NMR analysis of GPCR conformational landscapes and dynamics
Marina Casiraghi, Elodie Point, Alexandre Pozza, Karine Moncoq, Jean-Louis Banères, Laurent Catoire

To cite this version:

HAL Id: hal-03011639
https://hal.archives-ouvertes.fr/hal-03011639
Submitted on 21 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
NMR analysis of GPCR conformational landscapes and dynamics

Marina Casiraghi,¹, # Elodie Point,¹ Alexandre Pozza,¹ Karine Moncoq,¹ Jean-Louis Banères,² Laurent J. Catoire¹,*

¹Laboratoire de Biologie Physico-Chimique des Proteines Membranaires, UMR7099, CNRS/Université Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique (FRC 550), 13 rue Pierre et Marie Curie, 75005 Paris, France; ²Institut des Biomolécules Max Mousseron (IBMM), UMR 5247 CNRS, Université Montpellier, ENSCM, 15 av. Charles Flahaut, 34093 Montpellier, France. #Current address : Department of Molecular and Cellular Physiology, Stanford University School of Medicine, 279 Campus Drive, 94305 Stanford California, USA

*correspondence to : laurent.catoire@ibpc.fr

© 2019 published by Elsevier. This manuscript is made available under the CC BY NC user license https://creativecommons.org/licenses/by-nc/4.0/
Abstract

Understanding the signal transduction mechanism mediated by the G Protein-Coupled Receptors (GPCRs) in eukaryote cells represents one of the main issues in modern biology. At the molecular level, various biophysical approaches have provided important insights on the functional plasticity of these complex allosteric machines. In this context, X-ray crystal structures published during the last decade represent a major breakthrough in GPCR structural biology, delivering important information on the activation process of these receptors through the description of the three-dimensional organization of their active and inactive states. In complement to crystals and cryo-electronic microscopy structures, information on the probability of existence of different GPCR conformations and the dynamic barriers separating those structural sub-states is required to better understand GPCR function. Among the panel of techniques available, nuclear magnetic resonance (NMR) spectroscopy represents a powerful tool to characterize both conformational landscapes and dynamics. Here, we will outline the potential of NMR to address such biological questions, and we will illustrate the functional insights that NMR has brought in the field of GPCRs in the recent years.
1. Introduction

GPCRs are the largest family of cell-surface receptors, with more than 800 members identified in the human genome. They are integral membrane proteins constituted by 7 trans-membrane (TM) spanning helices (Fig. 1) which mediate the signal transduction from outside to inside the cell, translating stimuli from a broad panel of ligands (neurotransmitters, hormones, ions, photons, etc.) into biochemical responses (Rosenbaum et al., 2009). Ligands with different efficacies are responsible for the fine-tuning of GPCR signaling, as they can block, activate or modulate the activity of the receptors (Deupi and Kobilka, 2010). The diversity of GPCR-mediated biological outputs results from the ability of most receptors to activate G protein dependent and independent pathways, as a single GPCR can activate different G protein sub-types or other effectors such as arrestins (Luttrell and Kenakin, 2011). Due to their prominent role in physiological processes, GPCRs represent the target of around 30% of the drugs currently on the market (Rask-Andersen et al., 2014; Santos et al., 2017; Hauser et al., 2017), addressing a broad range of pathologies including neuropsychiatric, cardiovascular, pulmonary and metabolic disorders, cancer, obesity and AIDS (Lagerström and Schioth, 2008; Bockaert et al., 2010; Persaud, 2017; Nieto Gutierrez and McDonald, 2018). Despite recent major advances in the field, there is still an important lack of knowledge regarding GPCR signaling mechanism in association with their manifold pharmacology. This implies that GPCRs are the object of intense study in order to characterize the relationship between their structure, conformational dynamics and function, with important implications for drug design.

Numerous biological functions are closely related to changes in spatial and temporal locations of atoms inside biomolecules. The GPCR-mediated signal transduction phenomenon through the membrane of eukaryotic cells underlies the existence of a certain conformational plasticity that has yet to be understood. Indeed, the functional versatility of GPCRs cannot be explained by a simple two-state model of activation considering only an active and an inactive state, as these receptors are highly dynamic and can explore a wide range of conformations (Manglik and Kobilka, 2014; Latorraca et al., 2017). GPCRs are able to modify and adapt their shape in response to ligands with different pharmacological profiles, through a series of structural rearrangements that are crucial for the binding of various downstream intracellular partners (Urban et al., 2006). The ensemble of conformations that a receptor can adopt can be defined with the concept of conformational landscape (Frauenfelder et al., 1991), which describes how highly dynamic proteins like GPCRs can exist in an ensemble of conformations in the presence or absence of ligands. In this energy landscape, the ground states, or low-energy conformations, correspond to the global minimum of energy, an ensemble of conformers separated by small kinetic barriers, that are easily overcome in the range of motions in the picosecond to nanosecond (ps-to-ns) timescales (Fig. 2). While most biophysical techniques are able to sample the highly probable low-energy states, higher energy conformers, also known as excited states, are more difficult to study due to the difference in free energies between the ground and the excited states (Boltzmann law). Moreover, the lifetime of those states can be very short in the case of unfavorable activation energy from the ground to the excited state. Under the assumption of a conformational selection process triggering receptor activation, pre-existing conformations or sub-states around the basal state (tiers 1 and 2 in Fig. 2) may explain why different ligands are able to bind a same GPCR, as each conformation would have a differential propensity to interact with a given ligand compared to the others. The ligand binding event would subsequently lead to the selection and/or selective stabilization of a given conformation, thus causing a population shift in the conformational ensemble that would favor the interaction with a specific intracellular partner and therefore engage a biological response. To get a full picture of the functional implications of these different conformers, we need to fully characterize the conformations that populate the energy landscape and their associated relative probabilities (thermodynamics) and the dynamic barriers separating these different conformations during the activation and inactivation of the receptors (kinetics) (Henzler-Wildman and Kern, 2007; Smock and Giersch, 2009).

