New Prospects for Drug Discovery (IV) I.2

(Part of this chapter was done by Laura Iarriccio).

G-protein-coupled receptors (GPCR) are a major class of membrane proteins belonging to a continuously growing superfamily. These receptors play a critical role in signal transduction, and are among the most important pharmacological drug targets. The first structural model for the GPCR superfamily was the bacterial protein bacteriorhodopsin with its characteristic seven transmembrane (TM) helical architecture. The visual photoreceptor rhodopsin is a better model for GPCR, and the recent elucidation of the crystal structure of bovine rhodopsin has renewed the interest in this receptor as a template for molecular modeling of other GPCR, particularly for the implications in ligand design and drug discovery. Rhodopsin will continue to be a widely used model for GPCR but rhodopsin-based approaches have to be complemented by other theoretical and experimental approaches -while waiting for the crystal structure of other members of the superfamily- if these want to be successfully used for drug discovery.

A complete knowledge of the structural cavity for ligand binding will allow the design of new agonists and antagonists to be proposed. In this regard, new avenues for ligand design have been opened up by the resolution of the crystal structure of the first GPCR to be crystallized, i.e. the visual photoreceptor rhodopsin. In the wide family of GPCR, ligands bind to the extracellular part of the receptor, and can interact with the TM domains, the three extracellular loops and/or the N-terminus. A significant number of GPCR binding sites for their corresponding ligands have been mapped from a combination of site-directed mutagenesis –and other biochemical techniques- with homology modeling approaches. In this way an important number of small molecules able to bind to the receptors have been characterized. These ligands vary from small catecholamines such as serotonin and dopamine, to large peptides and small proteins, such as hormones and chemokines. Small-molecules binding sites are where most of orally bioavailable drug candidates bind, from the drug discovery point of view [Becker et al., 2003].

A number of examples from site-directed mutagenesis experiments indicate that many small-molecule agonist and antagonist molecules bind completely in the TM region of the receptor [Klabunde and Hessler, 2002]. For the β_2-adrenergic receptor, catecholamines agonists have been shown to bind to a region of the receptor comprising helices III, V, VI and VII [Tota et al., 1991; Savarese and Fraser, 1992]. In the case of the α_{1A}-adrenergic receptor similar studies were carried out to examine binding of dihydropyridine and 4-piperidyloxazole antagonists [Hamaguchi et al., 1996; Hamaguchi et al., 1998]. The detailed analysis of ligand-receptor binding revealed critical amino acid residues involved in this molecular recognition process. This is the case of D3.32 that was determined to be very important in ligand binding studies of the 5-HT$_{1A}$ serotonin receptor [Jacoby et al., 1999].

The binding domain of other ligand has been also mapped to the TM domain of the receptor. This is the case of losartan, an antihypertensive in type 1 angiotensin II receptor (AT1), which was mapped within the seven TM region of the receptor [Ji et al., 1995]. Another interesting example is that of substance P, an 11-amino acid peptide, that binds to its endogenous neurokinin-1 (NK1) receptor through interactions with three residues in the first extracellular segment (Asn-23, Gln-24 and Phe-25 in NK1 receptor sequence), two residues in the second extracellular loop (Asp-96 and His-108 in NK1 receptor sequence), and several residues in TM II and TM VII of this receptor [Fong et al., 1993]. However, the binding of small-molecule antagonists was mapped within the TM region between helices IV, V and VI [Cascieri et al., 1995].

The examples cited above correspond to receptors that bind their ligands primarily in the TM domain of the protein. However, some studies question the common ligand-binding site for the stabilization of the active form of the receptor [Klabunde and Hessler, 2002], and in some cases antibody binding to extracellular loops mimics the action of the ligand by activating the receptor [abu Alla et al., 1996]. Allosteric modulation of ligand binding – by increasing or reducing ligand binding effect- has also been described for several GPCR [Soudijn et al., 2001; Christopoulos et al., 2002]. These are factors that should be taken into account when studying the ligand-receptor binding process.

lthough the crystal structure of rhodopsin has provided important information useful for getting new insights into these other receptors structure and function it should be stressed again that rhodopsin is the only receptor that has its ligand bound in its inactive conformation, making this a unique case in the GPCR superfamily. It would be, thus, very interesting to have the structure of opsin (the apoprotein without the 11-*cis*-retinal ligand bound) and that of the photoactivated rhodopsin, MetaII, in order to study the conformational rearrangements accompanying receptor activation. Elucidating the molecular details of the receptor activation process is a key issue in drug design and in the development of new molecules with increased binding affinity and specificity.

