Peeking at G-protein-coupled receptors through the molecular dynamics keyhole

Deganutti, G., Moro, S. & Reynolds, C. A.

Published PDF deposited in Coventry University’s Repository

Original citation:
https://dx.doi.org/10.4155/fmc-2018-0393

DOI 10.4155/fmc-2018-0393
ISSN 1756-8919
ESSN 1756-8927

Publisher: Newlands Press

This work is licensed under the Creative Commons Attribution 4.0 License.

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.
Molecular dynamics is a state of the art computational tool for the investigation of biophysics phenomenon at a molecular scale, as it enables the modeling of dynamic processes, such as conformational motions, molecular solvation and ligand binding. The recent advances in structural biology have led to a bloom in published G-protein-coupled receptor structures, representing a solid and valuable resource for molecular dynamics studies. During the last decade, indeed, a plethora of physiological and pharmacological facets of this membrane protein superfamily have been addressed by means of molecular dynamics simulations, including the activation mechanism, allostery and, very recently, biased signaling. Here, we try to recapitulate some of the main contributions that molecular dynamics has recently produced in the field.

First draft submitted: 10 August 2018; Accepted for publication: 4 January 2019; Published online: 19 March 2019

Keywords: allostery • biased agonism • class B GPCRs • GPCRs • ligand binding • molecular dynamics • solvation

The G-protein-coupled receptors (GPCRs) are a superfamily of integral membrane proteins encoded by more than 800 different human genes [1]. GPCRs transduce extracellular chemical signals into the cellular cytoplasm, where they interact with a set of intracellular effector proteins able to trigger many different physiological effects. GPCRs engage heterotrimetric G proteins, generally able to modulate the cytoplasmatic release of calcium [2] and the production of cyclic adenosine monophosphate (cAMP), phosphatidylinositol-4,5-bisphosphate and diacylglycerol [3]; GPCR signaling can also involve arrestins [4], once the receptor C terminus has been phosphorylated by GPCR kinases [5]. GPCRs are structurally characterized by seven transmembrane (TM) helices, interrupted by three flexible extracellular loops (ECLs) and three intracellular loops (ICLs; Figure 1) [6,7]. It is the inherent flexibility of GPCRs that make molecular dynamics (MD) such a useful technique for their study.

It has been estimated that approximately 34% of approved drugs act on GPCRs [8], making GPCRs the most targeted protein family. Considering that these drugs largely act on a small proportion of the family and that the development of potential drugs for most of the GPCRs currently targeted is still in early or preclinical phases, there is considerable prospect and scope for designing novel therapeutic candidates. In recent years, x-ray crystallography first, and more recently cryo-electron microscopy (cryo-EM) techniques have led to a bloom of published GPCR structures [9], reinforcing the importance of structure-based drug design as a fundamental approach in both academic and industrial research environments. From this point of view, MD simulations represent a helpful computational tool for the rationalization of experimental findings as well as the generation of further working hypothesis at the molecular level. MD algorithms enable a description of the temporal evolution of a system according to its potential energy and Newton’s laws (knowing the acceleration, velocity and position of the atoms at a precise instant in time, $t$, it is possible to approximate its velocity and position at a successive instant $t + \Delta t$). In biological disciplines as well as other fields, computational sciences are taking advantage of the continuous advancement of both hardware and software [10]. It follows that the timescales that MD can cover have leaped from picoseconds, on systems with a few hundred atoms in the 1970s [11], to milliseconds (on purpose-dedicated computers [12]) for systems with hundreds of thousands of atoms. However, the duration of biological phenomena such as conformational activation and ligand unbinding usually go beyond the MD capability [13]. To tackle this issue, many so-called adaptive [14] and enhanced [15] sampling MD protocols have been developed, allowing simulations to overcome energetic barriers.
between stable basins and, in turn, to sample biologically relevant configurations kinetically inaccessible by means of classic MD simulations (Figure 2).

In this review, we try to outline and summarize the contribution of atomistic MD simulations to recent advances in the structural comprehension of GPCRs. The main focus will address crucial aspects of ligand binding, the allostery phenomena, solvation effects and the activation mechanism of this captivating receptor family.

**Ligand binding**

In equilibrium conditions, a binding process can be described as an ensemble of states that are populated with a probability proportional to the energetic stability gained by the ligand (Figure 2). According to transition state theory (which probably oversimplifies the actual free-energy surface of binding), the probabilities (and the timescales) of transitions between these basins are driven by the energy needed to overcome the saddle points (transition states). It follows that a ligand can be stabilized in several intermediate configurations (metastable states, Figure 3) along the path to the final bound state. The energy required for overcoming a transition state and to evolve toward another metastable state is usually due to desolvation phenomena, conformational rearrangements and breaking intermolecular interactions. From this standpoint, MD has proved to be a useful tool for describing the ligand (un)binding mechanism for several families of receptors [16,17].

