Unraveling the Structure and Function of G Protein-Coupled Receptors Through NMR Spectroscopy


Published in:
Current Pharmaceutical Design

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Unraveling the structure and function of G protein-coupled receptors through NMR spectroscopy

Irina G. Tikhonova and Stefano Costanzi*
Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, DHHS, Bethesda, MD 20892, USA

Abstract
G protein-coupled receptors (GPCRs) are a large superfamily of signaling proteins expressed on the plasma membrane. They are involved in a wide range of physiological processes and, therefore, are exploited as drug targets in a multitude of therapeutic areas. In this extent, knowledge of structural and functional properties of GPCRs may greatly facilitate rational design of modulator compounds. Solution and solid-state nuclear magnetic resonance (NMR) spectroscopy represents a powerful method to gather atomistic insights into protein structure and dynamics. In spite of the difficulties inherent the solution of the structure of membrane proteins through NMR, these methods have been successfully applied, sometimes in combination with molecular modeling, to the determination of the structure of GPCR fragments, the mapping of receptor-ligand interactions, and the study of the conformational changes associated with the activation of the receptors. In this review, we provide a summary of the NMR contributions to the study of the structure and function of GPCRs, also in light of the published crystal structures.

Keywords
G protein-coupled receptors (GPCRs); NMR; three-dimensional structures; ligand recognition; receptor activation

G protein-coupled receptors (GPCRs), also known as seven transmembrane-spanning receptors (7TMRs), are a large superfamily of signaling proteins expressed on the plasma membrane that function as receivers for extracellular stimuli [1]. About 1000 GPCRs have been identified in the human genome [2]. Since their signaling is involved in numerous physiological functions and pathological conditions, GPCRs constitute the molecular target of a significant percentage of the currently marketed drugs. Moreover, many additional members of the superfamily have been identified as potential targets for the treatment of a variety of diseases and are the object of substantial drug discovery efforts. From the molecular point of view, all GPCRs share a common molecular architecture, being composed of a single polypeptide chain folded into a bundle of seven α-helical transmembrane domains (TMs) connected by three extracellular and three intracellular loops (ELs and ILs). The N-terminus is located in the extracellular space, while the C-terminus is in the cytosol (Fig. (1)).

Given that structure-based drug discovery is an efficient method to rationally design novel drugs and improve the properties of old drugs, the scientific community has been striving for a long time to shed light onto the elusive structure-function relationships of GPCRs employing a variety of direct biophysical and indirect biochemical methods [3]. The most direct method that has been used is X-ray crystallography. However, up until now, this technique has been
applied successfully only to a limited number of receptors, namely rhodopsin, the unliganded opsin, the beta-adrenergic receptors (β-ARs), and the adenosine A2A receptor [4–13]. In particular, the low resolution projection map of bovine rhodopsin published by Schertler in 1993 [14] and the subsequent 2.8 Å resolution structure published by Paczewski in 2000 [15] provided the first static three-dimensional (3D) pictures of a GPCR, and have been widely used as templates for the construction of homology models. The subsequent recent publication of the X-ray structure of the β-ARs and the A2A receptor [6;7;10;11] proved conclusively that GPCRs share a common arrangement of the TMs, while suggesting a greater variability for the extracellular and intracellular regions. This includes the second extracellular loop (EL2), which, together with the portions of the TMs facing the extracellular side, is thought to line the ligand binding pocket for the great majority of GPCRs.

On the basis of the available crystallographic information, various studies reported insights into the structure and functions of GPCRs obtained through indirect methods such as site-directed mutagenesis, zinc crosslinking of histidines, photoaffinity labeling and site-specific chemical labeling (see as examples [16–22]). In this context, in silico molecular modeling conducted in an iterative manner with the gathering of experimental data became a widespread tool to infer the 3D structure of the receptors [23–25]. Like other researchers, we have successfully applied this approach to the identification of modulators of several GPCRs (see as examples [26–29]). Despite the relatively low sequence identity between rhodopsin and most other GPCRs, which instilled many doubts on the reliability and the scope of rhodopsin-based modeling, a recent comparison between in silico models and X-ray structure of the β2-AR firmly supported the applicability of homology modeling and molecular docking to the study of GPCR/ligand complexes [30].

Besides crystallography, other direct methods of investigation aimed at gathering atomic-resolution information have been also applied, among which electron crystallography, electron paramagnetic resonance, UV absorbance and fluorescence spectroscopy, and, as we will discuss at length in this review, nuclear magnetic resonance (NMR) spectroscopy (see as examples [14;31–38]).

NMR spectroscopy, conducted in solution or in solid-state, is a powerful method to study protein structure, protein dynamics, and protein-ligand interactions. NMR techniques are routinely used in drug discovery for soluble protein targets [39;40]. Although their application to GPCRs is hindered by difficulties related to the preparation of receptor samples and the interpretation of the complex spectra, many studies directed towards a structural characterization of GPCRs through NMR have been reported, providing useful information for structure-based drug discovery.