I.2.1. STRUCTURAL AND FUNCTIONAL ASPECTS RELEVANT TO DRUG DESIGN

OLIGOMERIZATION

Most of the studies aimed at designing new drugs for GPCR have assumed that receptors were in their monomeric form. This assumption has recently been changed by the description of a number of GPCR that can be found in oligomeric (homo-oligomeric or hetero-oligomeric) state [George et al., 2002].

Dimerization has been proposed as a general mechanism for modulating GPCR function. In the case of rhodopsin, the possible arrangement of this protein in the photoreceptor cell membranes as a dimer has recently been proposed [Fotiadis et al., 2003a]. New details on rhodopsin dimerization, and its implication in the receptor-G-protein binding process, have been recently provided [Liang et al., 2003]. The main molecular interactions involved in rhodopsin dimer formation have been described [Fotiadis et al., 2003a]. The organization of rhodopsin as dimers *in vivo* has been a matter of controversy [Chabre et al., 2003; Fotiadis et al., 2003b]. It has been argued that rhodopsin would be an arquetypical monomeric receptor dispersed in a random fashion and moving freely in the lipid membrane of disc photoreceptor cells as proposed previously by means of classical biophysical techniques [Cone, 1972; Roof and Heuser, 1982]. The recent AFM study describing dimer formation of rhodopsin has proposed formation of rows of highly packed rhodopsin molecules in isolated membranes at a density double of that described previously by *in situ* studies [Roof and

Heuser, 1982]. This model has been challenged because lipids would have been excluded from the packing and this would make it incompatible with the expected rapid rotational diffusion. However, Fotiadis et al. [Fotiadis et al., 2003b] argue against an artifact of the AFM technique because they can see the same effect at room temperature and although this is not physiological temperature, a temperature of 3ºC has been described for phase transition of bovine disc photoreceptor lipids. In any case the subject of dimer formation of rhodopsin in its native environment is still a matter of debate.

The process of GPCR dimer or oligomer formation, and its effect on receptor function, is not currently well understood, but it is generally agreed that correct formation of oligomers would be a requirement for receptor expression to the cell surface as well as for receptor function. Hetero-oligomer formation can lead to complex formation with different binding affinities for different ligands. This can change the basic features of the receptors and this can result in a new potential pharmacological target. In addition, ligand binding can play a role in oligomer formation of the receptor, promoting association or dissociation of the oligomer, or binding the oligomer and changing its conformation. These important aspects on ligand binding should be taken into account for effective drug design. New strategies based on GPCR oligomerization have been recently reviewed [George et al., 2002], including the development of dimeric ligands like those proposed for the 5-HT1 receptor [Halazy et al., 1996]. One of these strategies involves using new approaches, like enhancing and disrupting oligomerization for improving lead discovery and optimization, by using old drugs [Hillion et al., 2002]. One of the relevant questions concerning GPCR heterodimers is whether these generate new receptors in terms of the pharmacology of the ligand or its function [Milligan, 2004]. An example is that a number of antiparkinsonian agents may have higher affinity for the D3/D2 dopaminergic heterodimer than for the corresponding homodimer [Maggio et al., 2003]. It will be important to develop new experimental strategies that allow detecting ligand binding to heterodimers in the presence of homodimers, for effective and rapid screening of ligands [Milligan, 2004].

The design of dimeric ligands has also prospects of becoming important in the near future. In this case the oligomeric nature of GPCR should be carefully considered because these dimeric ligands can have increased affinity and changed potency with regard to its

constituent monomeric moieties. The dimeric ligand would stabilize the dimeric conformation of the receptor increasing the efficiency of the signal transduction process. In summary, oligomerization of GPCR, i.e. quaternary structure, is an important factor that has to be critically assessed for potential new drug design.