**GPCRs ligands (un)binding: a dynamic perspective**

During drug discovery campaigns, it is common to refer to the affinity or potency of a drug candidate, to quantify in vivo activity. However, living organisms are open systems, able to exchange both energy and matter with the surrounding environment. In such physiological conditions, pharmacokinetic phenomena do not allow chemicals to reach a steady state in the proximity of their biological targets after administration. During recent years, there has been a gradual shift from the axiomatic use of the thermodynamic constants $K_D$ (dissociation constant) and $K_A$ (association constant) to nonequilibrium parameters (i.e., the kinetics constants $k_{on}$ and $k_{off}$) as ligand-target affinity descriptors. In fact, emerging evidence has pointed out the importance of the binding kinetics on the pharmacology of both endogenous and xenobiotic compounds [18-21]. The dissociation kinetics and therefore the target residence time (RT, defined as the reciprocal of the $k_{off}$) should be a better predictor of in vivo efficacy in specific cases, as compounds with similar thermodynamic affinity for a biological target can have a very different
Peeking at G-protein-coupled receptors through the molecular dynamics keyhole

Figure 2. Schematic representation of different molecular dynamics approaches. Each state of the system, within all the possible ensemble of configurations (configurational domain), is associated with an energy value. The system can experience a transition from one basin (here indicated as a, b, c and d) to another if it is able to traverse the transition state (TS) between them. By means of classic MD, it is possible to sample minima separated by low-energy TSs (i.e., TS1). By monitoring one or more physical quantities during the simulations (adaptive sampling), it is possible to reduce the time needed to overcome higher energy transition states (TS2). Enhanced sampling methods can enable the system to visit all the states by adding an energetic bias: configurations kinetically unreachable for classic MD and adaptive sampling methods (TS3) can, therefore, be sampled during the simulation. Reported transition timescales are just indicative of the orders of magnitude that enhanced sampling methods allow, compared with classic MD. MD: Molecular dynamics; TS: Transition state.

Figure 3. Representation of metastable and transition states along the binding pathway of a G-protein-coupled receptor ligand. (A) The centroid of a GPCR ligand (shown as a small red sphere) populates different metastable states during the binding process, before reaching the orthosteric site. Energetically favored transitions from different metastable states describe complete binding paths (in this example, binding path 1 and 2). (B) Centroids corresponding to energetically unfavorable states (white spheres) are widely dispersed at the extracellular vestibule, indicating the receptor/ligand desolvation as a putative kinetic bottleneck, as proposed in [31]. GPCR: G-protein-coupled receptor; ECL: Extracellular loop; TM: Transmembrane.
RT, and therefore a kinetic selectivity (KS). An example of KS is ascribable to tiotropium [22,23], which when dissociating about ten-times slower from the M3-R (muscarnic receptor M3) than the M2-R subtype, is able to trigger the pharmacological effects of the former, limiting the side effects due to the inhibition of the latter one. A further case of how RT can tune in vivo effects could be represented by the atypical antipsychotic drugs [24]. Indeed, it has been proposed that antagonist of the D2-R (dopamine D2 receptor) that are characterized by a faster dissociation rate have reduced on-target extrapyramidal side effects [25].

Also the association kinetics is progressively more recognized as a significant factor in determining a ligand’s activity. On-target side effects, in theory, can be reduced by slow-binding ligands, as high target occupancy and activation may be limited. However, a slow binder would not be able to produce a maximal effect in vivo, as the pharmacokinetic elimination may limit the drug tissue concentration and therefore not allow the full occupation of the biological target. Fast binders, on the other hand, are more prone to rebinding phenomenon, which can result in prolonged receptor activity [26]. The necessity of drug candidates with a well-balanced kinetic is increasingly producing efforts on the development of quantitative structure–kinetic relationship protocols, [27–30].

In the GPCR field, the binding to the β2-AR (β2-adrenergic receptor) [31–33] and the M2-R and M1-R [33] were the first to be studied using MD, in 2011. In their pioneering work, Dror et al. [31] simulated several spontaneous binding events of the β2-AR antagonists propranolol, alprenolol, dihydroalprenolol as well as the agonist isoproterenol, observing a common binding path characterized by metastable states in the proximity of extracellular loop 2 (ECL2), extracellular loop 3 (ECL3) and the upper portions of transmembrane helix 6 (TM6) and TM7. Interestingly, the energy barrier of the process was ascribed to the rapid dehydration (within 1 ns) of both the ligand and the receptor in order to establish hydrophobic interactions at the extracellular vestibule. For the M2-R and M3-R [33], simulations highlighted the presence of several metastable states but suggested that conformational rearrangement involving the receptor may act as the rate-limiting step of the overall event. Interestingly, acetylcholine was instead able to bind the M3-R without experiencing any metastable state in concordance with its status as an ultrafast binder [34]. Metastable sites at the protein-solvent interface were also observed during the dissociation paths for β2-AR inverse agonists [32].

