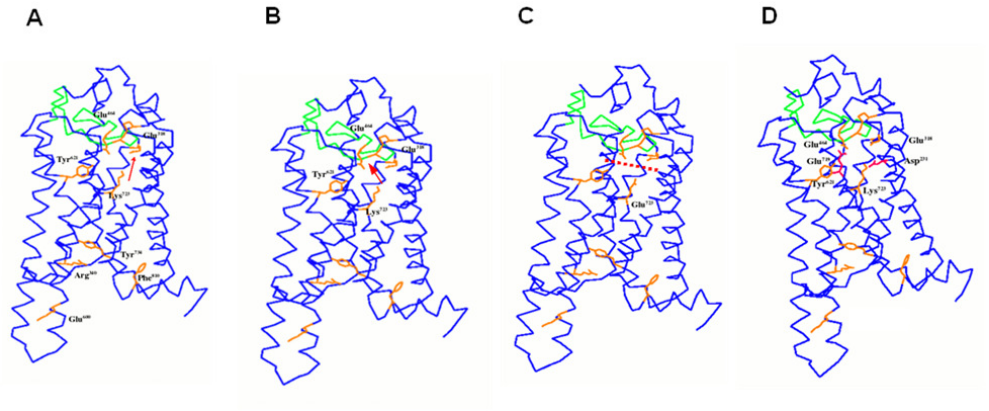

Figure 23. Structures of retinal-free opsin showing the EC2 loop hairpin in green. Note the disrupted interactions in the cytosolic ends of the structure between the pairs of residues, $Arg^{340(135)}$-$Glu^{600(247)}$ and Tyr734(306)-$Phe^{810(313)}$, indicating an activated form of protein. (A) Despite it, an inactive-like form of the protein may be assumed in the interface between 7TM loop and EC2 loop, with interactions between helix VI $Tyr^{621(268)}$ and EC2 loop $Glu^{464(181)}$, and between helix VII $Lys^{723(296)}$ and helix III $Glu^{318(113)}$ (see red arrow). (B) A metarhodopsin I-like form of opsin would be formed by making $Lys^{723(296)}$ interact with EC2 loop $Glu^{464(181)}$(see red arrow), thus releasing the side-chains of $Tyr^{621(268)}$ and $Glu^{318(113)}$. (C) The Lys723(296)Glu mutation would break the previous bond by a charge repulsion effect (red dotted line). (D) Gly231(90)Asp and Ala(719)292) mutations (see residues in red) may promote a similar effect, including due to the formation of a bond between $Lys^{723(296)}$ and $Asp^{231(90}$ residues.

Inactive rhodopsin was compared to retinal-free opsin by superposing the structures of these proteins and measuring the distances between their corresponding main chain atoms (N, C, O, Cα). A special preliminary requisite was that the superpositions had different initial templates, that is, the procedures would have to be centered not only on the whole opsin structure but also on different 7TM bundle domains (helices), separately. This protocol is likely to identify whether the full structures of opsins or parts of them are similar in a dependence on the templates used. For instance, (1) if the structures of inactive rhodopsin and retinal-free opsin are on the whole very different, procedures based on small regions of the proteins, as isolated 7TM helices, may eventually produce good superpositions but limited to the regions of templates; (2) if the inactive and active opsins have similar overall structures, superpositions are perfect even using templates spanning all over the molecules; (3) if parts of the opsin structures are similar, only superpositions based on templates of these regions are succeeded.

The hypothesis of item 3 above was found to be true as indicated experimentally. In fact, by superpositions of inactive rhodopsin to retinal-free opsin were satisfactory only when helices III, IV and V were used together, or in pairs, as templates. Superpositions centered on helices I, II, VI and VII led at sparse results. When isolated helices were used as templates, satisfactory superpositions were only found for the respective helix. However, when helices I, VI and VII were made templates together,

the superpositions of these three helices plus that of helix II were better but the superpositions of helices III, IV and V were now entirely unmade.

These studies enabled us to split the opsin 7TM bundle into two continuous blocks, the first consisting of helices III, IV and V (group 1) and the second of helices I, II, VI and VII (group 2) (Fig. 24A and 24B). A plausible interpretation of these results is that both inactive rhodopsin and retinal-free opsin have a conserved conformation in their 7TM bundle block formed by group 1 helices but different conformations in the structure block formed by group 2 helices. Thus, if inactive rhodopsin and retinal-free opsin are, respectively, the initial and final states of opsin activation, the conformational changes occurring along the steps of this mechanism, would rather occur in the group 2 block of the 7TM bundle structure (blue in Fig. 24B).

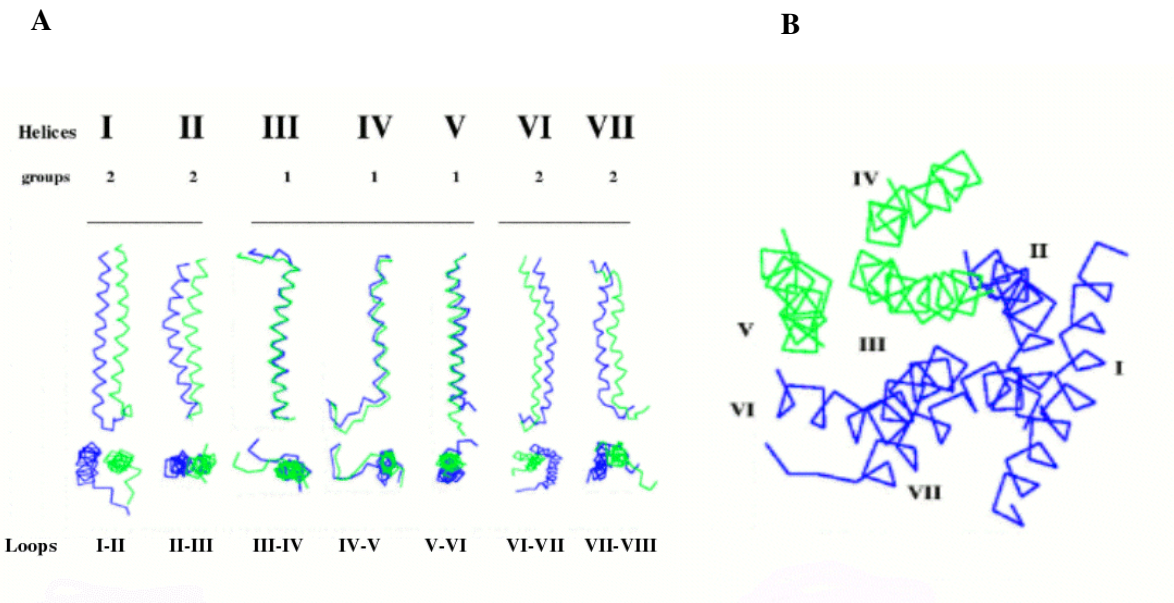

Figure 24. (A) Cα traces taken from superposed retinal-free (green) and inactive (blue) opsins according to a procedure centered in helices of group 1 (III, IV and V), Note that superposition for helices of group 2 (I, II, VI and VII) is imperfect all over the extension of helices. (B) Cross-section of the 7TM bundle of retinal-free opsin showing helices of groups1 (blue) and 2 (green).