ACTIVATION MECHANISM OF THE RECEPTOR

The inactive state of rhodopsin has 11-*cis*-retinal covalently bound to the apoprotein opsin. In the activation process, retinal is photoisomerized to the all-*trans* configuration by light absorption [Wang et al., 1994]. This initial photoisomerization of the ligand is followed by a number of discrete conformational states of the receptor that can be spectroscopically characterized [Okada et al., 2001; Matthews et al., 1963]. The active photointermediate that binds and activates the G-protein is called MetaII [Choi et al., 2002]. The conformational change that results in active Meta II from rhodopsin may be similar in the activation mechanism of other GPCR [Sakmar et al., 2002; Meng and Bourne, 2001; Bissantz, 2003; Gether et al., 2002]. MetaII binding to the G-protein transducin triggers a cascade of enzymatic reactions that leads to the hyperpolarization of the membranes of photoreceptor cells [Stryer, 1986].

In the activation mechanism of GPCR several proteins play a significant role, and distinct functional states are characterized by distinct conformational structures. It is of central interest for drug design to carefully investigate these functional states, focusing on the inactive and active conformations in order to have a structural basis for the study of agonist and antagonist binding to the receptor.

In the case of rhodopsin the 11-*cis*-retinal acts as a strong inverse agonist, by keeping the receptor in its inactive state. Light photoconversion of 11-*cis*-retinal to the all-*trans* form is an irreversible process because there is no equilibrium among the different states in contrast to other GPCR in which these states are interconvertible. During the activation process three phases can be observed: i) the *cis-trans* isomerization of the ligand; ii) thermal relaxation of the retinal protein complex; iii) the late equilibria that are affected by the interaction of rhodopsin with the G-protein. The MetaII active conformation formed decays

to free all-*trans*-retinal plus opsin, either directly or through the formation of MetaIII, a proposed storage form of the receptor [Heck et al., 2003]. Recently formation of MetaIII directly from another photointermediate, MetaI, has been proposed [Vogel et al., 2004]. MetaIII decays to free opsin and all-*trans*-retinal where the covalent bond between retinal and opsin is eventually hydrolyzed. A new molecule of fresh 11-*cis*-retinal will occupy the empty binding pocket of opsin and this would provide the molecular basis of the regeneration cycle of rhodopsin.

Several models have been proposed for the activation process of other members of the GPCR superfamily. Among them, the extended ternary complex model (eTCM) [Samama et al., 1993] and the cubic ternary complex model (CTC) [Weiss et al., 1996a; Weiss et al., 1996b; Kenakin, 2003] were proposed (Figure I.2.1). The CTC model, being more thermodynamically complete, also allows for the binding of G-protein with the resting-state receptors [Chen et al., 2003]. These models are interesting because they incorporate the active state of the receptor in the system allowing the study of the activation of G protein in the absence of ligand (constitutive activity).
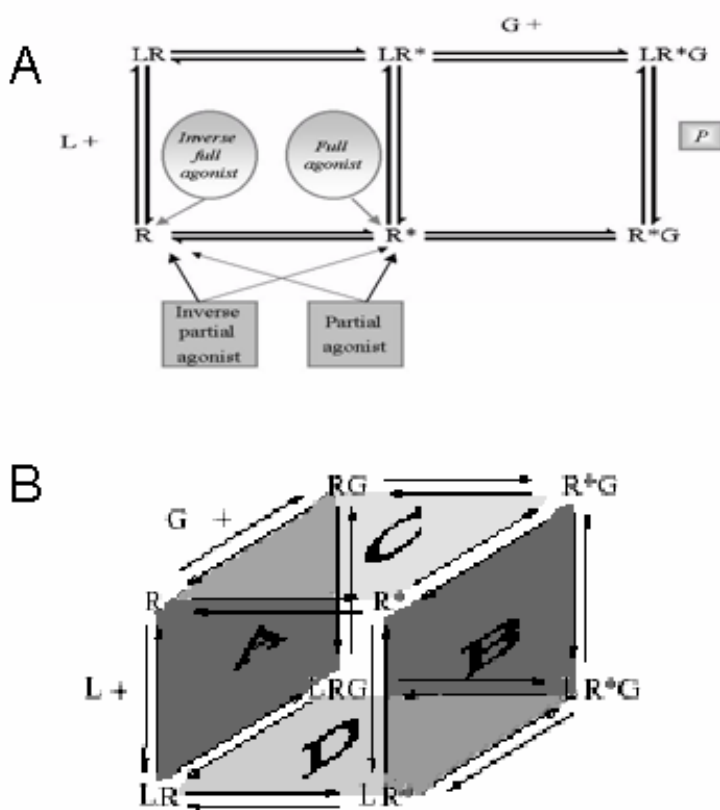