The methods and technical challenges underlying the application of NMR spectroscopy to membrane proteins and GPCRs have been covered in recently published reviews [41–43]. Here, we will focus on the advances in the understanding of the structure and function of GPCRs that have been made through NMR studies. In particular, the review is divided in three sections: in the first one, we will discuss attempts to determine 3D structures of GPCRs; in the second one, we will illustrate studies on receptor-ligand interactions conducted by labeling ligands and/or selected residues in the binding pocket; in the third section we will review studies intended to elucidate the motions and the structural changes consequent to the activation of the receptor, also conducted by labeling selected receptor residues.

1. Determination of the 3D structure of GPCRs

Many obstacles interfere with the determination of the complete 3D structure of a GPCR.

Structural studies by NMR require the assignment of all resonance picks in the spectra.

Since 1H is the only NMR active isotope abundantly present in proteins, a complete assignment
of the complex NMR spectra for a whole protein the size of a GPCR is contingent to the preparation of protein samples uniformly labeled with NMR active C and/or N isotopes.

NMR spectra are complicated by anisotropic nuclear spin interactions, such as dipole-dipole coupling, chemical shielding anisotropy, and quadrupole coupling. Solution-state NMR, which takes advantage of the averaging effect on anisotropic interactions due to the fast tumbling of the samples, is routinely applied to the analysis of proteins with a molecular weight between 30 and 50 kDa. Solid-state NMR spectra, ideal for membrane proteins such as GPCRs, are complicated by the restricted molecular movements that prevent the averaging of the anisotropic effects, inducing a peak-broadening that lowers significantly the resolution of the spectra [42;43].

The introduction of the magic-angle spinning and oriented-sample techniques substantially addressed these problems. However, the complexity of the solid-state NMR spectra and difficulties related to sample preparation have prevented, so far, the solution of the structure of whole GPCRs through NMR. For these reasons, most of the studies have been directed toward the solution of individual portions of the receptor structure either in aqueous solution or in different lipid-mimetic solvents [41]. In Table 1, we report a synoptic view of the available 3D structural data. Additionally, for the NMR-derived GPCR fragments with 3D coordinates available in the Protein Data Bank, we provide, in Fig. (2) and (3), a structural alignment with the X-ray structure of rhodopsin and the β-ARs.

Transmembrane domains

We now know from crystallographic data that some of the TMs present kinks in their α-helical structures, particularly pronounces in TM6 and TM7. These data emerged clear from NMR studies, even before the publication of GPCR X-ray structures. In particular, kinks were detected studying portions of TM6 of the Saccharomyces cerevisiae α-factor receptor (Ste2pR) [44] and TM7 of the tachykinin receptor NK1 [45] (Table 1). Solution and solid-state NMR studies of the TMs of the Ste2pR continued after the disclosure of the X-ray structure of rhodopsin and, among other findings, confirmed the existence of the helical kink of TM6 previously detected through solution NMR with a more sophisticated solid-state NMR of the 15N labeled TM6 peptide in 1,2-dimyristoyl-sn-glycero-phosphocholine bilayers [46–49]. As mentioned, the kinks of TM6 and TM7 postulated on the basis of the NMR results are consistent with the conformation of the homologous regions of the receptors that have now been crystallized (Fig. (2), panel A) [12;15], supporting the idea that even structural data gathered for isolated domains can be very informative. Probably, even greater insights could be obtained from the analyses of larger substructures or even whole receptors. At this purpose, Zheng and coworkers biosynthesized a quite large segment of the cannabinoid CB2 receptor, including TM1, IL1 and TM2, using a fusion protein overexpression strategy [50]. The preliminary results gathered by the authors for the 13C/15N labeled peptide argue in favor of the applicability of this strategy to the synthesis of GPCR helical bundles for NMR studies.

Extracellular loops

While homology modeling based on the available GPCR X-ray structures allows the construction of reliable models of the TMs, modeling the 3D structure of extracellular and intracellular regions remains a difficult task [30]. Low sequence similarity as well as numerous insertions and deletions, hamper an effective application of homology modeling to these regions. NMR spectroscopy could prove particularly useful in solving this problem: the structures of loops and termini of the receptor of interest could be derived through NMR and subsequently combined with homology models to build complete 3D structures.
To obtain structures of ILs and ELs close to the naturally occurring ones, many experimentalists turned to the application of structural constraints meant to keep the termini at a distance compatible with their role of connecting two TMs. In some cases, portions of the adjacent TMs have been added to act as anchors for shackleing loops or termini to the lipid environment. Alternatively, the loops have been constrained with disulfide bridges between cysteines [51] or S-carboxymethylcysteines [52] added to the peptide termini, with or without methylene linkers [53].

Since for most of the GPCRs activated by peptides ligand binding is thought to occur at the level of the extracellular domains, NMR analyses have been extensively applied to study receptor-peptide interactions. For example, the contact interactions of CCK8–the natural cholecystokinin peptide, composed of eight amino acids–with the N-terminus and EL3 of the CCK1 and with EL3 of the CCK2 subtypes of the cholecystokinin receptors have been studied by NMR by monitoring chemical shift perturbations and intermolecular NOEs [54–56] (Table 1). For the CCK2 the analysis has been complemented by integrating NMR-derived structural information with rhodopsin-based homology modeling. With a similar strategy, the interactions between the CCK1 and synthetic agonists have also been studied [57]. The integration of homology modeling and NMR-derived information provided also a structural hypothesis for the binding of substance P to the NK1 receptor, confirming and rationalizing the results previously obtained through a variety of indirect biochemical methods [58]. A similar approach also corroborated the postulated existence of a disulfide bridge between EL1 and EL2 in the thromboxane A2 receptor (TP), highlighting also the involvement in ligand binding of the WCF motif of EL2 [59;60].