5.2.2. Conformational changes on retinal-free opsin relative to inactive opsin

Cα traces (Fig. 24A) shows that the conformational changes supposedly linked to activation of opsins, are lateral displacements of the entire extension of group 2 helices (I, II, VI and VII) and some adjoining loops of the retinal-free opsin in relation to the corresponding domains of the inactive rhodopsin, while group 1 helices (III, IV and V) and adjoining loops, including the N-terminal half of loop V-VI, are less moved or are practically kept still.

Considering the 7TM bundle structure shown in Fig. 24B, one may figure out that, upon the postulated activation, half of the inactive rhodopsin 7TM bundle comprising group 2 helices and some intermediate loops, are laterally moved specially at the cytosolic ends (Fig. 25). For instance, direct measurements in superposed structures of inactive rhodopsin and retinal-free opsin (Fig. 24A, 24B and 25) show that the averaged shifts are of 7-9 A in the C-terminal segment of loop V-VI and N-terminal segment of helix VI; 4.5 A in the cytosolic ends of VII, VIII and loops and of 3 A in the cytosolic end of helix II. Thus, along an evolution going from inactive to active forms of opsin (retinal-free structure), expressive expansion of the 7TM bundle is rather produced in the cytosolic ends of helix VI, in the interface between the cytosolic ends of helices VI and VII and between the cytosolic end of helix VII and the helix VIII.

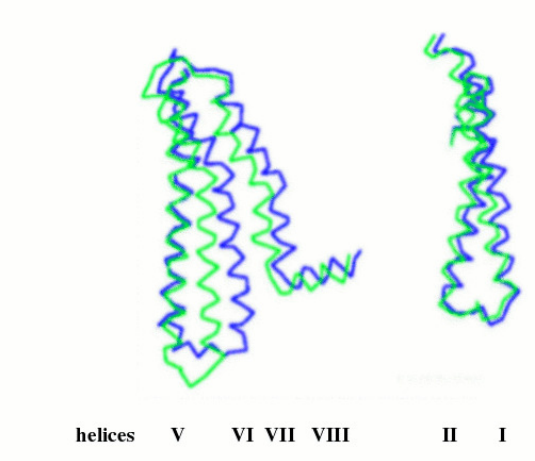

Figure 25. Cα traces taken from superposed retinal-free (green) and inactive (blue) opsins. Side view of the 7TM bundle structure showing helices of group 2: I , II, VI, VII and VIII plus helix V as reference. In the hypothesis now under consideration, blue and green stand for inactive and activated forms of opsins, respectively. Thus, the lacking of superposition indicates movemtns of helices upon activation. Helices V perfectly superposed are the references.

5.2.3. Features of activation in rhodopsins

Opsin conformational changes leading to activation may be spectroscopically monitored since along this process the retinal has maximal values and wavelenghts of absorption changed. For instance, products observed in the process of rhodopsin light-mediated activation have maximal absorptions at 500 nm (rhodopsin itself), 480 nm (metarhodopsin I MI) and 380 nm (metarhodopsin II MII) (Matthews *et al.*, 1963). As a consequence of this experimental advantage, studies of rhodopsin activation have attained an advanced stage as denoted by the detailed description available for the mechanism of activation of this protein (Knierim *et al.*, 2007). Here, a summary of this mechanism is reproduced as follows.

In the inactive structure of rhodopsin, the retinal site is stabilized by bonds between a protonated SB and helix III $Glu^{318(113)}$ (Kim *et al.*, 2004), and between loop IV-V hairpin $Glu^{464(181)}$ and helix VI $Tyr^{621(268)}$ (Li *et al.*, 2004) (bonds 1 and 2 in Fig. 26A, respectively). Following retinal light-mediated 11-cis-all-trans isomerization, some short lifetime opsin states evolve in the system, before rhodopsin MI state is formed upon breaking of the two bridges above (Yan *et al.*, 2002; Ruprecht *et al.*, 2004; Schertler, 2005). As a consequence of this change, the retinal SB link has its position changed so that it is now stabilized by its interaction with loop IV-V $Glu^{464(181)}$ (indicated by arrow in Fig. 26B) being the helix VI released from connection with the EC2 loop IV-V hairpin.

Large changes in the environment around the SB drive a transition of rhodopsin structure from a MI state to a MII state of rhodopsin. Kinetical studies (Knierim *et al.*, 2007) have revealed the precise order of events in this process. It starts with SB deprotonation in which SB pK is seen decaying step-wisely from about 16.0 to 2.5 and follows with a motion of helix VI all over the 7TM bundle structure. Finally, it ends with a proton uptake by $Glu^{339(134)}$ residue in the cytosolic end of helix III due to a pK decay linked to the helix motion and breakage of previous interactions between cytosolic ends (compare the state of these interactions in Fig. 26B and 26C). Metarhodopsin MII is the active state of rhodopsin which is believed to couple G-proteins (Hamm *et al.*, 1988; Kisselev *et al.*, 1999: Ernst *et al.*, 2000) and be target for phosphorylation (Maeda *et al.*, 2003). Whether a same form of MII is related or not to these two events is matter for discussion in next sections.

The mechanism described for rhodopsin is believed to be common for activation of all opsins and thus the bridge involving $Glu^{464(181)}$ and protonated SB cannot be so relevant to regulate MI/ MII equilibrium as described for rhodopsin since the residue in question is not fully conserved along the sequences of other opsins (Ludeke *et al.*, 2005). In red or green visual pigments, $Glu^{464(181)}$ is mutated to $His^{464(181)}$ and is part of a chloride site that, via binding of this ion, can discriminate colour by regulating the light wavelenght absorbed by the opsin upon activation (Wang *et al.*, 1993).