Fig. I.2.1. **Different states and equilibria of GPCR**. These figures represent the different equilibria in which the receptor is involved (ligand binding and G-protein activation). (A) eTMC model. Different classes of L are shown in the figure, indicating how this can affect the different conformational states of the receptor. For example a partial inverse agonist can act upon R and R*, although is shows stronger affinity for R (indicated by thick arrow). The contrary happens with a partial agonist that shows stronger preference for binding to the activated R* state. The potency or efficiency of the ligand is also represented (P) [Weiss et al., 1996b]. (B) CTC model. It corresponds to a modified version of the CTC model previously described [Kenakin, 2003]. The activated states are represented by +. Side A of the cube represents conformational states of the inactive receptor. Side B corresponds to the active states of the receptor. Side C would illustrate constitutive activity of the receptor (activation in the absence of activating stimulus, i.e. ligand binding). Side D corresponds to the typical ligand-induced activation of the receptor. L= ligand, R= Receptor, G= G-protein.

Other concepts are introduced like "agonist-directed trafficking of receptor stimulus' [Weiss et al., 1996a; Weiss et al., 1996b] or 'strength of signaling' (in which different receptor conformations can be detected depending on the potency of the agonist), that account for the fact that an individual receptor may couple to a variety of different G-proteins, leading to activation of multiple signaling pathways. It is therefore possible that an agonist -by stabilizing the different activated states of a receptor- may have preference for binding to a particular receptor-G-protein complex, directing signals to one of these pathways [Kenakin, 2003]. Receptor binding drugs can behave as full or partial agonists, inverse antagonists or neutral antagonists depending on how they act in the receptor function. The significance of these concepts in current rational drug therapy has been previously discussed [Brink et al., 2004]. Several agonists can function as partial agonists in certain systems and as inverse agonists in others; this has been termed protean behavior of agonists [Weiss et al., 1996a; Weiss et al., 1996b]. Several experimental lines of evidence have been reported for this kind of behavior [Brink, 2002].

In the proposed models the receptor and the G protein exist in a continuous equilibrium between exchangeable states, the active and inactive states, G protein bound and unbound receptors that are fluctuating very rapidly between them. The concept of agonist/ligand directed activation was first proposed for the α1A-adrenergic receptor [Perez et al., 1996]. The inactive state is the predominant state in the absence of ligand and the active state it the one conformationally coupled to the physiologic effector system. When the receptor binds the ligand the equilibrium state is changed increasing the activation levels (in the case of an agonist) or decreases them (in the case of an inverse agonist). The activated receptor is very unstable and tends to go back to the inactive state or to stabilize itself by binding to the corresponding G protein. This is a serious drawback for crystallizing the active conformation of GPCR. In the absence of ligand, there is always a very low basal level of G protein activation that is the equilibrium is somehow slightly shifted to the active conformation. This is called constitutive activity of the receptor and is also a factor that can be increased by mutation and needs to be considered in drug design.

In rhodopsin a number of mutations have been associated with retinal diseases like RP and CSNB [Farrar et al., 2002; van Soest et al., 1999]. Three of those Gly90Asp (G2.57D),

Ala292Glu (A7.39E) and Thr94Ile (T2.61I) [Ramon et al., 2003] cause CSNB and the molecular mechanism of the disease is proposed to involve constitutive activation of the receptor. These mutations could be affecting the electrostatic network of interactions in the vicinity of the Schiff base linkage between the protonated N and E3.28 counterion causing constitutive activity of the mutant rhodopsins (as schematically represented by face C of the CTC model, see Figure I.2.1.) Other receptors show naturally occurring mutations that are related to disease and this is also important when analyzing the possible physiological consequences of GPCR structure alterations.