Lastly, Pham and coworkers developed an alternative methodology for studying GPCR loops based on the computational design of a peptide containing segments that mimic the self assembling of the ends of two TMs connected by the loop under examination [61]. When applying their strategy to the analyses of EL1 of the sphingosine 1-phosphate receptor S1P4, the authors obtained the best results with the coiled-coil strategy, i.e. by substituting the TM2 and TM3 segments with sequences from structurally characterized water-soluble proteins with a pair of antiparallel helices oriented in a way consistent with the S1P4 homology model. The addition of a disulfide bridge appeared also necessary to keep the helices together. The ability of the peptide to bind an analog of the agonist’s headgroup suggested that EL1 adopted a biologically relevant conformation. Here, we superimposed this NMR-derived structure to the homologous regions in the crystal structure of the β2 receptor (Fig. (2), panel A) and, in agreement with the conclusions of the authors, revealed a sensible conformation and orientation for the TM2 and TM3 mimetics.

### Intracellular loops

Since the ILs, and IL2 in particular, are directly involved in the coupling of the receptors with the G proteins, [62–64] the experimental elucidation of the structure of these binding interfaces would provide insights into the molecular mechanisms of the selectivity toward the various G protein subtypes.

NMR studies of IL2 in the adrenergic α2A receptor, rhodopsin, the bradykinin B2 receptor, and the vaspressin V1a receptor have been performed (Table 1) [52;65;66]. In particular, NMR studies of the peptide corresponding to the IL2 of the α2A receptor revealed the presence of a cytoplasmic helix not detected in the analogous rhodopsin peptides [65] or in the crystal structure of rhodopsin. X-ray crystallography revealed that this helix, although not detected in the β2, is indeed present in the β1 subtype [7;11].

Conversely, the NMR-derived structure of IL2 of the B2 receptor revealed a “U” shape [66] similar to that detected in rhodopsin. Also in the case of the vaspressin V1A receptor, the
NMR analysis of a linear and a cyclic IL2 peptide revealed a conformation in TFE/H$_2$O solvent similar to that shown in rhodopsin, particularly with the cyclic peptide [52].

The structure of IL3 has also been determined for several receptors, including the cannabinoid CB$_1$ receptor, the β$_2$ receptor, and the parathyroid hormone PTH1 receptor (Table 1) [53;67;68]. Of note, the IL3 peptide of the CB$_1$ receptor (Fig. (2), panel A) assumed a helical conformation in aqueous solution when G$_{a1}$ was added, while a mutated IL3 formed just a single helixturn [68]. The authors proposed the importance of this portion of IL3 in the formation of the G protein binding interface.

Also for the ILs, hybrid solutions composed of molecular modeling and NMR studies have been applied. In particular, Wu and coworkers derived by NMR the structure of all three ILs of the thromboxane A$_2$ receptor and subsequently incorporated them into a rhodopsin-based homology model [51]. They consequently proposed a list of amino acid residues potentially involved in the formation of the G protein-binding interface that proved consistent with available mutagenesis data.

**N-/C- termini**

In addition to the mentioned NMR structure of the CCK$_1$ N-terminus, the structure of the N-terminus of PTH1 receptor has also been solved [69], revealing to be different from that of rhodopsin (Fig. (2), panel A). These data confirm the idea that each GPCR may be characterized by unique extracellular regions that contribute to the selectivity of the receptors for specific ligands.

All the available GPCR X-ray structures are characterized by the presence of αα-helix, known as helix 8 (H8), in the portion of the C-terminus proximal to TM7. NMR in conjunction with CD spectroscopy confirmed the presence of an H8 for the CB$_1$ and CB$_2$ receptors, the Ste2p receptor, the β$_1$ and β$_2$ receptors, the angiotensin AT$_{1A}$ receptor, and the bradykinin B$_2$ receptor in detergent [70–75]. In Fig. (2), panel B, we superimposed the NMR-derived the β$_1$ receptor H8 to the corresponding segment of the crystal structure of the same receptor, revealing a very good agreement between crystallographic and NMR data. Conversely, a random coil conformation was detected in water for this region [71–73]. These conformational changes of H8 have been speculated to be indicative of possible structural rearrangements consequent to the activation of the receptors. For instance, partial unfolding of H8 upon light activation has been detected in rhodopsin by Fourier transform infrared (FTIR) and fluorescence spectroscopy [76].