These additional facts allow to conclude that the mechanism of opsin activation is more general. Formation of MII state, involving previous SB deprotonation, has a high activation energy and thus it is a process that requires broad conformational changes supposedly occurring at the retinal pocket and other regions of the 7TM bundle. These changes probably result from the isomerization itself. Results from Nuclear Magnetic Resonance studies (Ahuja *et al.*, 2009) reveal that upon light activation the extracellular third of helix V and the EC2 loop are substantially moved. Also, the all-trans-retinal is inserted into the 7TM bundle cavity; its β-ionone ring reaches the cytosolic half of the structure (Ahuja *et al.*, 2009) and thus might be a vital factor to driving the motion of helices. Anyway, conformational changes thus produced are transmitted to the other side of the 7TM bundle leading to extensive displacements of the structure cytosolic ends (Altenbach *et al.*, 2008).

5.2.4. Structure of retinal-free opsin is consistent with the mechanism of rhodopsin activation

Two features of retinal-free opsin structure may be perfectly matched to events of rhodopsin activation. First, the helix VII $Lys^{723(296)}$ side-chain (free in the absence of retinal) is found shifted towards the IV-V loop $Glu^{464(181)}$ residue due to the lateral movement of group 2 helices (I, VI and VII) (Fig. 24A and 25) thus mimicking the exchange of salt bridges observed with SB in rhodopsin light-mediated activation. Second, the space amid cytosolic ends is significantly enlarged in retinal-free opsin as denoted by increases of $Arg^{340(135)}$-$Glu^{600(247)}$ and $Tyr^{734(306)}$-$Phe^{810(313)}$ distances in relation to inactive rhodopsin, from 8.7 A to 14.7 A and from 7.4 A to 10.4 A, respectively (compare distances of corresponding residues in the cytosolic ends of inactive rhodopsin and retinal-free opsin in the bottom of Fig. 26A and 26B, respectively). Thus, many conditions that would be expected as products of light-activation of opsins are right developed upon simple withdrawal of retinal.

Given the higher conservation of residues in the cytosolic regions of the AGPCR 7TM bundles, the expansion of structure at these regions may be a common feature of most of the receptor classes (Oliveira *et al.*, 2007). In fact, the stability of the $Arg^{340(135)}$-$Glu^{600(247)}$ salt bridge between the cytosolic ends of helices III and VI (Ballesteros *et al.*, 2001), and even of the $Tyr^{734(306)}$-$Phe^{810(313)}$ aromatic-ring interaction between the cytosolic end of helix VII and helix VIII (Fritze *et al.*, 2003), are crucial for keeping the AGPCRs inactivated (Fig. 3A, 26A and 26B).

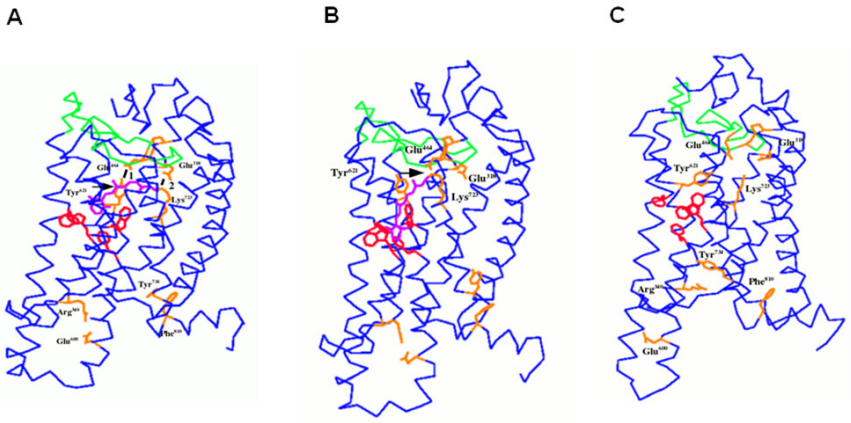

Figure 26. Structures of rhodopsin in which retinal is drawn in magenta, the EC2 loop hairpin in green, residues bound to retinal or making interactions between cytosolic ends in gold, aromatic side-chain interacting with retinal in red. (A) Inactive form (Li et al., 2004) showing interactions between EC2 loop $Glu^{464(181)}$ and helix VI $Tyr^{621(268)}$ (interaction1) and between the cis-retinal Schiff base and helix III $Glu^{318(113)}$(interaction 2). Cytosolic interactions (below) between helix III $Lys^{340(135)}$ and helix VI $Glu^{247(600)}$ orbetween helix VII-helix VIII bend $Tyr^{306(734)}$ and helix VIII $Phe^{313(810)}$ are intact. (B) A hypothetical form of rhodopsin created by pulling the 7TM bundle helices away from each other (Oliveira et al., 2007). Due to retinal displacement produced by isomerization, the interaction 1 of previous figure was disrupted and the Schiff base is now bound to EC2 loop $Glu^{464(181)}$ (see arrow). Cytosolic interactions (below) are now disrupted. (C) Retinal-free opsin (Park et al., 2008). The free $Lys^{723(296)}$ is making salt-bridge with EC2 loop $Glu^{464(181)}$. Cytosolic interactions (bottom of the figure) are also disrupted.

5.2.5. A same state or different states of activated rhodopsin may drive G-protein activation and phosphorylation

The cytosolic loops IC2 and IC3 of AGPCRs contain sites for coupling cytosolic proteins and are certainly regions where extensive conformational changes occur in the course of the activation (Hamm *et al.*, 1988; Altenbach *et al.*, 1996; Ernst *et al.*, 2000; Ballesteros *et al.*, 2001; Ridge *et al.*, 2006). The IC3 loop is found already accessible in inactive rhodopsin and its accessibility is significantly increased in retinal-free opsin (Fig. 25) thus reinforcing the idea that it is vital in signaling mechanisms.

Two pathways are described driving interactions of AGPCR cytosolic ends with cytosolic proteins, the first leading to activation of G-protein system (release of GDP from Gα chains of types t,s,q and others and separation of the respective Gβγ complexes) and the second one enabling the protein kinase attack mainly to the C-terminal tail of receptor with consequent creation of sites for high-affinity binding of adaptors as arrestins (Hofmann *et al.*, 1992; Pulvermüller *et al.*, 1993; Han *et al.*, 2001; Luttrell & Lefkowitz, 2002: Kisselev *et al.*, 2004), and finally for receptor internalization (Nabi & Le, 2003; Mousavi *et al.*, 2004).

A problem relative to these interactions that remains to be solved is whether a same conformational state of activated AGPCRs, or different states of these proteins, would have to be formed in order to give rise to G-protein activation and receptor phosphorylation. In case of rhodopsin, it has been demonstrated from spectroscopic studies that MII is the form leading to activation of Gαôchain (Hamm *et al.*, 1988; Kisselev *et al.*, 1999; Ernst *et al.*, 2000; Herrmann *et al.*, 2004; Knierim *et al.*, 2007; Scheerer *et al.*, 2008) and other forms exist with ability to couple G protein but with different efficiencies in order to promote GDP release (Morizumi *et al.*, 2005). However, if this problem is thus well tested and documented at the side of G protein activation events, it is, on the hand, less visited at the phosphorylation side.