2. NMR: a powerful tool to investigate energy landscapes at the atomic scale

NMR is one of the most powerful techniques to investigate the energy landscapes of biomolecules, with a spectrum of different methodologies to address i) conformational ensembles, including hidden or sparsely and transiently populated conformers (i.e. with a relative population below 0.5% and a lifetime of less 1 millisecond), and ii) timescales of motions from ps to s and beyond. To do so, NMR can interrogate a vast
number of reporters within the protein, the NMR active nuclei, which are $^1$H, $^{15}$N and $^{13}$C in proteins, and also $^{19}$F thanks to the use of chemical modifications. The position of a signal, named chemical shift, represents an atom of the receptor in the NMR spectrum, and is dependent on the micro-environment that this atom experiences within the protein. In proteins, $^1$H, $^{15}$N and $^{13}$C nuclei may experience a variety of environments that depend on both local and global protein structure. As a result, depending on the localization of the nucleus, the position of the NMR peak can be very sensitive to changes in protein conformation and to the rates of chemical exchange between the different conformations sampled by the protein (Fig. 3). Thus, it is possible to directly observe several co-existing conformations or variations in the conformational ensemble, if these conformations are populated enough and provided that the inter-converting rates of exchange between the different conformers are slow enough compared to the window of NMR observation (directly related to the value of the static magnetic field) experienced by the nucleus under investigation. In other words, two peaks in a spectrum corresponding to two distinct conformations, like in Fig. 3 (right panel), can be observed if the difference in chemical shifts between these two peaks is much greater than the rate of chemical exchange.

NMR is a very useful tool to study the chemical exchanges towards lowly populated (<5%) and transiently formed conformations (few milliseconds at most). This typically concerns excited-state conformers separated by high kinetic barriers that are usually transient to the traditional biophysical methods (Baldwin and Kay, 2009). A number of NMR approaches have been developed to characterize these invisible states (Palmer et al., 2001; Hansen et al., 2008; Clore, 2011; Fawzi et al., 2011; Vallurupalli et al., 2012). These methods are based on the use of the relaxation properties of the spin under study that are sensitive to the chemical exchanges (Mittermaier and Kay, 2009). The use of NMR relaxation experiments allows also to quantitatively determine the kinetic barriers separating the different conformations that occur on different timescales by determining the rates of inter-conversion and the populations involved (Palmer et al., 2005). Importantly, when repeated at equilibrium at various temperatures, these NMR experiments open the possibility to i) determine the enthalpic and entropic contributions to the different chemical exchanges detected in the receptor energy landscape and, ii) measure the forward and backward energies of activation involved in these different processes (Kay 2005).

### 3. The use of artificial membrane environments

Atomic-scale NMR studies of membrane proteins (MPs), and GPCRs in particular, are performed in vitro with a purified material, in conditions that are far from their native membranes, despite some promising in cellulo studies (Selenko and Wagner, 2007; Ito and Selenko, 2010; Renault et al., 2012; Theillet et al., 2016). To date, the most used artificial environments for NMR studies on GPCRs are detergents (Tanford, 1972; Grisshammer, 2009), micelles/bicelles (Vold and Prosser, 1996; Sanders and Prosser, 1998; Thompson et al., 2011), amphipols (Tribet et al., 1996; Dahmane et al., 2009) and lipid nanometric bilayers, also named nanodiscs or HDL particles (Bayburt et al., 2002; Whorton et al., 2007) (see e.g. Catoire et al. 2014 for a more detailed description of artificial membrane environments used for in vitro studies of MPs by NMR).

Usually, the choice of the artificial environment is based on its compatibility with the downstream biophysical application, sometimes at the expense of the protein integrity. To date, most of the NMR studies on GPCRs have been conducted in detergent solutions, because the size of the GPCR embedded in a detergent micelle is smaller compared to other membrane-like systems, considering that the quality of signals in solution-state NMR is very sensitive to the size of the protein or protein complex under study. The size dependence is closely related to the overall correlation time, which is inversely proportional to the rate of rotation of the object on itself, which impacts the nuclear spin relaxation properties. The longer the protein takes to rotate in solution on itself, the shorter the lifetime of the NMR signal is, resulting in broad NMR peaks that can even disappear in the spectral background noise.