Several NMR studies of the rhodopsin C-terminus have been published [33;77;78]. Beyond H8, this region resulted unstructured in aqueous solution, both in unphosphorylated or phosphorylated form, while has adopted a defined 3D structure when phosphorylated and in the presence of arrestin (Fig. (2), panel C) [78]. Here we could only compare this NMR-derived structure (pdb code: 1nzs) with the crystal structures of lumirhodopsin and batorhodopsin (pdb code: 2hpy and 2g87), for which this distal portion of the C-terminus has been determined. The detected fairly high root mean square deviations of 5.5–6 Å suggest that the C-terminus of rhodopsin undergoes a conformational change when bound to arrestin.

**Yeagle’s approach to derive the structure of rhodopsin through NMR**

Before the obtainment of high resolution X-ray structures of GPCRs, Yeagle and coworkers devised a methodology intended to solve the 3D structure of portions of the receptors and subsequently assemble them together. Through two dimensional homonuclear $^1$H NMR analyses in solution, the authors determined the structure of overlapping peptides spanning the entire sequence of the rhodopsin and subsequently computationally assembled the fragments.
into a single structure [79–81]. The information about the packing of the helical bundle, necessary to put the pieces together, was gathered from electron paramagnetic spectroscopy (EPS) [36;82;83], zinc cross-linking of histidines [16], and electron cryo-microscopy data [84]. In the form of distance constraints, this was then applied to the NMR-derived fragments through simulated annealing. Using different sets of experimental data collected for the ground-state and activated (META II) rhodopsin, the authors intended to generate models for both states (PDB codes: 1jfp and 1ln6). After the publication of the crystal structure of rhodopsin, the authors discovered that the largest TM peptides yielded the best superimposition with the corresponding segments of the crystal structure [85].

Here we superimposed Yeagle’s ground state rhodopsin to the crystal structures of the ground state rhodopsin (PDB code: gzm), detecting a RMSD of the backbone of about 6 Å (Fig. (3), panel A). The packing of the TM helical bundle is similar in the two structures, however, Yeagle’s model predicts the presence of several non helical portions, in particular in TM7, not confirmed by the crystallography. The intracellular and extracellular regions appear significantly divergent in the Yeagle’s model and the crystal structure. We also superimposed Yeagle’s Meta II rhodopsin to the recently solved crystal structure of opsin in its G protein interacting conformation [8], detecting, also in this case, an RMSD of the backbone of about 6 Å (Fig. (3), panel B).

Whole GPCRs

Despite the difficulties, several attempts to solve the structure of whole GPCRs by NMR are currently being undertaken. Solution NMR spectroscopy experiments, conducted for the uniformly $^{15}$N-labelled vasopressin V$_2$ receptor, solubilized with lysomyristoylphosphatidylcholine, indicated the potential suitability of the technique for structural studies of GPCRs. Notably, the authors observed over 250 amide peaks out of the expected 349 in their $^1$H,$^{15}$N-TROSY spectrum [86]. Additionally, solid-state NMR studies of a uniformly $^{15}$N-labeled and selectively $^{15}$N-Ile-labeled chemokine CXCR1 receptor in magnetically aligned bicelles provided spectra that suggested the potential applicability of NMR to structural determinations and the elucidation of structure-activity relationships (vide infra) [32].

2. Mapping receptor-ligand interactions

Direct atomic resolution information on the interactions of GPCR with their ligands is available through X-ray crystallography only for a limited number of receptors, while, for the great majority of them, indirect methods of analysis have been used. A synergistic application of mutagenesis studies, photoaffinity labeling, chemical modifications of the ligands, and computational modeling has been the most common way of studying the determinants of ligand recognition. Taken together, these results have contributed to the general definition of a common binding pocket for GPCRs located towards the extracellular opening of the helical bundle [87;88]. Also for receptors naturally stimulated by ligands that bind to their N-terminal ectodomains, such as the calcium sensing or the glycoprotein hormone receptors, the possibility of modulation through ligands that bind within their helical bundle has been demonstrated [89–93].

In this context, structure-activity relationships (SAR) analyses and drug-discovery studies would greatly benefit from direct experimental proofs of receptor-ligand interactions, such as those that NMR could provide. Solid-state NMR spectroscopy has been applied successfully to membrane proteins to detect ligand binding and to analyze protein-ligand interactions, especially when coupled to selective isotopic labeling of specific residues located in the binding pocket and/or of the ligand [94]. Through this expedient, significantly simplified spectra with a manageable amount of signals have been obtained, thus rendering NMR applicable to the
analysis of bioactive ligand conformations, protonation states, receptor/ligand interactions and residue/residue interactions. Moreover, although not yet for GPCRs, NMR coupled to selective isotopic labeling has successfully been applied to rational fragment-based drug discovery. In particular, ligands or fragments that bind to specific pockets of a protein have been identified by monitoring the changes in the chemical shifts of $^{15}$N or $^1$H-amide atoms in $^{15}$N-labelled protein, an approach known as SAR by NMR [95;96].

Rhodopsin has been extensively studied by solid-state NMR coupled to selective labeling. Before any high-resolution crystal structures were released, selective $^{15}$N-labeling of Lys residues led to the conclusion that the distance between the protonated Schiff base by which the retinylidene chromophore is covalently bound to K296 (7.43 according to the Ballesteros and Weinstein residue indexing [97]) and the E113 (3.28) counterion (Fig. (4)) was greater than 4 Å [98]. The subsequent publication of crystal structures, revealed that the distance is, in fact, about 3.5 Å.