Cell experiments planned to study activation-mediated phosphorylation of rhodopsin were at a first stage unstable so that contradictory results were initially obtained showing either that a single MII-like conformation is driving both G protein activation and phosphorylation pathways (Bhattacharya *et al.*, 1992; Li et al, 1995; Rim & Oprian, 1995) or the reverse condition on which different rhodopsin states are needed for both of these post-activation events (Robinson *et al.*, 1994). This last result coincides with results from specific experiments carried out with other AGPCRs (Perez & Karnik, 2005) and specially angiotensin II AT1 receptors (Thomas *et al.*, 2000).

It is difficult by means of initially rudimentary experiments to check the validity of the two hypotheses but more precise spectroscopic techniques applied to solubilized rhodopsin have allowed important advances in these studies. In the process of MII formation, a previous finding showing that there is an equilibrium between two forms of activated rhodopsin, called MII_a and MII_b (Arnis & Hofmann, 1993), could be more recently analyzed with more details (Herrmann *et al.*, 2004). The opsin MII_a has SB already deprotonated but not helix VI moved and proton associated by $Glu^{339(134)}$at the bottom of helix III. The opsin MII_b has the motion of helix VI and and proton uptake processes completed (Ernst *et al.*, 2007). As it has been verified that helix VI motion and proton uptake are required for Gα coupling to the opsin (Arnis *et al.*, 1994), one may conclude that MII_b is the more likely form of activated opsin related to G-protein activation. But a crucial question still remains. Would the MII_a opsin the specialized form of this protein with ability of triggering the phosphorylation pathway?

A way to address this problem is to focus attention on the broad structural changes occurring at the cytosolic ends of AGPCR 7TM bundle upon agonist-mediated activation. The analysis of retinal-free opsin structure (Parker *et al.*, 2008) as a prototype for opsin activation, reinforced previous assumptions (Oliveira *et al.*, 2007) that activated AGPCRs have their cytosolic ends moved away from each other, with notable displacements of helices VI and VII that ultimately lead to breakage of interactions between helix III $Arg^{340(135)}$ and helix VI $Glu^{600(247)}$ (Ballesteros *et al.*, 2001), and helix VII-helix VIII bend $Tyr^{734(306)}$and helix VIII $Phe^{810(313)}$ (Fritze *et al.*, 2003) (See Fig. 26A and 26C).

Thus, it is plausible to assume that upon rhodopsin activation, the breakage of interactions between helices III and VI (first interaction above) might exactly express the helix VI motion step in the mechanism of MII formation, that ultimately drives the system to the G-protein activation pathway.

When, for reasons not analyzed yet, the process of conformational changes linked to rhodopsin activation is oriented towards the second interaction between helix VII and the 7TM bundle itself (second interaction above), the phosphorylation pathway would be favored. Interestingly, there are indications that the breakage of the interaction between $Tyr^{734(306)}$and helix VIII $Phe^{810(313)}$ can produce a realignment of helix VIII, relative to the 7TM bundle, that is deleterious for the mechanism of G-protein activation (Fritze *et al.*, 2003) but is extremely favorable for phosphorylation as long as it might release the C-terminal tail of receptors for attack by protein kinases.

5.2.6. Structural modifications of retinal or agonists in general select post-receptor-activation pathways

The selection of G-protein activation and phosphorylation pathways based on the side of the 7TM structure that is more modified by conformational changes upon activation, is quite attractive. More interesting is the possibility that this way of selecting post-receptor-activation events might be dictated early by agonist binding to the specific site at the extracellular side of the 7TM bundle structure. A great deal of quite reliable information consistently point to this hypothesis.

In case of rhodopsin, if chemical changes are made in the retinal structure, mainly in its ionone ring, a condition of partial agonist emerges characterizing upon light activation a shift of MI/MII equilibrium towards the inactive MI (Bart *et al.*, 2005; Vogel *et al.*, 2005). No data are available yet regarding the status of the phosphorylation pathway elicited upon light activation of rhodopsin docked with modified retinal molecules.

In other AGPCRs for which the activation is mediated by larger molecules as peptides and proteins, it has been verified that the total agonistic activity results from the sum of selective effects determined by binding of specific regions of the agonist to receptor. For instance, as it has been already referred to above, in AT1 receptors for angiotensin II Gaq chain activation is due to binding of Tyr4 and Phe8 side-chains (Miura *et al.*, 1999), a condition that is not required for activating phosphorylation (Thomas *et al.*, 2000). Analogs or homologs of angiotensin II, devoid of aromatic rings due to selective mutation or remotion, are only able to trigger phosphorylation and internalization pathways (Le *et al.*, 2000; Feng *et al.*, 2005). By applying this reasoning to the rhodopsin system, one may argue that the region of retinal ionone ring might be a probable candidate to elicit activation of Gαô chain whereas parts of retinal which might be linked to increment of opsin phosphorylation remain to be determined.

Unfortunately, to the best of our knowledge, no precise information about relationships between structures of agonist and efficiency to trigger activation of G-protein or phosphorylation, have been reported for other systems of AGPCRs. Thus, a more comprehensive discussion of this subject will have to wait for the conclusion of specific experiments.

6. References

Ahuja S, Crocker E, Eilers M, Hornak V, Hirshfeld A, Ziliox M, Syrett N, Reeves PJ, Khorana HG, Sheves M, Smith SO. Location of the retinalchromophore in the activated state of rhodopsin. J Biol Chem 2009;284:10190-201.

Altenbach C, Yang K, Farrens DL, Farahbakhsh ZT, Khorana HG, Hubbell WL. Structural features and light-dependent changes in the cytoplasmic interhelical E-F loop region of rhodopsin: a site-directed spin-labeling study. Biochemistry 1996;35:12470-8.

Altenbach C, Kusnetzow AK, Ernst OP, Hofmann KP, Hubbell WL. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. Proc Natl Acad Sci USA 2008;105:7439-44.

Anfinsen CB. Principles that govern the folding of protein chains. Science 1973;181:223-30.

Angelova K, Fanelli F, Puett D. A model for constitutive lutropin receptor activation based on molecular simulation and engineered mutations in transmembrane helices 6 and 7. J Biol Chem 2002;277:32202-13.