As stated above it would be extremely desirable, for the discovery and development of new drugs, to have a well-defined structure for the inactive and active conformation of the different GPCR of pharmacological interest. This could allow the elucidation of the molecular details of the activation mechanism of the receptor. The structure of the inactive state of the receptor is important for the study of antagonists, whereas the structure of the active state would allow the study of different agonists. Unfortunately we only have at present the structure of the inactive state of rhodopsin [Palczewski et al., 2000; Teller et al., 2001; Okada et al., 2002]. For molecular modeling of the active state it is necessary to follow the published literature in order to build a reasonable model that incorporates all the experimental data available. It is not clear at this stage whether information derived for a given GPCR can be directly extrapolated to other members of the family, or whether there are significant differences in the activation process of different GPCR. In spite of this limitation, and according to our present knowledge, it is believed that there should not be critical differences in the activation mechanism that would justify different schemes for receptor activation [Meng and Bourne, 2001; Gether et al., 2002]. Modeling of the active conformation has to start from the hypothesis of the activation mechanism and the activation models should take into account that the ligand has to be admitted to the ligand binding pocket, that the receptor should convert from the inactive to the active form, and finally that the G-protein should become activated by the activated receptor.

A mechanism of activation for family A GPCR has been proposed in accordance with experimental data on rhodopsin activation [Okada et al., 2001; Meng and Bourne, 2001; Hubbell et al., 2003; Nikiforovich et al., 2003]. Helix movements are the major components

of the transition from the inactive to the active states. In particular, the movement of the C-terminal part of helix VI, by rotating in a clock-wise fashion, apart from the hydrophobic core of the TM domain (particularly away from helix III), has been described. Less pronounced movements of helices II, III and VII relative to the TM core accompany the movement of helix VI. Thus, a site for G protein binding is generated by opening-up of the cytoplasmic site of the protein as a result of the receptor expansion induced by helical movements. Cytosolic loops II and III would also be involved in generating this binding site for the G-protein [Acharya et al., 1997; Arimoto et al., 2001].

In the rhodopsin activation process, MetaII formation results in a spectral shift in the visible spectrum and a concomitant breakage of the salt bridge between E3.28 and the protonated Schiff base linkage with K7.43, that constraints the receptor in its inactive form. This is one of the causes of the TM helical movement. MetaII formation is associated with capture of a proton from the cytoplasmic face and probably results in the protonation of E3.49 [Fahmy et al., 2000; Arnis and Hofmann, 1993] that belongs to the highly conserved D/ERY motif. E3.49 is forming a salt bridge with R3.50, and is also interacting with E6.30 and T6.34 in helix VI. These electrostatic interactions are important for the structure of the inactive receptor and may also play a role in receptor activation [Ramon et al., In Press]. Protonation of E3.49 would destabilize the network of electrostatic interactions and induce the movement of the TM helices. A similar effect has been described for the β_1-adrenergic receptor [Scheer et al., 1997] and for the thrombin receptor [Seibert et al., 1999] with an analogous proton translocation.

A structural model for the active MetaII conformation of rhodopsin has been proposed that takes into account all experimental data available from spectroscopic as well as mutational studies [Hubbell et al., 2003]. This model shows a good agreement with experimental data on conformational changes induced by light, like those observed in EPR studies on distance changes of spin labels upon rhodopsin photoactivation, and changes in relative mobility and accessibilities of Cys residues in rhodopsin mutants. The location of the cytosolic loop III described in this model would also be suitable for G-protein interaction. The model is a good starting point for the study of the activated form of rhodopsin and for the study of other active forms of different GPCR, although there are some limitations like the

assumption that the non-TM segments of the protein retain the conformation of the initial state.

I.2.2. NOVEL DRUG DESIGN

GPCR represent over 50% of the drugs on the market today and 26 of the top 100 pharmaceutical products are compounds that target GPCR [Klabunde and Hessler, 2002; Wise et al., 2002]. The Human Genome Project [Venter et al., 2001; Lander et al., 2001] has identified around 400 GPCR (except olfactory receptors) that are considered to be possible pharmacological targets, but currently marketed drugs target only about 30 of them. Natural ligands have been identified for 210 receptors but remain still unknown for 160, called orphan receptors, the natural ligands and the function of which have not been reported yet [Wise et al., 2002]. A major interest of the pharmaceutical industry is on the knowledge of the (patho) physiologic roles of these receptors, in order to be used as targets for novel drugs.

A number of modern approaches to drug discovery are currently being used. These comprise different research areas like proteomics, bioinformatics, screening techniques, combinatorial chemistry, compound library design, ligand- and structure-based drug design and pharmacokinetic approaches. In particular, methodological approaches using high throughput techniques, that allow processing of a high number of samples, are being investigated.