The conformation of retinal and its interactions with rhodopsin have also been studied extensively through NMR coupled to selective labeling. For example, high-resolution solid state deuterium ($^2$H) NMR experiments provided detailed information on the orientation of retinal in the binding pocket and on the conformational changes that occur with the transition from the ground state to the Meta I intermediate of the activation cycle. The study applied three retinal analogues labeled with deuterium at three different pairs of adjacent carbons [99]. Ultra high field solid-state magic angle spinning 2D homonuclear and 2D heteronuclear NMR spectroscopy has also been used to study rhodopsin reconstituted with a uniformly $^{13}$C-labeled 11-cis-retinal. Complete assignment of the $^{13}$C and $^1$H chemical shifts for retinal highlighted nonbonding interactions between the protons of the methyl groups of its ionone ring and the nearby aromatic acid residues F208 (5.43), F212 (5.47), and W265 (6.48) (Fig. (4)). Furthermore, it was shown that binding of retinal involves a chiral selection of the ring conformation, resulting in equatorial and axial positions for CH$_3$-16 and CH$_3$-17 [100]. Additional evidences of the interactions between retinal and rhodopsin in the ground and Meta I states have also been obtained through NMR experiments conducted variously labeling the ligand with $^2$H [101–104]. These data support the hypothesis that a strain in the polypeptide around the cis bond assists the photoisomerization of retinal. A $^{13}$C-labelled 9-cis retinal isomer, typical of isorhodopsin, has also been investigated, leading to similar conclusions [105].

Application to rhodopsin of an NMR technique named by the authors “selective interface detection spectroscopy” (SIDY), based on the detection of the correlations between $^{13}$C atoms of labeled ligands and $^1$H atoms of unlabelled receptors, led to the detection of several contacts between the aliphatic carbons of the 11-cis-retinal ionone ring and residues in the binding pocket. Although, SIDY data do not provide sequence-specific assignments of the contacts, these could be identified by means of additional data, and, in the case of rhodopsin, resulted in good agreement with the available crystallographic structures [106].

NMR spectroscopy provides also a very effective way of studying the conformational changes that a ligand undergoes upon binding to a receptor. In this context, the analysis of the unbound and receptor-bound $^{13}$C, $^{15}$N-labelled neurotensin, a 13-residue peptide, revealed that the ligand is in a disordered state in the absence of the receptor, while adopts a beta-strand conformation when bound to the NTS$_1$ receptor [107]. Similarly, the structure of the bradykinin (BK) peptide bound to the bradykinin B$_2$ receptor has been studied through solid-state NMR, revealing a double S-shape structure [108]. These type of studies can be also conducted to rationalize the biological activities of compounds on the basis of their bioactive conformations, as it has been done in the case of two natural peptides active at the neuropeptide NPR-1 receptor one [109].
In addition to conformational changes, NMR spectroscopy can also detect changes in the protonation state of a ligand as a consequence of binding to a receptor. In this context, solid state NMR experiments conducted with uniformly labeled $^{13}$C,$^{15}$N-histamine bound to the human histamine H$_1$ receptor revealed that the ligand can bind either in a monocationic or a dicaticionic form. On the basis of their data, the authors hypothesized that, in analogy with rhodopsin, also in the case of the H$_1$ receptor a protonation switch might be part of the activation mechanism [110].

3. Detection of motions and conformational changes associated with receptor activation

Upon binding of agonists, which typically occurs in proximity of the extracellular opening of the helical bundle, GPCRs undergo a series of structural changes that cascade from the extracellular to the intracellular part of the receptor and ultimately lead to G protein activation. Detailed structural knowledge of the ground and activated state of the receptor and mechanistic insights into the activation process would significantly assist drug discovery, allowing a more rational design of compounds capable of stimulating or blocking the receptor.

A wealth of information has been derived by the analysis of mutations that affect the basal activity of GPCR, either naturally occurring or generated through site-directed mutagenesis. Mutations that cause an increase of the basal activity are likely to stabilize the activated state of the receptor and are referred to as constitutively active mutants. Those that prevent the activation of the receptor are, instead, called uncoupling mutants and are likely to stabilize its inactive state or to disrupt the cascade of conformational changes that lead to signaling. Through such mutational analyses, various molecular switches for GPCR activation have been proposed [111;112]. Alternatively, the structural changes that occur upon receptor activation have been monitored through electron paramagnetic resonance spectroscopy (EPR) [34;36; 83;113], UV absorbance spectroscopy [35], engineering of metal-ion-binding sites [16], as well as site-specific chemical labeling coupled to fluorescence spectroscopy [37], and, as we will see in the coming paragraphs, NMR spectroscopy. These data shed light onto a number of intramolecular distances that are characteristic of the activated receptors. We recently used these biophysically measured distances to construct a computational structure of the activated rhodopsin based on coarse-grained and all atom simulations, and subsequently study the dynamics of the activation process [114]. The substantial conformational changes predicted by these indirect methods were not detected in the crystal structure of a photoactivated deprotonated intermediate (PDI) of rhodopsin that shows absorption maxima consistent with the META II state [13]. The crystal lattice may have prevented large-scale structural rearrangements. However, a significant opening of the intracellular surface of the receptor and a furthering of the intracellular ends of TM3 and TM6 have been captured in the crystal structure of opsins with a fragment of the C-terminus of transducin. At least relatively to the cytosolic half, this may represent the first detailed structure of an activated GPCR [8].