Angelova K, Narayan P, Simon JP, Puett D. Functional role of transmembrane helix 7 in the activation of the heptahelical lutropin receptor. Mol Endocrinol 2000;14:459-471.

Arnis S, Hofmann KP. Two different forms of metarhodopsin II: Schiff base deprotonation precedes proton uptake and signaling state. Proc Natl Acad Sci USA 1993;90:7849-53.

Arnis S, Fahmy K, Hofmann KP, Sakmar TP. A conserved carboxylic acid group mediates light-dependent proton uptake and signaling by rhodopsin. J Biol Chem 1994;269:23879-81.

Avlani VA, Gregory KJ, Morton CJ, Parker MW, Sexton PM, Christopoulos A. Critical role for the second extracellular loop in the binding of both orthosteric and allosteric G-coupled receptor ligands. J Biol Chem 2007;282:25677-86.

Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, Gether U, Javitch JA. 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;276:29171-7.

Balmforth AJ, Lee AJ, Warburton P, Donnelly D, Ball SG. The conformational change responsible for AT1 receptor activation is dependent upon two juxtaposed asparagines residues on transmembrane helices III and VII. J Biol Chem 1997;272:4245-51.

Bartl FJ, Fritze O, Ritter E, Herrmann R, Kuksa V, Palczewski K, Hofmann KP, Ernst OP. Partial agonism in a G protein-coupled receptor: role of the retinal ring structure in rhodopsin activation. J Biol Chem 2005;280: 34259-67.

Berthold M, Kahl U, Juréus A, Kask K, Nordvall G, Langel U, Bartfai T. Mutagenesis and ligand modification studies on galanin binding to its GTP-binding-protein-coupled receptor GalR1. Eur J Biochem 1997;249:601-6.

Bhattacharya S, Ridge KD, Knox BE, Khorana HG. Photoregeneration of rhodopsin is sensitive to interactions with Gt and certain Gt-derived peptides. J Biol Chem 1992;267:6763-8.

Bhattacharya S, Hall SE, Vaidehi N. Agonist-induced conformational changes in bovine rhodopsin: insight into activation of G-protein-coupled receptors. J Mol Biol 2008;382:539-55.

Blanpain C, Doranz BJ, Bondue A, Govaerts C, De Leener A, Vassart G, Doms RW, Proudfoot A, Parmentier M. The core domain of chemokines binds CCR5 extracellular domains while their amino terminus interacts with the transmembrane helix bundle. J Biol Chem 2003;278:5179-87.

Bokoch MP, Zou Y, Rasmussen SG, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJ, Thian FS, Kobilka TS, Puglisi JD, Weis WI, Pardo L, Prosser RS, Mueller L, Kobilka BK. Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. Nature 2010;463:108-12.

Bywater RP. Location and nature of the residues important for ligand recognition in G-protein coupled receptors. J Mol Recognit 2005;18:60-72.

Caetano-Anollés G, Wang M, Caetano-Anollés D, Mittenthal JE. The origin, evolution and structure of the protein world. Biochem J 2009;417:621-37.

Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, StevensRC. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 2007;318:1258-65.

Correa SA, Pignatari GC, Ferro ES, Pacheco NA, Costa-Neto CM, Pesquero JB, Oliveira L, Paiva AC, Shimuta SI. Role of the Cys18-Cys274 disulfide bond and of the third extracellular loop in the constitutive activation and internalization of angiotensin II type 1 receptor. Regul Pept 2006;134:132-40.

Costa-Neto CM, Miyakawa AA, Oliveira L, Hjorth SA, Schwartz TW, Paiva ACM. Mutational analysis of the interaction of N- and C-terminal ends of angiotensin II with the AT1 receptor. Br J Pharmacol 2000;130:1263-8.

De Amici M, Dallanoce C, Holzgrabe U, Tränkle C, Mohr K. Allosteric ligands for G protein -coupled receptors: A novel strategy with attractive therapeutic opportunities. Med Res Rev 2009; E-ahead publication, Jun 25.

Dror RO, Arlow DH, Borhani DW, Jensen MO, Piana S, Shaw DE. Identification of two distinct inactive conformations of the beta2-adrenergic receptor reconciles structural and biochemical observations. Proc Natl Acad Sci U S A 2009;106:4689-94.

Ernst OP, Meyer CK, Marin EP, Henkleini P, Fu W-Y, Thomas P. Sakmar TP, Hofmann KP. Mutation of the Fourth Cytoplasmic Loop of Rhodopsin Affects Binding of Transducin and Peptides Derived from the Carboxyl-terminal Sequences of Transducin a and g Subunits. J Biol Chem 2000;275:1937-43.

Ernst OP, Gramse V, Kolbe M, Hofmann KP, Heck M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. Proc Natl Acad Sci USA 2007;104:10859-64.

Fan QR, Hendrickson WA. Structure of human follicle-stimulating hormone in complex with its receptor. Nature 2005;433: 269-77.

Fathy DB, Kyle DJ, Leeb-Lundberg LM. High-affinity binding of peptide agonists to the human B1 bradykinin receptor depends on interaction between the peptide N-terminal L-lysine and the fourth extracellular domain of the receptor. Mol Pharmacol 2000;57:171-9.

Feng YH, Ding Y, Ren S, Zhou L, Xu C, Karnik SS. Unconventional homologous internalization of the angiotensin II type-1 receptor induced by G-protein-independent signals. Hypertension 2005;46:419-25.

Feng YH, Zhou L, Qiu R, Zeng R. Single mutations at Asn295 and Leu305 in the cytoplasmic half of transmembrane -helix domain 7 of the AT1 receptor induce promiscuous agonist specificity for angiotensin II fragments: a pseudo-constitutive activity. Mol Pharmacol 2005;68:347-55.

Findlay JB, Pappin DJ. The opsin family of proteins. Biochem J 1986;238:625-42.

Fritze O, Filipek S, Kuksa V, Palczewski K, Hofmann KP, Ernst OP. Role of the conserved NPxxY(x)5,6F motif in the rhodopsin ground state and during activation. Proc Natl Acad Sci USA 2003;100:2290-5.

Fukuda M.N., Papermaster D.S., Hargrave P.A. Rhodopsin carbohydrate. Structure of small oligosaccharides attached at two sites near the NH2 terminus. J Biol Chem 1979;254:8201-7.

Gerber BO, Meng EC, Dotsch V, Baranski TJ, Bourne HR. An activation switch in the ligand binding pocket of the C5a receptor. J Biol Chem 2001;276:3394-400.

Gether U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. Endocr Rev 2000;21:90-113.

Gimpl G, Reitz J, Brauer S, Trossen C. Oxytocin receptors: ligand binding, signalling and cholesterol dependence. Prog Brain Res 2008;170:193-204.