Although they represent the easiest alternative, detergents are known to destabilize membrane proteins (Popot, 2010; Chipot et al., 2018), potentially altering their activity and function. Also, detergents undoubtedly affect the dynamics of the proteins under investigation (e.g. Chung et al., 2012), possibly compromising the interpretation of the conformational landscape, and, above all, they do not recapitulate the features of a native membrane, which has a prominent role in receptor dynamics and conformational modulation (Pucadyil and Chattopadhyay, 2006; Mondal et al., 2014; Casiraghi et al., 2016). Alternatives exist to avoid the use of detergent solutions, like nanodiscs, that mimic a more membrane-like environment that can still be compatible with the acquisition of highly resolved NMR data of GPCRs (Kofuku et al., 2014; Casiraghi et al, 2016; Casiraghi et al, 2018).
4. Isotope-labeling schemes for large proteins: deuteration

Coordinated with improvements in NMR methodology and hardware, the capacity of NMR to describe the energy landscape of large proteins or protein complexes in solution is closely related to the isotope enrichment method chosen. As mentioned in the former paragraph, the quality of the NMR signal in solution, which concerns both the sensitivity and the resolution, is predominantly determined by the overall size of the object. As the molecular weight increases, NMR spectra will display a lack of spectral resolution, due to the larger number of signals and the unfavorable relaxation properties of the system (increased relaxation rates caused by the slow overall tumbling rate), at the point that the signal deteriorates and might become impossible to detect. The best strategy to overcome those issues is to express the recombinant protein in deuterated media (Gardner and Kay, 1998). Deuteration improves resolution and the signal-to-noise by reducing the relaxation rates of the remaining protons and other nuclei under investigation such as $^{15}$N and $^{13}$C (Kerfah et al., 2015). This is due to a fundamental physical constant, the gyromagnetic ratio, which is directly related to the frequency of precession of the spin in an external magnetic field. Deuterons have a gyromagnetic ratio 6.7 times weaker than protons, which considerably reduces the dipolar interactions with the neighboring nuclei (remaining $^1$H, $^{15}$N and $^{13}$C). The main consequence is that the transverse relaxation rates of observed nuclei become much slower, substantially improving the quality of the NMR signals.

Importantly, the level of deuteration must be as high as possible in the case of large tumbling objects in solution, because the signal deteriorates dramatically due to the remaining protons around the nuclei under investigation. Perdeuteration occurs when all protons have been replaced, except for those that are aimed to be detected, like exchangeable amide protons along the backbone of the protein or protons that have been specifically incorporated during amino-acid biosynthesis or subsequently added by chemical reactions. The use of deuteration, together with the advent of ad hoc NMR experiments dedicated to the study of large protein complexes, opened the way to address complex biological questions like the study of conformational landscapes (e.g. Ruschak et al., 2010; Velyvis et al., 2012; Sekhar et al., 2016; Huang et al., 2016), or dynamics and allostery properties (Religa et al., 2010; Ruschak and Kay, 2012; Zhuravleva et al., 2012; Rosenzweig et al., 2013) of very large protein complexes, with a size that can exceed 1 MDa.

5. GPCR energy landscapes by NMR

Solution-state NMR has been widely applied to the study of the conformational dynamics of a series of GPCRs mostly obtained from expression in the membrane of insect cells (sf9 cells) and subsequently extracted from the membrane by detergents. GPCRs that have been investigated by NMR following this approach are the $\beta_2$-adrenergic receptor ($\beta_2$AR) (Bokoch et al., 2010; Liu et al., 2012a; Kofuku et al., 2012; Kim et al., 2013; Nygaard et al., 2013; Kofuku et al., 2014; Manglik et al., 2015), $\mu$-opioid receptor ($\mu$OR) (Okude et al., 2015; Sounier et al., 2015) and $\beta$-AR (Isogai et al., 2016; Solt et al., 2017). More recently, NMR investigations have been conducted on a few receptors expressed in other organisms, such as the $\alpha_2$AR, expressed in yeast (Ye et al., 2016; Clark et al., 2017; Ye et al., 2018; Eddy et al., 2018) and BLT2, expressed in Escherichia coli (E. coli) (Casiraghi et al., 2016) (see also Casiraghi et al., 2018 on the advantages and drawbacks of GPCRs expression in different hosts). The body of data collected in these studies highlights the co-existence of different GPCRs conformations at equilibrium, i.e. inactive, active and intermediate states, whose probability and lifetime are modulated by the presence of ligands and effectors with different properties.

5.1. GPCRs expressed in insect cells

$\beta_2$AR, one of the main target of investigation within the GPCR family, has been the first and most extensively characterized receptor by NMR (Bokoch et al., 2010; Liu et al., 2012; Kofuku et al., 2012; Kim et al., 2013; Nygaard et al., 2013; Kofuku et al., 2014; Manglik et al., 2015). Pioneering NMR conformational studies have been conducted by Bokoch and collaborators following the rearrangements of the extracellular surface of the
receptor upon ligand binding (Bokoch et al., 2010). They studied the conformational changes around a salt bridge between the extracellular loop 2 (ECL2) and ECL3, previously observed in crystal structures. Their NMR probes consisted in two $^{13}$C-labeled methyl groups chemically attached through reductive methylation to the ε-NH$_2$ of lysine side chains. The $^{13}$C-dimethyllysines served as conformational probes in 2D $^1$H-$^{13}$C correlation NMR experiments that uncovered the existence of three conformations for the receptor in n-Dodecyl β-D-maltoside (βDDM) detergent solution: one conformation (I) predominantly populated in the unliganded state of the receptor or in the presence of a neutral antagonist, another conformation (II) populated in the presence of an inverse agonist and a third one (III) stabilized in the presence of the agonist. These conformations are associated with different functional behaviors: when conformation (I) is predominant, the receptor is able to couple to Gs, in agreement with the concept of basal activity and the efficacy of the neutral antagonist alprenolol; when population (II) is predominant, carazolol (inverse agonist) prevents the coupling of the receptor to Gs; conformation (III), which is predominant in the presence of a full agonist, is associated with a strong coupling to Gs. Although this investigation of β2AR conformational landscape was limited to the solvent-accessible lysines, this study represents an elegant approach to the characterization of ligand-induced changes in the extracellular surface of the receptor (Bokoch et al., 2010).