Compared to crystallography, NMR spectroscopy offers the possibility of conducting structural analyses in a less constrained membrane-like environment. In this context, NMR is well suited to the study of receptor activation through isotopic labeling of specific residues located in areas affected by the structural changes. In this context, the specific labeling of rhodopsin led to observance of the transition between its inactive and activated states.

$^{12}$D-dipolar-assisted rotational resonance NMR measurements between $^{13}$C-labels on the C14, C15, C19, C20 atoms of retinal and $^{13}$C-labeled G114(3.29), T118(3.33), G121(3.36), Y178(4.68), G188(EL2), Y191(EL2), S196(EL2), and Y268(6.51), all located in the retinal binding pocket, was performed for the ground state and META II rhodopsin (Fig. (5), panel A, purple and green spheres indicate retinal and rhodopsin atoms, respectively) [115]. The results...
highlighted that the extracellular portion of the receptor undergoes conformational changes upon activation. In particular, a large rotation of the C20 methyl group of retinal toward EL2 and also a translation of about 4–5 Å of the retinal chromophore toward TM5 were observed. This displacement of retinal has been associated with the motions of the TM5, TM6, and TM7 predicted by different biophysical and biochemical methods [115]. Subsequent measurements between the C19 and C20-methyl groups of retinal and $^{13}$C-labeled W265(6.48) confirmed that, in agreement with what suggested by atomic microscopy [38], the movement of the betainion ring causes a reorientation of the side chain of W265(6.48) upon activation (Fig. (5), panel A, a blue sphere indicates W265). In agreement with what seen in the crystal structure of opsin [8], the study also suggested that TM6, but not TM3, undergoes a significant rearrangement [116]. Most likely, the local movements triggered by the isomerization of retinal cause a perturbation of the complex network of hydrogen bonds that stabilizes the helical bundle of the ground state rhodopsin. This is suggested also by the comparison of the chemical shifts of $^{15}$N and $^{13}$C-labeled wild type and mutant rhodopsin in the ground and META II states [117], which, among other observations, led to the detection of the disruption of the strong hydrogen bond between Glu 122 in TM3 and H211 in TM5 (Fig. (5), pink spheres indicate Glu122 and H211), thus indicating a motion of TM5. A more recent study demonstrated that the transition to META II is accompanied also by a displacement of EL2 from the retinal binding site, and suggested that this displacement is coupled to the rotation of TM5 and the breakage of the ionic lock connecting TM3 and TM6 [118]. As mentioned, all these interconnected conformational changes caused by the photoisomerization of retinal eventually cascade down towards the cytosolic portion of the TMs, thus leading to G protein activation. To specifically monitor the changes that occur in this region, NMR spectroscopy of $^{19}$F labeled rhodopsin has been applied after mutation of specific residues to Cys and subsequent attachment of trifluoroethylthio groups via disulfide bridges (C65, C139, C140, C251, C316, indicated by yellow spheres in Fig. (5), panel A). The results support the idea that the tertiary structure of the cytoplasmic face of the receptor changes significantly upon light activation [119;120].

Beyond the activation process, NMR analyses of $^{15}$N labeled rhodopsin have been applied also to the study of the dynamics of the dark adapted receptor. In particular, a study on α,$^{15}$N-Lys-labeled rhodopsin revealed that, while the single Lys residue located in the C-terminus is endowed with nanosecond scale movements, those located in different regions of the receptors present micro- to millisecond timescale motions (Fig. (5), panel B). In contrast with Lys residues, which in rhodopsin are located mostly in the cytosolic domains, Trp residues are, for the most part, present in the TMs (Fig. (5), panel B). A subsequent study of α and ε,$^{15}$N-Trp-labeled rhodopsin, confirmed that there are micro- to millisecond timescale backbone motions in the inactive dark state, while suggesting a substantial restriction of the Trp side chains to a single specific conformation [121;122].

Conclusion

Despite the numerous technical challenges, NMR spectroscopy has the potential of becoming a leading technique in the study of the structure-function relationships of GPCRs and their interactions with ligands. In contrast to crystallographic methods, which provide high-resolution but static molecular snapshots, NMR techniques can provide dynamic pictures of receptor structures and receptor-ligand interactions, thus offering insights into the molecular mechanisms of ligand recognition and receptor activation.

Through this article, we reviewed many of the contributions that NMR studies have given to the understanding of the molecular structure and functioning of GPCRs. In particular, NMR spectroscopy has been extensively applied to the study of individual portions of GPCRs, also in combination with homology modeling. Through selective labeling of ligands and/or specific
receptor residues, NMR studies have also been applied to the study of GPCR-ligand interactions and to the investigations of the molecular changes coupled to activation of the receptor.