Greenfeder S, Cheewatrakoolpong B, Billah M, Egan RW, Keene E, Murgolo NJ, Anthes JC. The neurokinin-1 and neurokinin-2 receptor binding sites of MDL103,392 differ. Bioorg Med Chem 1999;7:2867-76.

Grishin, N. V. Estimation of evolutionary distances from protein spatial structures, J Mol Evol 1997;45:359-69.

Grishin NV. Fold Change in Evolution of Protein Structures. J Struct Biol 2001;134:167-85.

Gurevich VV, Gurevich EV. The molecular acrobatics of arrestin activation. Trends Pharmacol Sci 2004;25:105-11.

Hamm HE, Deretic D, Arendt A, Hargrave PA, Konig B, Hofmann KP. Site of G protein binding to rhodopsin mapped with synthetic peptides from the alpha subunit. Science 1988;241:832-5.

Han M, Smith SO, Sakmar TP. Constitutive activation of opsin by mutation of methionine 257 on transmembrane helix 6. Biochemistry 1998;37:8253-61.

Han M, Gurevich VV, Vishnivetskiy SA, Sigler PB, Schubert C. Crystal structure of beta-arrestin at 1.9 A: possible mechanism of receptor binding and membrane translocation. Structure 2001;9:869-80.

Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EY, Velasquez J, Kuhn P, Stevens RC. A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. Structure 2008;16:897-905.

Hargrave PA, McDowell JH, Curtis DR, Wang JK, Juszczak E, Fong SL, Rao JK, Argos P. The structure of bovine rhodopsin. Biophys Struct Mech 1983;9:235-44.

Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. J Mol Biol 1990:213:899-929.

Herrmann R, Heck M, Henklein P, Henklein P, Kleuss C, Hofmann KP, Ernst OP. Sequence of interactions in receptor-G protein coupling. J Biol Chem 2004;279:24283-90.

Higginbottom A, Cain SA, Woodruff TM, Proctor LM, Madala PK, Tyndall JD, Taylor SM, Fairlie DP, Monk PN. Comparative agonist/antagonist responses in mutant human C5a receptors define the ligand binding site. J Biol Chem 2005;280:17831-40.

Hjorth SA, Schambye HT, Greenlee WJ, Schwartz TW. Identification of peptide binding residues in the extracellular domains of the AT1 receptor. J Biol Chem 1994;269:30953-9.
Hoffmann SH, ter Laak T, Kühne R, Reiländer H, Beckers T. Residues within transmembrane helices 2 and 5 of the human gonadotropin-releasing hormone receptor contribute to agonist and antagonist binding. Mol Endocrinol 2000;14:1099-115.
Hofmann KP, Pulvermüller A, Buczy³ko J, Van Hooser P, Palczewski K. The role of arrestin and retinoids in the regeneration pathway of rhodopsin. J Biol Chem 1992;267:15701-6.
Holm L, and Sander C. Mapping the protein universe. Science 1996;273:595-603.
Holm L, and Sander C. New structure - Novel fold? Structure 1997;5:165-71.
Horn F, Bettler E, Oliveira L, Campagne F, Cohen FE, Vriend G. GPCRDB information system for G protein-coupled receptors (www.gpcr.org). Nucleic Acids Res 2003;31:294-7.
Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 2008;322:1211-7.
Kalatskaya I, Schüssler S, Blaukat A, Müller-Esterl W, Jochum M, Proud D, Faussner A. Mutation of tyrosine in the conserved NPXXY sequence leads to constitutive phosphorylation and internalization, but not signaling, of the human B2 bradykinin receptor. J Biol Chem 2004;279:31268-76.
Kedzie KM, Donello JE, Krauss HA, Regan JW, Gil DW. A single amino-acid substitution in the EP2 prostaglandin receptor confers responsiveness to prostacyclin analogs. Mol Pharmacol 1998;54:584-90.
Khorana HG. Two light-transducing membrane proteins: bacteriorhodopsin and the mammalian rhodopsin. Proc Natl Acad Sci U S A 1993;90:1166-71.
Kim JM, Altenbach C, Kono M, Oprian DD, Hubbell WL, Khorana HG. Structural origins of constitutive activation in rhodopsin: Role of the K296/E113 salt bridge. Proc Natl Acad Sci U S A 2004;101:12508-13.
Kisselev OG, Meyer CK, Heck M, Ernst OP, Hofmann KP. Signal transfer from rhodopsin to the G-protein: evidence for a two-site sequential fit mechanism. Proc Natl Acad Sci USA 1999;96:4898-903.
Kisselev OG, Downs MA, McDowell JH, Hargrave PA. Conformational changes in the phosphorylated C-terminal domain of rhodopsin during rhodopsin arrestin interactions. J Biol Chem 2004;279:51203-7.
Knierim B, Hofmann KP, Ernst OP, Hubbell WL. Sequence of late molecular events in the activation of rhodopsin. Proc Natl Acad Sci U S A 2007;104: 20290-5.
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;103:21-80.
Kyle DJ. Structural features of the bradykinin receptor as determined by computer simulations, mutagenesis experiments, and conformationally constrained ligands: establishing the framework for the design of new antagonists. Mol Pharmacol 2000;57:171-9.
Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105-32.
Le MT, Vanderheyden PM, Szaszak M, Hunyady L, Vauquelin G. Angiotensin IV is a potent agonist for constitutive active human AT1 receptors. Distinct roles of the N- and C-terminal residues of angiotensin II during AT1 receptor activation. J Biol Chem 2000;277:23107-10.
Leeb T, Mathis SA, Leeb-Lundberg LM. The sixth transmembrane domains of the human B1 and B2 bradykinin receptors are structurally compatible and involved in discriminating between subtype-selective agonists. J Biol Chem 1997;272:311-7.
Lefkowitz RJ, Whalen EJ. Beta-arrestins: traffic cops of cell signaling. Curr Opin Cell Biol 2004;16: 162-8.
Li J, Edwards PC, Burghammer M, Villa C, Schertler GF. Structure of bovine rhodopsin in a trigonal crystal form. J Mol Biol 2004;343: 1409-38.
Li T, Franson WK, Gordon JW, Berson EL, Dryja TP. Constitutive activation of phototransduction by K296E opsin is not a cause of photoreceptor degeneration. Proc Natl Acad Sci USA 1995;92: 3551-5.