I.2.2.1. LIGAND-BASED DRUG DESIGN

The increasing knowledge available on GPCR and their ligands enables novel drug design strategies to accelerate the finding and optimization of GPCR leads. Generally, drug design approaches may be regarded as being guided by two major strategies. One of these is the classical strategy using a chemical approach in which the ligand is being used as a powerful method for lead finding and optimization. It is based on the structural and chemical similarity principle of ligands. This method is based on the availability of one or several

known ligands or active molecules without taking into account the 3D-receptor structure. Natural ligands for many GPCR are taken as a starting point for lead finding, and information from structure-activity relationships (SAR) can be derived [Klabunde and Hessler, 2002]. Taking this ligand as a starting point, a hypothetical pharmacophore of the receptor is constructed, that can be virtually assembled mimicking the pattern present on the original template structure (pharmacophore modeling) [Cacace et al., 2003]. Through 3D pharmacophore-based design a preselection of the compounds that will bind to the target can be made. The clear advantage of work based on a pharmacophore is that it can be applied to ligands for which no three-dimensional structure of the corresponding receptor is known, although 3 or 4 pharmacophoric point descriptions are required [Bissantz et al., 2003; Mason et al., 1999]. The results can be used in a virtual screening process for lead finding by using a database of small ligands that target a given receptor-binding site. Virtual screening, or *in silico* screening, is a new approach of interest for the pharmaceutical industry as a productive and cost-effective technology in the search for novel lead compounds [Waszkowycz et al., 2001].

Ligand-based 3D quantitative structure-activity relationship (3D-QSAR) methods have supported the chemical optimization of numerous GPCR lead compounds. In addition, the cross-target analysis of GPCR ligands has revealed more and more common structural motifs and 3D pharmacophores.

I.2.2.2. STRUCTURE-BASED DRUG DESIGN

An important approach to drug design is based on the structure of the target receptor, that is structure-based design relying on a 3D-receptor model. Structural information comes from NMR, X-ray crystallographic and spectroscopic methods, and molecular modeling. New opportunities for the development of therapeutics have risen from the crystal structure of rhodopsin determined experimentally. This has stimulated the availability of reasonable structural models for the binding site of a number of GPCR, and this allows potential interactions at the atomic level to be explored from a computational approach [Ballesteros and Palczewski, 2001a]. Using the complete 3D structure of a target should be more efficient than merely using the pharmacophore because in the first case the protein environment is

also being considered. Docking methods score compounds on how well they fit the shape and chemical features of the ligand-binding site, and these are also used to predict the binding modes and binding affinities of every compound in the data set to a given biological receptor. However these techniques have also some problems associated with docking-scoring inaccuracies [Dixon et al., 1997; Tame et al., 1999].

Two main applications of rhodopsin structural approaches to drug design can be considered: i) lead identification based on the crystal structure of retinal rhodopsin recently solved. The availability of this structure has fostered the appearance of structural models for the binding site of different GPCR, mainly of its family A, and this is guiding the current pathway of rational chemical design. Interestingly, conservation of key amino acid residues in the ligand-binding pocket has been observed [Ballesteros et al., 2001b; Waszkowycz et al., 2001]. Once a ligand binding cavity is identified several approaches can be used, like the aforementioned virtual screening techniques, or *de novo* generation approaches [Bohacek and McMartin, 1997] that experimentally combine high throughput screening with combinatorial chemistry in order to develop a good lead based on the target receptor structure; and subsequent lead optimization. Once a particular ligand is selected, the detailed interactions between the ligand and the receptor have to be determined for proper docking procedures to be carried out effectively [Shoichet et al., 2002]. The current rhodopsin structural model can be used as a template to probe specific interactions at the atomic level starting from the 3 Å to 5 Å resolution [Ballesteros and Palczewski, 2001a]. SAR, together with experimental data obtained mainly from site-directed mutagenesis studies in the receptor with minimal ligand modifications, is one of the preferred approaches undertaken to identify the key interactions between ligand and receptor. A representative example of this approach has been reported for binding of epinephrine to β2-adrenergic receptor [Ballesteros and Palczewski, 2001a]. Several factors have to be taken into account due to the diversity among the different GPCR subfamilies. Among those: i) energetic calculations between the complex and the lipid membrane environment where the receptor is located; ii) for this kind of studies it is necessary to confront the molecular models with experimental data. Computational calculations are important to ascertain that the proposed molecular structure is valid and consistent with experimental data in order to take further decision along the modeling process [Ballesteros and Palczewski, 2001a].