Recently, dynamic nuclear polarization (DNP), a technique intended to increase the NMR sensitivity by transferring the polarization of electron spins to nuclei, has been successfully applied to the study the photocycle of bacteriorhodopsin, a 7TM light-driven ion pump with some structural analogies to rhodopsin [123]. Due to the enhanced sensitivity of the experiments, DNP allows the analysis of low concentration samples in a lower timeframe, and the detection of low populated conformations.

The introduction of DNP, as well as other advances in solution and solid-state NMR technology, together with progresses in the preparation of GPCR samples, are expected to lead in the future to the determination of the 3D structure of whole GPCRs, and to provide further mechanistic insights into the complex cascade of motions and rearrangements characteristic of the activation process. Furthermore, these methodological advances are also likely to foster the application of SAR by NMR techniques to the discovery of lead compounds for GPCRs through NMR-based high throughput screenings and their subsequent structure-based optimization.

Acknowledgments

This research was supported by the Intramural Research Program of the NIH, National Institute of Diabetes and Digestive and Kidney Diseases.

References


Figure 1.
The topology of G protein coupled receptors.
Figure 2.
Structural alignment of the NMR coordinates of GPCR portions deposited in the PDB and the crystal structures of the and bovine rhodopsin. Panel A: superimposition of the β2-AR crystal structure (2RH1, red) with the NMR-derived structures of the N-termini of PTH₁ (1BL1–white) and CCK₁ (1D6G, orange), EL1 of S₁P₄ (2DCO, aquamarine), IL3 of CB₁ (1LVQ, light blue), TM6 of Ste2pR (1PJD, yellow), and EL3 of CCK₁ (1HZN, pink) and CCK₂ (1L4T, dark blue). Panel B: superimposition of the crystal structure of the β₁-AR (2VT4, plum) with the NMR-derived structure of H₈ of the same receptor (1DEP, green). Panel C: superimposition of the crystal structure of rhodopsin (2HPY, white) and the NMR-derived structure of the C-terminus of the same receptor (1NZS, blue purple). Pictures prepared with Maestro 8.0.308, Schrodinger.
Figure 3.
Superimposition of the crystal structures and the NMR-based three-dimensional models of bovine rhodopsin. Panel A: Superimposition of the X-ray structure (1GZM, white) and the NMR-based model of the ground state (1LN6, purple) receptor. Panel B: Superimposition of the X-ray structure of opsin in its G-protein-interacting conformation (3DQB, orange) and NMR-based model of Meta II rhodopsin (1JFP, green). Picture prepared with Maestro 8.0.308, Schrodinger.
Figure 4.
Two dimensional schematic representation of the retinal binding site in ground state rhodopsin. Picture prepared with MOE 2008.10, Chemical Computing Group.
Figure 5.
Panel A: Positions of the ligand atoms and residues labeled in order to study the activation-related conformational changes of rhodopsin. The retinal atoms (C5, C9, C11, C13, C14, C15, C19, C20) are in purple; residues 114, 118, 121, 178, 188, 191, 196, and 268 are in green; residue 265 is in blue; residues 122 and 211 are in pink; the residues mutated to Cys to be labeled with trifluoroethylthio groups are in yellow. For simplicity, only one carbon of the residue backbone is shown. Panel B: All Trp and Lys residues of rhodopsin. Picture prepared with Maestro 8.0.308, Schrodinger.
A synoptic view of the available NMR structural data for GPCRs. A list of the abbreviations is provided in footnote\textsuperscript{a}.