Lu D, Vage DI, Cone RD. A ligand-mimetic model for constitutive activation of the melanocortin-1 receptor. Mol Endocrinol 1998;12:592-604.
Lüdeke S, Beck M, Yan EC, Sakmar TP, Siebert F, Vogel R. The role of Glu181 in the photoactivation of rhodopsin. J Mol Biol 2005;353:345-56.
Luttrell LM, Lefkowitz RJ. The role of beta-arrestins in the termination and transduction of G-protein-coupled receptor signals. J Cell Sci 2002;115:455-65.
Maeda A, Kandori H, Yamazaki Y, Nishimura S, Hatanaka M, Chon YS, Sasaki J, Needleman R, Lanyi JK. Intramembrane signaling mediated by hydrogen-bonding of water and carboxyl groups in bacteriorhodopsin and rhodopsin. J Biochem. 1997;121:399-406.
Maeda T, Imanishi Y, Palczewski K. Rhodopsin phosphorylation: 30 years later. Prog Retin Eye Res 2003;22:417-34.
Matthews RG, Hubbard R, Brown PK, Wald G. Tautomeric forms of metarhodopsin. J Gen Physiol 1963;47:215-40.
Maudsley S, Martin B, Luttrell LM. The origins of diversity and specificity in G protein-coupled receptor signaling. J Pharmacol Exp Ther 2005;314:485-94.
Meng F, Wei Q, Hoversten MT, Taylor LP, Akil H. Switching agonist/antagonist properties of opiate alkaloids at the delta opioid receptor using mutations based on the structure of the orphanin FQ receptor. J Biol Chem 2000;275:21939-45.
Min L, Ascoli M. Effect of activating and inactivating mutations on the phosphorylation and trafficking of the human lutropin/choriogonadotropin receptor. Mol Endocrinol 2000;14:1797-810.
Miura S, Karnik SS. Angiotensin II type 1 and type 2 receptors bind angiotensin II through different types of epitope recognition. J Hypertens 1999;17:397-404.
Miura S, Feng YH, Husain A, Karnik SS. Role of aromaticity of agonist switches of angiotensin II in the activation of the AT1 receptor. J Biol Chem 1999;274:7103-10.
Miura S, Fujino M, Hanzawa H, Kiya Y, Imaizumi S, Matsuo Y, Tomita S, Uehara Y, Karnik SS, Yanagisawa H, Koike H, Komuro I, Saku K. Molecular mechanism underlying inverse agonist of angiotensin II type 1 receptor. J Biol Chem 2006;281:19288-95.
Mirny LA, Shakhnovich EI. Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. J Mol Biol 1999;291:177-96.
Morizumi T, Imai H, Shichida Y. Direct observation of the complex formation of GDP-bound transducin with the rhodopsin intermediate having a visible absorption maximum in rod outer segment membranes. Biochemistry 2005;44:9936-43.
Mousavi SA, Malerød L, Berg T, Kjeken R. Clathrin-dependent endocytosis. Biochem J 2004;377:1-16.
Murakami M, Kouyama T. Crystal structure of squid rhodopsin. Nature 2008;453:363-7.
Muto T, Tsuchiya D, Morikawa K, Jingami H. Structures of the extracellular regions of the group II/III metabotropic glutamate receptors. Proc Natl Acad Sci USA 2007;104:3759-64.
Nabi IR, Le PU. Caveolae/raft-dependent endocytosis. J Cell Biol 2003;161:673-7.
Natochin M, Gasimov KG, Moussaif M, Artemyev NO. Rhodopsin determinants for transducin activation: a gain-of-function approach. J Biol Chem 2003;278:37574-81.
Nakagawa M, Miyamoto T, Kusakabe R, Takasaki S, Takao T, Shichida Y, Tsuda MO. Glycosylation of G-protein-coupled receptor, octopus rhodopsin. Direct analysis by FAB mass spectrometry. FEBS Lett 2001;496:19-24.
Neve KA, Cumbay MG, Thompson KR, Yang R, Buck DC, Watts VJ, DuRand CJ, Teeter MM. Modeling and mutational analysis of a putative sodium-binding pocket on the dopamine D2 receptor. Mol Pharmacol 2001;60:373-81.
Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. J Mol Biol 2004;342:571-83.
Oliveira L, Paiva ACM, Vriend G. A common motif in G-protein coupled seven transmembrane helix receptors. J Comp-Aid Mol Des 1993;7:649-58.
Oliveira L, Paiva AC, Vriend G. A low resolution model for the interaction of G proteins with G protein-coupled receptors. Protein Eng 1999;12:1087-95.
Oliveira L, Paiva PB, Paiva AC, Vriend G. Identification of functionally conserved residues with the use of entropy-variability plots. Proteins 2003a;52:544-52.