I.2.2.3. MOLECULAR MODELING OF GPCR

This approach is very useful because it gathers all available information on SAR for GPCR. The process of building a model can be based on different kinds of approaches that are outlined. Experimental findings from structure-activity, mutagenesis and affinity labeling studies have been used to revise and refine the models. By using this large data information of the binding site, receptor-ligand complexes may be generated. A number of examples have been reported on this approach that allowed mapping of the binding site of GPCR with their antagonists.

RHODOPSIN-BASED GPCR HOMOLOGY MODELS

Homology modeling is used to model new structures of GPCR based on known GPCR structures previously solved by X-ray crystallography [Fiser et al., 2001]. In order to verify homology modeling, it has to be confirmed by site-directed mutagenesis and ligand binding assays. Currently rhodopsin is the unique solved structure and it is used to design new 3D models. Molecular models of GPCR were based on the structure of bacteriorhodopsin before the crystal structure of rhodopsin was reported. However, helical packing of both proteins is clearly different in spite of the presence of a seven TM helical domain. Nevertheless, in some specific cases, like that of the 5-HT$_{2B}$ receptor, bacteriorhodopsin was found to yield better modeling results than rhodopsin [Manivet et al., 2002]. A 3D model of AT1 receptor (small hormone peptide), based on bacteriorhodopsin, has been also reported [Underwood et al., 1994]. For receptor binding of large proteins, like the chemokine receptor CCR2, mutagenesis studies showed that an acidic E7.39 residue in the TM VII was critical for binding of the spiropiperidine series using a bacteriorhodopsin-based structural model of a new lead bound to the CCR2 receptor [Mirzadegan et al., 2000].

Rhodopsin has been used as a template for GPCR of the A family like the D2 dopamine [Ballesteros et al., 2001b] and M1 muscarinic receptors [Lu et al., 2002b] in order to interpret available structure-function data and for constructing realistic models. Several of these examples have been recently reported in homology models based on rhodopsin and

bacteriorhodopsin [Marvin et al., 2001]. Models of the TRH receptor on a generic template for GPCR [Osman et al., 1999] and for the histamine H1-receptor model have been also reported [Ter Laak et al., 1995]. Molecular models based on rhodopsin are useful but some aspects have to be taken into account for its use for drug design [Fiser et al., 2001]. It should be stressed again that the low primary sequence identity between rhodopsin and other receptors is an issue that complicates GPCR modeling. This is a general drawback for efficient use of primary structure sequences of GPCR for homology modeling. In addition, rhodopsin holds the ligand covalently bound in the inactive conformation of the receptor and this is a different feature from other GPCR that should be taken into account when using rhodopsin as a template [Becker et al., 2003]. In spite of some reported limitations for the wide use of rhodopsin as a template, there are several outstanding homology modeling studies, supported by site-direct mutagenesis and ligand binding assays, in which ligand-receptor interactions derived from the 3D-crystal structure can be successfully applied to specific ligand design for GPCR. An example for the binding site of small molecules like biogenic amines is the β-2 adrenergic homology model [Green et al., 1996]. In this case amino acid residues in TM IV may be involved in binding of catecholamine antagonists.

The A1 adenosine receptor has been recently modeled using the rhodopsin crystallographic structure. This has allowed the proposal of a more realistic model than previously attempted because it has taken into account data from high throughput screening, SAR determination and mapping of contacts by site-directed mutagenesis. In this case the binding pocket would include conserved residues like D2.50 and S7.46 and other residues specific for the adenosine receptors subfamily like S1.46 and H7.43. H6.52 and H7.43 should be protonated at positions ε and δ respectively in order to agree with experimental data [Gutiérrez-de-Terán et al., 2004].