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Source</th>
<th>Domain</th>
<th>Sequence</th>
<th>Constraints</th>
<th>Solvent</th>
<th>Optimization</th>
<th>Summary of Results</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>β\textsubscript{2} adrenergic receptor</td>
<td>turkey</td>
<td>IL3</td>
<td>284–295</td>
<td>TMA</td>
<td>LPC</td>
<td>DPC</td>
<td>290–294: α-helix and flexible Nt in LPC; α-helix in TFE; unstructured in AS; random coil in TFE/AS.</td>
<td>[67]</td>
</tr>
<tr>
<td></td>
<td>human</td>
<td>Helix 8</td>
<td>324–357</td>
<td>AS</td>
<td>DMSO</td>
<td>SA/MD</td>
<td>Unstructured in AS. α-helix in DMSO and DPC.</td>
<td>[72]</td>
</tr>
<tr>
<td>β\textsubscript{1} adrenergic receptor</td>
<td>turkey</td>
<td>Helix 8</td>
<td>345–359</td>
<td>LMPC</td>
<td>DPC</td>
<td>DG/MD</td>
<td>α-helix (PDB code: 1dep).</td>
<td>[73]</td>
</tr>
<tr>
<td>B\textsubscript{2} bradykinin receptor</td>
<td>rat</td>
<td>Ct</td>
<td>309–366</td>
<td>DPC</td>
<td></td>
<td></td>
<td>311–326, 333–345, and 348–363: α-helices.</td>
<td>[75]</td>
</tr>
<tr>
<td>CB\textsubscript{1} cannabinoid receptor</td>
<td>human</td>
<td>IL3</td>
<td>338–346</td>
<td>AS</td>
<td>DG</td>
<td></td>
<td>Unstructured in AS; α-helix with G\textsubscript{0i} (PDB code: 1lvq).</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td>Helix 8</td>
<td>397–418</td>
<td>TMA</td>
<td>SDS</td>
<td>DG/MD</td>
<td></td>
<td>Unstructured in AS; α-helix in DPC/SDS.</td>
<td>[71]</td>
</tr>
<tr>
<td>CB\textsubscript{2} cannabinoid receptor</td>
<td>human</td>
<td>TM1-IL1-TM2</td>
<td>27–101</td>
<td>DMSO</td>
<td></td>
<td></td>
<td>Preliminary results</td>
<td>[50]</td>
</tr>
<tr>
<td></td>
<td>Helix 8</td>
<td>298–319</td>
<td>AS</td>
<td>DMSO</td>
<td>SA/MD</td>
<td></td>
<td>Unstructured in AS; α-helix in DMSO and DPC.</td>
<td>[70]</td>
</tr>
<tr>
<td>CCK\textsubscript{2} cholecystokinin receptor</td>
<td>human</td>
<td>EL3</td>
<td>352–379</td>
<td>TMA</td>
<td>DPC/AS</td>
<td>DG/MD-wd</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receptor</td>
<td>Source</td>
<td>Domain</td>
<td>Sequence</td>
<td>Constraints</td>
<td>Solvent</td>
<td>Optimization</td>
<td>Summary of Results</td>
<td>Ref.</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>----------</td>
<td>-------------</td>
<td>----------</td>
<td>--------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>α-Factor receptor (Ste2pR)</td>
<td><em>S. cerevisiae</em></td>
<td>TM6</td>
<td>252–269, C252A</td>
<td>TFE/AS</td>
<td>DG/EM</td>
<td>α-helix with a proline kink.</td>
<td>[44]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TM6</td>
<td>252–269, C252A</td>
<td>DMPC</td>
<td></td>
<td>α-helix with a proline kink (PDB code: 1pjd).</td>
<td>[49]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TM7, EL3, Cs</td>
<td>267–339</td>
<td>TFE/AS CDCl3/CD3 OH/AS LPC DPC</td>
<td>DG/SA</td>
<td>High helical content of TM7 and Ct in TFE/AS and CDCl3/CD3 OH/AS. Sample stability in detergents was not insufficient for NMR experiments.</td>
<td>[48]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EL2</td>
<td>162–198</td>
<td>TMA</td>
<td>MD-wd</td>
<td>176–182: α-helix. 187–189: turn.</td>
<td>[125]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ct</td>
<td>330–348</td>
<td>AS</td>
<td></td>
<td>Unstructured peptide.</td>
<td>[33]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphorilated Ct</td>
<td>330–348</td>
<td>AS/Arrestin</td>
<td>DG/MD</td>
<td>Unstructured peptide in AS. The N-terminal portion is flexible and the C-terminal portion assumes a stable helix-loop structure in the presence of arrestin (PDB code: 1nzs).</td>
<td>[77,78,127]</td>
<td></td>
</tr>
<tr>
<td>V1a Vaspressin receptor</td>
<td>human</td>
<td>IL2</td>
<td>146–166</td>
<td>LF, CF1</td>
<td>AS</td>
<td>Random coil conformation in AS; α-helix followed by a β-turn in TFE/AS</td>
<td>[52]</td>
<td></td>
</tr>
<tr>
<td>S1P1 Sphingosine 1-phosphate receptor</td>
<td>human</td>
<td>EL1</td>
<td>95–125</td>
<td>DB</td>
<td>TMA</td>
<td>Partially helical structure (PDB code: 2dco).</td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>Receptor</td>
<td>Source</td>
<td>Domain</td>
<td>Sequence</td>
<td>Constraints</td>
<td>Solvent</td>
<td>Optimization</td>
<td>Summary of Results</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------</td>
<td>--------</td>
<td>----------</td>
<td>-------------</td>
<td>------------------</td>
<td>--------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>NK₁ tachykinin receptor</td>
<td>TM7</td>
<td></td>
<td>AMSSTM</td>
<td>YNPISHS</td>
<td>DMSO</td>
<td>EM/MD</td>
<td>Kinked α-helix as a minor conformer in DMSO; increased folded content in DMSO/CDCl₃; two-α-helical regions linked by a hinge around the NP motif in PFTB/CD₃OD.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L</td>
<td></td>
<td>DMSO/CD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L</td>
<td></td>
<td>DM/OD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L</td>
<td></td>
<td>DMSO/OD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L</td>
<td></td>
<td>DMSO/Cl₃</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A₂ tromboxane receptor</td>
<td>human</td>
<td>EL1</td>
<td>88–104</td>
<td>DB</td>
<td>AS</td>
<td>DG/MD</td>
<td>β-turns.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>173–193</td>
<td>AS</td>
<td>DG/MD</td>
<td></td>
<td>A two-turns loop in correspondence of 183 and 188.</td>
<td></td>
</tr>
</tbody>
</table>

---