A different labeling scheme was exploited by Liu and collaborators for the investigation of the conformational changes of the β2AR in βDDM in response to ligand binding (Liu et al., 2012a). Three cysteines in the cytoplasmic part of the receptor, on trans-membrane helices VI (TMVI) and VII (TMVII), were covalently labeled with 2,2,2 trifluoroethanethiol (TET) (Klein-Seetharaman et al., 1999). $^{19}$F-NMR data were deconvoluted using a double-Lorentzian function in order to obtain quantitative information about the conformational equilibria. The observation of $^{19}$F-NMR spectra upon addition of ligands with different pharmacological properties revealed the presence of two conformations for each of these two helices, whose relative populations were modulated by various ligands. Hence, the binding of an agonist would favor one of the conformations of helix VI that would correspond to the G protein specific active state, while β-arrestin-biased ligands would populate the conformational state of helix VII that would favor the interaction with β-arrestin (Liu et al., 2012a).

$^{19}$F-NMR has been successfully used in other studies on the β2AR. One study introduced the use of a novel detergent, the maltose neopentyl glycol-3 (MNG-3), which displays a very low critical micellar concentration (Chae et al., 2010). $^{19}$F-NMR investigations on chemically modified Cys265 at the intracellular extremity of TMVI of β2AR revealed the presence of different conformations: two distinct inactive states, an intermediate state en route to activation, and a predominant active state in the presence of the G protein mimic (Kim et al., 2013). The observation of four states instead of the previous two was possible thanks to the use of this new generation of detergents, which compared to βDDM display a slower off rate, resulting in a better spectral resolution (Chung et al., 2012). Experiments conducted with the unliganded receptor also highlighted the coexistence of three populations, two predominant one and a minor one (Kim et al., 2013).

Those findings were further completed by another study, published by Manglik and collaborators in 2015 (Manglik et al., 2015). This time, β2AR reconstituted in MNG micelles was studied by $^{19}$F-NMR in combination with double electron-electron resonance (DEER) experiments. Cys265 at TMVI was labeled with $^{13}$C$^{19}$F$_3$-3-bromo-1,1,1-trifluoracetone ($^{13}$C$^{19}$F$_3$-BFTA) for NMR investigations, while 3-(2-Iodoacetamido)-Proxyl (IA-Proxyl) was used to label TMIV and TMVI of the β2AR for DEER studies. They confirmed the coexistence of two inactive states for the unliganded receptor, that interconverted in the ms regime. The addition of an agonist shifted the equilibrium towards the active state but in an incompletely manner without the presence of the G protein, resulting in the co-existence of different conformational states. Complete transition to the active state was only observed in the presence of the G protein or a nanobody, a G protein mimic. The two inactive conformations in the unliganded state seemed to correspond to the intact and broken "ionic lock" (Nygaard et al., 2009), which interconvert due to the low energy barriers between the two populations. They hypothesized that the break of the ionic lock could represent an important step towards activation, a mechanism that might be conserved among other receptors of the family (Manglik et al., 2015).

$^{13}$C-labeled methyl groups of methionines were used to detect the presence of two inactive states at equilibrium, and a third active one in the presence of the agonist in the β2AR (Kofuku et al., 2012). Interestingly, with the addition of a partial agonist, the receptor was in equilibrium between inactive and active conformations. In the presence of an inverse agonist they observed a minor active conformation in equilibrium with the inactive one, corresponding to the presence of a residual basal activity.

Nygaard and collaborators used $^{13}$CH$_3$ methionines to gain access to the TM part of the receptor, not accessible by $^{19}$F labeling, to study β2AR conformational changes (Nygaard et al., 2013). They specifically followed Met82, located below the orthosteric pocket in TMII, and Met215 and Met279, located halfway between the ligand binding pocket and the cytoplasmic side of TMV and TMVI (Nygaard et al., 2013). In the unliganded state or in the presence of an inverse-agonist (corazolol), the receptor showed the co-existence of two peaks in the ms timescale, that according to molecular dynamics simulations seemed to correspond to
two distinct inactive conformations. They could also observe that agonist alone (BI-167107) did not stabilize the fully active conformation, that became the most populated one only in the presence of the G protein mimetic, the nanobody Nb80, as confirmed by other studies of the group (Manglik et al., 2015).

The studies mentioned so far represent milestones in the description of GPCR conformational landscape. However, they are characterized by some caveats: i) as mentioned in § 4, the protonated environment, which precludes the detection of subtle but important variations in the conformational landscape, and ii) the lack of a native environment, as all the studies mentioned above were conducted in detergents. The choice of detergents is mostly related to the size of the object under study, which needs to be as small as possible taking into account the impossibility to perdeuterate the receptor. Unfortunately, detergents undoubtedly affect the dynamics of the protein under investigation compared to a lipid environment (Chung et al., 2012). To face those issues, the group of Shimada published a study with two major breakthroughs: the use of lipid nanodiscs (Bayburt et al., 2002; Ritchie et al., 2009) and a partial deuteration of the βAR receptor in insect cells (Kofuku et al., 2014). Partial deuteration was made possible through the addition of 14 types of deuterated amino acids whose choice was based on the 3D crystal structure of the receptor to identify the most common residues around the βH3 groups of methionines. Only selective amino acid deuteration could be achieved, as the growth of insect cell is highly affected in D2O solutions. The amino acid deficient medium was then supplemented with 13C-methionines. Upon partial deuteration, spectra sensitivity was increased by more than 5 folds. Compared to a similar previous study of the group, in the presence of a partial agonist two signals were detected in nanodiscs, while only one was observed previously in βDDM micelles, suggesting lower exchange rates between inactive and active conformations in nanodiscs compared to βDDM. Similarly, the active population detected in nanodiscs was higher compared to the one observed in a detergent solution (Kofuku et al., 2014).