Some recent attempts at using the rhodopsin crystal structure as a template for modeling of other GPCR for application in drug design have encountered difficulties due in part to the overall lack of homology in the crucial ligand binding loop regions and the 'inactive' conformation of the rhodopsin structure. This may be the case for the cholecystokinin CCK1 receptor homology model based on the rhodopsin crystal structure, docked with its

natural ligand CCK in accordance with experimental data that identify contact points between CCK and its binding site on the CCK1 receptor. It was proposed that a refining process has to be carried out that takes into account the available biological, pharmacological and biophysical data in order to validate every step of the modeling process [Archer et al., 2003]. Thus, rhodopsin may not be an optimal model for certain GPCR, like that of CCK1, simply because these receptors have a number of structural divergences from rhodopsin that are necessary to ascertain their individual specificity [Sakmar, 2002]. A homology-based model of the human 5-HT$_{2A}$ receptor has been recently developed based on the structure in an attempt to solve this low homology problem. This model is based on the structure of the activated form of rhodopsin obtained *in silico* and it provides evidence that the ligand-receptor complex model is compatible with experimental data and docking studies with different ligands [Chambers and Nichols, 2002].

NEW MODELING APPROACHES

New approaches are recently being developed for the optimization of models. These include studies of 3D molecular modeling based on geometrical packing methods that allow optimization of individual helices in a structure of GPCR [Vaidehi et al., 2002]. In the case of models based on the structure of rhodopsin [Vaidehi et al., 2002], methods have been developed for predicting the ligand binding sites and relative binding affinities. This has been carried out for the β1-adrenergic receptor, the endothelial gene 6 receptor and different sensory receptors, like olfactory receptors and human sweet taste receptors. Modeling of the extramembranous loops is not very precise but this is not a major problem in drug discovery because the binding pocket for small ligands is presumable located in the TM domain.

Another study, which investigates homology-based receptor-ligand complexes, has been recently used for virtual screening of chemical databases [Bissantz et al., 2003]. Homology models based on the 2.8 Å-resolution X-ray structure of bovine rhodopsin, simulating an "antagonist-bound" form of three human GPCR (dopamine D3 receptor, muscarinic M1 receptor, and vasopressin V1a receptor) were constructed. Three different docking programs (Dock, FlexX,Gold), in combination with seven scoring functions (ChemScore, Dock, FlexX,

Fresno, Gold, Pmf, Score), were used. After testing the results, these were not as accurate as expected. Such a model was also used for an "agonist-bound" form but since the screening results were not satisfactory enough, manual refinements were applied to construct a pharmacophore-based modeling procedure rotating the TM VI simulating the conformational changes that occur during receptor activation. Three receptors, D3, B2 and opioid, were selected to perform these agonist-bound models for future agonist screening [Bissantz et al., 2003].

The methods that have been critically assessed use the structure of rhodopsin as a model, but due to the large variety of ligands for GPCR, novel approaches independent on the crystallization and structural determination of GPCR should also be devised. However, it is expected that in the next years a number of GPCR can be crystallized and their atomic structure determined. This would enormously facilitate new drug design and development. A recent innovative *in silico* screening method resulting in numerous novel compounds has also been proposed [Shacham et al., 2001]. This is the PREDICT algorithm that works at atomic levels to predict the 3D structures of GPCR without requiring a solved model for the prediction. The results obtained with this method were consistent whit the mutagenesis data available. Over 45 molecular models of members of the A family of GPCR have been constructed with this method, and ten of them are being used for the *in silico* screening process. This is a good starting point for drug discovery and lead optimization, for example the hit identification of a new compound for dopamine D2 receptor [Shacham et al., 2001; Stojanovic et al., 2004].

I.2.2.4. FUTURE PROSPECTS

The recent results on rhodopsin structure-function studies have been reviewed in the context of drug design for GPCR targets. The structural knowledge derived from the recent crystal structure of rhodopsin has allowed modeling of other receptors of the same family and has opened new avenues for the rational design of new drugs. Some of the different approaches for drug design have been also discussed, particularly those based on the

receptor and those based on the ligand, including some specific examples. Although the usefulness of rhodopsin structure-function studies in drug design approaches is certainly undisputed, the limitations of rhodopsin as a general model for the whole GPCR superfamily have been enumerated and highlighted. A reference has also been given to new approaches that try to avoid the intrinsic limitations of the rhodopsin model. The rhodopsin model will continue to be a useful reference template in the next future, particularly when new structural information is available. However, not a single approach seems to be of universal validity and in this regard a wise combination of using rhodopsin as a model and other approaches will certainly be valuable in drug design and development.