Oliveira L, Paiva PB, Paiva AC, Vriend G. Sequence analysis reveals how G protein-coupled receptors transduce the signal to the G protein. Proteins 2003b;52:553-60.
Oliveira L, Costa-Neto CM, Nakaie CR, Schreier S, Shimuta SI, Paiva AC. The angiotensin II AT1 receptor structure-activity correlations in the light of rhodopsin structure. Physiol Rev 2007;87: 565-92.
Ovchinnikov YuA. Rhodopsin and bacteriorhodopsin: structure-function relationships. FEBS Lett 1982;148:179-91.
Ovchinnikov IuA, Abdulaev NG, Fe-gina MIu, Artamonov ID, Bogachuk AS. Visual rhodopsin. III. Complete amino acid sequence and topography in a membrane. Bioorg Khim 1983;9:1331-40.
Ovchinnikov YuA, Abdulaev NG, Zolotarev AS, Artamonov ID, Bespalov IA, Dergachev AE, Tsuda M. Octopus rhodopsin. Amino acid sequence deduced from cDNA. FEBS Lett. 1988;232:69-72.
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. Crystal structure of rhodopsin: A G protein-coupled receptor. Science 2000;289:739-45.
Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP. Crystal structure of the ligand-free G-protein-coupled receptor opsin. Nature 2008;454:183-7.
Pauwels PJ, Wurch T. Review: amino acid domains involved in constitutive activation of G-protein-coupled receptors. Mol Neurobiol 1998;17:109-35.
Perez DM, Karnik SS. Multiple signaling states of G-protein-coupled receptors. Pharmacol Rev 2005;57:147-61.
Pioszak AA, Parker NR, Suino-Powell K, Xu HE. Molecular recognition of corticotropin-releasing factor by its G-protein-coupled receptor CRFR1. J Biol Chem 2008;283:32900-12.
Pulvermüller A, Palczewski K, Hofmann KP. Interaction between photoactivated rhodopsin and its kinase: stability and kinetics of complex formation. Biochemistry 1993;32:14082-8.
Rao VR, Cohen GB, Oprian DD. Rhodopsin mutation G90D and a molecular mechanism for congenital night blindness. Nature 1994;367:639-42
Ridge KD, Marino JP, Ngo T, Ramon E, Brabazon DM, Abdulaev NG. NMR analysis of rhodopsin-transducin interactions. Vision Res 2006;46:4482-92.
Rim J, Oprian DD. Constitutive activation of opsin: interaction of mutants with rhodopsin kinase and arrestin. Biochemistry 1995;34:11938-45.
Robinson PR, Buczylko J, Ohguro H, Palczewski K. Opsins with mutations at the site of chromophore attachment constitutively activate transducin but are not phosphorylated by rhodopsin kinase. Proc Natl Acad Sci USA 1994;91:5411-5.
Roth BL, Willins DL, Kristiansen K, Kroeze WK. 5-Hydroxytryptamine2-family receptors (5-hydroxytryptamine2A, 5-hydroxytryptamine2B, 5-hydroxytryptamine2C): where structure meets function. Pharmacol Ther 1998;79:231-57.
Ruprecht, J., Mielke, T., Vogel, R., Villa, C. & Schertler, G. F. X. Electron crystallography reveals the structure of metarhodopsin I. EMBO J 2004;23:3609-20.
Russell RB, Ponting CP. Protein fold irregularities that hinder sequence analysis. Curr Opin Struct Biol 1998;8:364-71.
Russell RB, Sasieni PD, Sternberg MJ. Supersites within superfolds. Binding site similarity in the absence of homology. J Mol Biol 1998a;282:903-18.
Russell, R B, Saqi, M A, Bates, P A, Sayle R A, and Sternberg M J. Recognition of analogous and homologous protein folds—Assessment of prediction success and associated alignment accuracy using empirical substitution matrices, Protein Eng 1998b;11:1-9.
Sander C, Schneider R. Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins 1991;9:56-68.
Santos EL, de Picoli Souza K, Cibrián-Uhalte E, Oliveira SM, Bader M, Costa-Neto CM, Oliveira L, Pesquero JB. Essential role of TM V and VI for binding the C-terminal sequences of Des-Arg-kinins. Int Immunopharmacol 2008;8:282-8.
Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP. Crystal structure of opsin in its G-protein-interacting conformation. Nature 2008;455:497-502.

Schertler GF. Structure of rhodopsin and the metarhodopsin I photointermediate. Curr Opin Struct Biol 2005;15:408-15.
Schertler GF, Villa C, Henderson R. Projection structure of rhodopsin. Nature 1993;362:770-2.
Shi W, Sports CD, Raman D, Shirakawa S, Osawa S, Weiss ER. Rhodopsin arginine-135 mutants are phosphorylated by rhodopsin kinase and bind arrestin in the absence of 11-cis-retinal. Biochemistry 1998;37:4869-74.
Shieh T, Han M, Sakmar TP, Smith SO. The steric trigger in rhodopsin activation. J Mol Biol 1997;269:373-84.
Thomas WG, Qian H, Chang CS, Karnik S. Agonist-induced phosphorylation of the angiotensin II (AT(1A)) receptor requires generation of a conformation that is distinct from the inositol phosphate-signaling state. J Biol Chem 2000;275:2893-900.
Tobin AB. G-protein-coupled receptor phosphorylation: where, when and by whom. Br J Pharmacol 2008;153:S167-76.
Trumpp-Kallmeyer S, Hoflack J, Bruinvels A, Hibert M. Modeling of G-protein-coupled receptors: application to dopamine, adrenaline, serotonin, acetylcholine, and mammalian opsin receptors. J Med Chem 1992;35:3448-62.
Unger VM, Hargrave PA, Baldwin JM, Schertler GF. Arrangement of rhodopsin transmembrane alpha-helices. Nature 1997;389:203-6.
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;280:17135-41.
Vlaeminck-Guillem V, Ho SC, Rodien P, Vassart G, Costagliola S. Activation of the cAMP Pathway by the TSH Receptor Involves Switching of the Ectodomain from a Tethered Inverse Agonist to an Agonist. Mol Endocrinol 2002;16:736-46.
Vogel R, Siebert F, Ludeke S, Hirshfeld A, Sheves M. Agonists and Partial Agonists of Rhodopsin: Retinals with Ring Modifications. Biochemistry 2005;44:11684-99.
Vriend G. WHAT IF: a molecular modeling and drug design program. J Mol Graph 1990;8:52-6.
Yan EC, Kazmi MA, De S, Chang BS, Seibert C, Marin EP, Mathies RA, Sakmar TP. Function of extracellular loop 2 in rhodopsin: glutamic acid 181 modulates stability and absorption wavelength of metarhodopsin II. Biochemistry 2002;41:3620-7.
Yarden Y, Rodriguez H, Wong SK, Brandt DR, May DC, Burnier J, Harkins RN, Chen EY, Ramachandran J, Ullrich A, *et al.*The avian beta-adrenergic receptor: primary structure and membrane topology. Proc Natl Acad Sci U S A 1986;83:6795-9.
Yu J, Prado GN, Taylor L, Piserchio A, Gupta A, Mierke DF, Polgar P. Global chimeric exchanges within the intracellular face of the bradykinin B2 receptor with corresponding angiotensin II type Ia receptor regions: generation of fully functional hybrids showing characteristic signaling of the AT1a receptor. J Cell Biochem 2002;85:809-19.
Wang Z, Asenjo AB, Oprian DD. Identification of the Cl(-)-binding site in the human red and green color vision pigments. Biochemistry 1993;32:2125-30.
Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF. Structure of a beta1-adrenergic G-protein-coupled receptor. Nature 2008;454:486-91.
Watson S, Arkinstall, S. The G-protein linked receptor. Facts Book. 1994. Academic Press London.
Webel R, Menon I, O'Tousa JE, Colley NJ. Role of asparagine-linked oligosaccharides in rhodopsin maturation and association with its molecular chaperone, NinaA. J Biol Chem 2000;275:24752-9.
Wei H, Ahn S, Shenoy SK, Karnik SS, Hunyady L, Luttrell LM, Lefkowitz RJ. Independent beta-arrestin 2 and G protein-mediated pathways for angiotensin II activation of extracellular signal-regulated kinases 1 and 2. Proc Natl Acad Sci USA 2003;100:10782-7.
Wilden U, Kühn H. Light-dependent phosphorylation of rhodopsin: number of phosphorylation sites.Biochemistry 1982;21:3014-22.
Zhukovsky EA, Robinson PR, Oprian DD. Transducin activation by rhodopsin without a covalent bond to the 11-cis-retinal chromophore. Science 1991;251:558-60.