A similar approach was used to investigate another receptor of the class A family, the μOR (Okude et al., 2015). In this case, 8 amino acids were deuterated, and the 2D 1H,13C correlation experiments were conducted by following 7 methyl-labeled methionines of the receptor in a detergent solution. The authors suggested an equilibrium between an open conformation, stabilized by the full agonist (DAMGO), and a closed one in the presence of the antagonist (naloxone). When the agonist is present, multiple open conformations coexisted at equilibrium, linked to both the G-protein and β-arrestin signaling pathways. In the presence of the agonist morphine, the receptor was in equilibrium between the closed and open conformations. Biased ligands were able to further switch the equilibrium towards one or the other pathway according to their pharmacology (Okude et al., 2015).

The μOR in a detergent solution was also investigated with the use of dymethylated lysines (Sounier et al., 2015). Nine dimethylated extracellular and intracellular lysines were studied with 2D 1H,13C correlation experiments which confirmed that the binding of the high-affinity agonist, BU72, was not sufficient to fully stabilize the active conformation at the intracellular extremities of TMV and TMVI, which required G protein stabilization (or the nanobody in this case), in agreement with previous studies (Nyggaard et al., 2013; Manglik et al., 2015). The study also underlined the weak allosteric coupling between the extracellular μOR binding domain and the G-protein coupling interface, as observed for the β2AR (Manglik et al., 2015). They also investigated structural modifications at the intracellular loop 1 (ICL1) and TMVIII upon addition of the BU72 ligand, concluding that the allosteric coupling between the ligand binding pocket and the ICL1/TMVIII seems to be stronger compared to TMV/TMVI (Sounier et al., 2015). This effect was not ligand-dependent, as observed with other ligands. They hypothesized that ICL1 and TMVIII are probably the first domains to engage in G protein binding, followed by stabilization of the TM core. This study nicely completed the X-ray structure obtained for this receptor in the presence of the agonist BU72 (Huang et al., 2015).

Another approach that has been implemented for the NMR studies of GPCRs is the use of thermostabilizing mutations, a strategy that was initially developed to improve the quality of crystals without necessarily the need to add sub-domains from other proteins to increase the crystalline contacts (Serrano-Vega et al., 2008; Lebon et al., 2011; Magnani et al., 2016). A first study conducted in detergent investigated the 28 15N-labeled valines distributed in the β1AR in a detergent solution to get insights into ligand binding and receptor activation (Isogai et al., 2016). They observed that agonist binding, even in the absence of a G protein mimic, had a heterogeneous outcome at the extracellular side, while it induced initial changes in the conformational equilibrium of TMV, favoring the conformation observed in the G protein complex of β2AR (Rasmussen et al., 2011). They also reported that the stabilization of GPCR fully active conformation required the binding of an agonist and an intracellular partner, as already highlighted in other studies (Manglik et al., 2015; Sounier et al., 2015). However, to observe a complete shift towards the active state in
the presence of the nanobody, two thermostabilizing mutations had to be reverted (Y227S and Y347S), significantly reducing the thermal stability of the receptor (Isogai et al., 2016).

Another recent study on a thermostabilized β2AR in detergent investigated, through 13C-labeled methionines, the role of different ligands and G-protein mimic nanobodies in GPCR activation (Solt et al., 2017). The study suggested an equilibrium between an inactive (i) and a pre-active state (A), the latter being more populated in the presence of a full agonist. This pre-active state was believed to sample dynamics in the order of the μs-ms timescale, that were associated to large amplitude motions at both the extracellular and intracellular side of the receptor. In contrast to observations with the β2AR (Nygaard et al., 2013; Kim et al., 2013; Manglik et al., 2015), the unliganded receptor or the receptor bound to low-affinity ligands displayed more restricted dynamics compared to the full agonist-bound receptor. According to these observations, agonist-bound receptor appeared to be more dynamic than the agonist-receptor-nanobody complex or the receptor coupled to the nanobody only. They interpreted the data by hypothesizing that the highly dynamic form of the agonist-bound receptor is able to sample different active-like conformational states that are primed to bind the different intracellular signaling effectors (Solt et al., 2017). To be noticed, as the analysis of the dynamic of the receptor was based on variations in peak intensities only, it would be interesting to incorporate spin relaxation properties to confirm this interpretation.

5.2. GPCRs expressed in yeast (Pichia pastoris) or E. coli

All the aforementioned studies were based on the expression of GPCRs in the membrane of insect cells (sf9 cells) followed by the extraction of the receptor from the membrane with the help of detergents. Despite the tremendous amount of information obtained, this methodology remains limited by the impossibility of perdeuterating the receptor, which hampers the detection of some conformations and subtle changes in the conformational landscape that might have important functional implications in the context of receptor activation and signaling. Also, i) chemical modifications of amino acids within the receptors might alter the dynamics of the system under study to an extent that is difficult to be judged, even when the coupling to ligands or G protein is still retained; ii) thermostabilization usually reduces or modifies the dynamics of the system; iii) except for one case (Kofuku et al., 2014), all the above-mentioned studies were conducted in detergent solutions, that are far from reproducing a membrane-like environment that undoubtedly impact membrane protein dynamics and conformational modulation (Pucadyil and Chattopadhyay, 2006; Mondal et al., 2014; Chipot et al., 2018). To overcome those limitations, in the recent years, other expression systems for GPCRs production have come to the spotlight, in particular yeast and E. coli, whose main advantage is the capacity to grow in very hostile conditions, typically 100%-D2O media. Thanks to the use of appropriate labeled precursors associated with the expression a growth in 100%-D2O media, high-resolution NMR has been shown to be accessible to large systems, up to 1MDa (Kay, 2005; Religa et al., 2010; Ruschak et al., 2010; Kay, 2016).

Yeast has often been used for expression of membrane proteins due its inexpensiveness, ease of use, good scalability and genetic flexibility. Yeast has grown particular interest for NMR studies on GPCRs because it is a eukaryotic system and it is able to grow in a deuterated environment (Clark et al., 2018). However, achieving a good protein expression yield is not always straightforward, particularly in D2O media. This is an important factor if considering that NMR is a technique that requires a lot of material (i.e. few tens of μM to look at conformational ensembles to few hundreds of μM to investigate kinetic barriers, for a sample volume of 150-300 μL). Usually the combination of multiple strategies is required to get a good yield of expressed protein in yeast, and to date only a few GPCRs have successfully been expressed in mg amount in P. pastoris, among which the A2A adrenoceptor (O’Malley et al., 2007; Hino et al., 2012) and the histamine H1 receptor (Shimamura et al., 2012). The first NMR study published on the A2AAR consisted in the trifluoromethyl labeling of a modified V229C residue located at the cytoplasmic end of TMVI (Ye et al., 2016). 19F-NMR experiments revealed the presence of two inactive and two active states for the unliganded receptor, in contrast to what observed in the case of the β2AR (Manglik et al., 2015). In this "loosely coupled" ensemble, the equilibrium was shifted towards the two active states upon addition of partial or full agonists. The exchange observed between the two inactive states was in the ms regime, while between the inactive and active states the exchange occurred at a slower rate, in the second timescale, probably due to higher activation barriers. Remarkably different from the case of the β2AR, which is biased towards the inactive state unless the concomitant addition of the agonist and nanobody (Manglik et al., 2015), 70% of the unliganded A2AR was already in the active states. The authors suggested that probably the active states associated with the unliganded β2AR are very weakly or shortly populated, and therefore extremely exchange-broadened to the point that they cannot be easily detected.
A recent study from the same group used \(^{19}\text{F}-\text{NMR}\) to study the effect of cations on the functional states of A\(_2\)AR in a detergent solution (MNG) (Ye et al., 2018). The data collected confirmed the findings published in the previous article, with a conformational landscape consisting in an ensemble of inactive, intermediate and active states modulated by the addition of ligands. Sodium was reported to have a negative allosteric effect on the activation of GPCRs (Christopoulos and Kenakin, 2002; Liu et al., 2012b; Fenalti et al., 2014), as it seemed to favor the inactive and intermediate states at the expense of the active one, an effect that seemed to be concentration-dependent. \(^{23}\text{Na}\) NMR measurements revealed that the binding of sodium was transient, characterized by a residency time of 480 \(\mu\text{s}\) on the receptor. They also hypothesized that the release of sodium from the binding site in the unliganded receptor might contribute to the activation process. In the same study the role of other cations, like calcium and magnesium, was also investigated by MD simulations, which suggested a positive allosteric role for the cations through their binding to the acidic residues located in ECL2 and ECL3, described as long-lived electrostatic interactions. Cation binding to these residues would also bring TMV and TMVI closer, facilitating TMVI intracellular rotation (Ye et al., 2018).

The role of sodium was also addressed with another approach, by looking at the protonated and \(^{13}\text{C}\)-labeled methyl groups located at the extremity of the side-chain of isoleucines of the A\(_2\)AR in a detergent solution. The receptor was produced in 100%-\(\text{D}_2\text{O}\) leading to an almost perdeuterated receptor (~90%) (Clark et al., 2017). The improved deuteration allowed the acquisition of 2D \(^1\text{H},^{13}\text{C}\) NMR experiments, even though the incorporation efficiency of the isotopic labeled probes was estimated to be only around 50%. The 2D \(^1\text{H},^{13}\text{C}\) correlation spectra were recorded in the presence of the agonist NECA or the inverse agonist ZM241385. 20 peaks out of 29 of the \(\delta\)-CH\(_3\) of isoleucines labeled were assigned and the ligand-binding changes observed by NMR were analyzed in the context of previous structural data. A too high sodium concentration seemed to suppress the large structural changes that occurred to the A\(_2\)AR receptor when bound to ligands, and the study of the side-chains motions of several regions of the receptor showed different conformational patterns depending on whether the receptor was bound to an agonist or an inverse agonist (Clark et al., 2017).

The A\(_2\)AR was again investigated in a detergent solution in a partially deuterated environment (Eddy et al., 2018). \(^{15}\text{N}\) uniformly labeled A\(_2\)AR was used to investigate the role of Asp52\(^{2.55}\) as a part of an allosteric hinge between the orthosteric binding pocket and the intracellular side of the receptor. NMR signals from residues located at both the extracellular and at the intracellular sides of the receptor were recorded both for the A\(_2\)AR and its version with the allosteric switch off, A\(_2\)AR[D52N] (Massink et al., 2015). To confirm that the A\(_2\)AR from \(P.\ pastoris\) had the same properties of the receptor expressed in insect cells, ligand-binding experiments were performed together with the crystallization of the receptor in the presence of an antagonist. Among the tryptophan and glycine residues that they could assign by NMR, they followed Trp246\(^{\delta6.48}\) which belongs to the "toggle switch" at the bottom of the ligand binding pocket on TMVI (Rosenbaum et al., 2009; Nygaard et al., 2009). They suggested a coupling between the toggle switch and Asp52\(^{2.52}\), based on the loss of the signaling-related conformational dynamics at the intracellular surface for the A\(_2\)AR[D52N], which correlated with changes in the local environment explored by the Trp246\(^{\delta6.48}\), in association with the activation of the receptor (Eddy et al., 2018).

\(E. coli\) has also been employed to conduct NMR studies on GPCRs. This host is closely related to the most robust and cost-effective strategy to study large proteins or protein complexes by high-resolution solution-state NMR (Goto et al., 1999; Tugarinov et al., 2004; Tugarinov and Kay, 2005; Tugarinov et al., 2006; Kay, 2011). This bacterial host displays a series of advantages: it is an established system for heterologous membrane proteins production, inexpensive, easy to genetically manipulate, and above all, it is capable of expressing MPs at a high yield in 100%-\(\text{D}_2\text{O}\) media (for GPCRs, typically 0.5 to 5 mg per liter of culture of purified protein). The main disadvantages are the lack of post-translational modifications and the fact that high levels of expression may result in inclusion bodies formation, for which an efficient folding strategy is needed (for more details, see Casiraghi et al., 2018).

To maximize production within \(E. coli\), GPCRs can be expressed as inclusion bodies (Banères et al., 2003; Arcemisbehere et al., 2010; Banères et al., 2011). Indeed, folding of membrane proteins from inclusion bodies no longer represents a bottleneck (Popot, 2014) and this includes also the GPCR family (Dahmane et al., 2009). Such strategy allowed to study the conformational ensemble of the low affinity leukotriene receptor BLT2 (Casiraghi et al., 2016). BLT2 was expressed in \(E. coli\) in 100% \(\text{D}_2\text{O}\) in the presence of labeled precursors which led to the \(^{13}\text{C}\)-labeling of protonated methyl groups of methionines and isoleucines (\(\delta\)-position). This concerned five methionines (including the two TM residues Met105\(^{3.36}\) and Met197\(^{5.54}\) and one single isoleucine (Ile229\(^{6.40}\), on the intracellular side of TMVI) (Casiraghi et al., 2016). The receptor was reconstituted into lipid nanodiscs and underwent a full panel of characterization assays to demonstrate that BLT2 from \(E. coli\) was fully active and stable prior to NMR. Ligand binding data and G protein activation experiments confirmed that the receptor recapitulated the functional characteristics of a GPCR expressed in the detergent solution [MNG] (Ye et al., 2018).
eukaryotic cells. The high-quality NMR data recorded allowed to monitor the subtle changes in BLT2 conformational ensemble in response to different ligands. For the unliganded receptor, different co-existing conformations at equilibrium were detected: two inactive states and an intermediate one, in agreement with previous studies (Kim et al., 2013; Manglik et al., 2015) and an active or active-like conformation. This active-like state observed for the unliganded receptor was never observed before and probably corresponds to a low-populated conformer that become visible thanks to the improved resolution of this method. The addition of an agonist, in parallel with measurements of GTP binding to the G protein, confirmed that this population corresponded to the active state (Casiraghi et al., 2016). The use of nanodiscs allowed the investigation at the atomic scale of the role of the lipid bilayer in GPCR modulation (Pucadyil and Chattopadhyay, 2006). In particular, they confirmed that negatively charged cholesteryl hemisuccinate, an analogue of cholesterol used for GPCR structural studies (Kuszak et al., 2009), has a positive allosteric effect on BLT2 activity. This study proved the importance of receptor deuteration, that otherwise would have hindered the observation of several co-existing sub-states with possible important functional implications. Perdeuteration also allows to work with nanodiscs, one of the best membrane-like systems for the in vitro study of the membrane impact on GPCRs function.

6. Concluding remarks

During the past decade, NMR proved to be a powerful technique to investigate the energy landscape of GPCRs. The pivotal studies aforementioned have highlighted some fundamental features that occur during the activation of these proteins. NMR investigations supported the evidence that agonist binding and the consequent receptor activation is a multi-step process, that involves different conformational states. Depending on the level of resolution achieved by NMR several conformations have been detected, corresponding to inactive states, intermediates towards activation, and active or active-like states. The probability and duration of those states depends on the presence of ligands with different pharmacology properties, intracellular effectors and allosteric modulators like ions or lipids. NMR is a very efficient tool to characterize the shift in the populations equilibrium and the relative stability between conformations, as the rearrangements that occur at a structural level on the receptor are detected as chemical shift variations at physiological temperatures. Investigations on GPCR dynamic landscape have also confirmed important insights into receptor function, as the stabilization of a given conformation seems to preferentially engage some signaling pathways rather than others. While GPCR energy landscape has been widely studied for several receptors, the next step in the characterization of GPCR dynamics by NMR consists in the determination of the kinetic barriers that separate the different conformational sub-states. This will help to shed light on the transitions between the distinct states that populate the conformational landscape and will possibly detect the presence of additional sub-states not observed yet because of too lowly populated or transiently formed. To do so, NMR relaxation experiments have to be conducted in order to assess if a chemical exchange exists between the observed conformations or towards lowly and transient states (Sekhar and Kay, 2013). The characterization of kinetic barriers within these complex allosteric proteins will provide important information regarding activation energies separating the different sub-states and their modulation by ligands. Above fundamental issues, this should greatly help the design of new drugs to block or activate these receptors.

Acknowledgments

This work was supported by grants from the Centre National de la Recherche Scientifique, Universités de Montpellier and Paris Diderot, Agence Nationale de la Recherche (ANR-17-CE11-0011), Laboratoire d’Excellence (LabEx) DYNAMO and Equipements d’Excellence (EQUIPEX) CACSICE from the French Ministry of Research, and a postdoctoral fellowship Marie Curie Global Fellowship for M.C. (“AlloGPCR”-799376). The authors acknowledge Soumaya Zlitni for the diligent proofreading of this manuscript.
References


Captions

Fig. 1. Topology of GPCRs. (A) GPCRs are integral membrane proteins constituted by seven transmembrane helices (in orange), three extracellular loops (ECLs, in green) and three intra-cellular loops (ICLs, in purple). In the case of class B, C or F GPCRs, the N-terminus presents a structured extra-cellular domain (schematically represented by a dotted circled). (B,C,D) Cartoon representations of the crystal structure of leukotriene B4 BLT1 receptor (PDB ID: 5X33; Hori et al., 2018) parallel to the membrane (B), from the extracellular side (C) and from the cytoplasmic side (D). α-helices and β-sheet secondary structure elements are represented with cylinders and arrows, respectively.

Fig. 2. Energy landscape. Illustration with one-dimensional cross-section through the hypersurface energy landscape of a biomolecule. Formally, an energy landscape describes all the possible conformations adopted by a biomolecule and the energy levels associated. The free energy of Gibbs, which corresponds to a thermodynamic potential, is traditionally used in such diagrams to describe chemical exchanges. This diagram describes a chemical exchange between two main populations, $p_A$ and $p_B$, where each state contains several sub-states separated by low energy barriers. Populations $A$ and $B$ follow a Boltzmann distribution based on their difference in their free energies, indicated by $\Delta G_{AB}$. The barriers between these states ($\Delta G^2(A \rightarrow B)$ and $\Delta G^2(B \rightarrow A)$) determine the rates of interconversion ($k_{A \rightarrow B}$ and $k_{B \rightarrow A}$), which are directly related to the forward and backward energies of activation. Importantly, the energy landscape defines the amplitude and timescale of protein motions in direct relation with the energy barriers than can be classified in three tiers, following the definition introduced by Frauenfelder and collaborators (Ansari et al., 1985). Tiers 2 and 1 concern fast motions between a large number of structurally similar sub-states, in the ps and ns timescales respectively. In these tiers, the kinetic barriers are low, which also means that the probabilities of transition between the different sub-states are quite high. Conversely, in Tier 0, the probability of transition is much lower and concerns slower motions, in the μs and ms timescales leading to a conformation $B$ that can be quite different than conformation $A$ because it concerns a priori collective motions of atoms. In the case of GPCRs, the binding of a ligand could modify the energy landscape as illustrated here (dark to light blue curve), by stabilizing the excited state $B$ and increasing its population. According to the energy landscape formalism, GPCRs that display a substantial basal activity will display more favorable energy of activation toward state $B$ compared to GPCRs that display a low or a lack of basal activity. The subsequent binding of a G protein could also stabilize the receptor by decreasing its Gibbs free energy, as showed in various NMR studies (e.g. Manglik et al., 2015). Hence, to fully understand how GPCRs function, in complement to atomic structures of the different sub-states, the relative probabilities between these different sub-states (thermodynamics) and the different dynamic barriers (kinetics) separating these sub-states have to be characterized (figure adapted from Henzler-Wildma and Kern, 2007; reprinted by permission from Springer Nature).

Fig. 3. The shape and chemical shift of two-dimensional NMR peaks depend on the exchange rate between two conformations $A$ and $B$. The chemical shift is very sensitive to structural changes in a biomolecule and can be measured with high accuracy by NMR. Most of the NMR studies of conformational ensembles of GPCRs are based on 2D $^1$H,$^1$C or $^1$H,$^15$N correlation experiments. In the simulated 2D $^1$H,$^13$C experiment represented here, the simulated peaks indicated with various colors correspond to one or more protons (i.e. chemically equivalent like the three protons in a methyl group for instance) attached to a same $^{13}$C. (Left) When the exchange rate is fast compared to the chemical shift difference between conformations $A$ and $B$, the position of the observed peaks is directly proportional to populations $A$ and $B$. At the limit of very fast exchange on the NMR chemical shift timescale, peaks have the same shape. In the intermediate exchange regime, i.e. when the exchange rate becomes similar to the chemical shift difference, the signals broaden and shift at the same time. (Right) In the condition of a slow exchange, with populations $A$ and $B$ decreasing and increasing, peak $A$ (blue) will for instance decrease in intensity as peak $B$ (red) increases, with no change in chemical shifts occurring for both peaks. When peaks $A$ and $B$ do not display the same shape, it becomes necessary to take into account the relaxation properties of both states in order to accurately determined the populations involved in the chemical exchange (figure adapted from Williamson, 2013; reprinted by permission from Elsevier).
Fig. 1
Fig. 2

Tier 0

Tier 1

Tier 2

Free energy, $G$

Conformational coordinate

$\Delta G^\ddagger (k_{A\rightarrow B})$

$\Delta G^\ddagger (k_{B\rightarrow A})$

$\Delta G_{AB}$

(ns)

(ps)

(A)

(B)

$\langle p_A, p_B \rangle$
Fig. 3