The A2A AR (A2A adenosine receptor) is among the most characterized class A GPCRs, as to date more than 40 structures have been published [35]. Several computational studies have addressed the dynamics of this receptor from different points of view. Guo et al. [36] employed temperature-accelerated MD simulations (i.e., a form of enhanced sampling) to sample the dissociation mechanism of the antagonist ZM241385 and identified several amino acids, grouped in two different topological clusters, able to modulate the unbinding kinetics rather than the equilibrium affinity of the ligand. The authors proposed a multistep dissociation, indicating the disruption of the ionic interaction between ECL2 and ECL3 as the early stage of the event. The influence of this intramolecular protein interaction on the koff of different antagonists was addressed in the work of Segala et al. [37], which reported how different ligand substituents can modulate the stability of the salt bridge and, in turn, the ligand RT.

The process of adenosine binding to its A2A AR, both in the intermediate-active [38] and active [39] state, was recently investigated using the supervised MD (SuMD) method, an adaptive method [40,41]. During the simulations the endogenous agonist was able to reach the crystallographic bound conformation only in the presence of the engineered G protein, in accord with the scenario in which receptor pre-coupling with intracellular effectors allosterically influences the association of agonists [42,43]. Very recently, it has been hypothesized that the A2A AR may contemporarily engage two molecules of its natural ligand, leading to a 1:2 complex [44]. This scenario has been partially supported by the experimental observation that adenosine, in presence of the potent agonist CGS-21680, binds to a site alternative to the orthosteric one [45]. SuMD simulations were also combined with a biased approach to characterize putative kinetics bottleneck during the (un)binding of A2A AR antagonists [46], proposing a critical role of the solvation/desolvation of the protein and the ligand in transition states of entropic nature.

Besides the bulk solvent, the lipid environment also allows the diffusion of (lipophilic) ligands towards GPCR orthosteric sites. From this standpoint, MD simulations have shed light on the binding processes to the P2Y1 R (purinergic P2Y1 receptor) [47], the CB2R (cannabinoid receptor 2) [48], and the S1PR (sphingosine-1-phosphate receptor 1) [49]. Interestingly, the latter study reported ionic interactions at the protein surface in contact with lipid environment receptor membrane vestibule as well as the desolvation required for the ligand entrance to the binding site as kinetics bottlenecks of the event. Temperature-accelerated MD was used to study the unbinding paths of the PAR1 (protease-activated receptor 1) antagonist vorapaxar [50,51], proposing the putative orthosteric access of the ligand through the cleft between TM6 and TM7.
Figure 4. Structural hallmarks of a class A G-protein-coupled receptor activation mechanism. The most striking structural change is the outward movement of TM6 (about 6–14 Å) from the TM bundle that allows the intracellular effector to bind. Further conformational switches (ribbon representations) involve the ‘ionic lock’ between R3.50 and D/E6.30, the NPxxY (N7.49, P7.50 and Y7.53) motif, and the highly conserved W6.48. The residue numbering follows the Ballesteros–Weinstein scheme [92].

ECL: extracellular loop; TM: Transmembrane.

MD is increasingly employed as a tool for the challenging direct estimation of binding kinetics parameters [52]. Besides the construction of Markov state models [53] (which are usually based on long timescale-unbiased simulations [54,55]), other computational approaches allow simulating GPCR (un)binding events by keeping track of the energy needed to achieve the transition between basins. Bortolato et al. [56] ranked a set of CRF1R (corticotropin-releasing factor 1 receptor) ligands on the basis of the maximum energy required for simulating the dissociation, gaining good agreement with the experimental RTs. In a subsequent work [57], the same authors proposed that a binding pathway from the membrane bilayer, between TM5 and TM6, was the energetically favored pathway for the allosteric modulator CP-376395.

**GPCR activation & signaling**

In order to accomplish their membrane transducer functions [58,59], GPCRs undergo structural transitions capable of transferring the extracellular chemical information (i.e., the agonist binding) to the intracellular side of the cytoplasmic bilayer, where cognate proteins bind. A few decades after the ternary model was formulated [60], structural biology and related disciplines [3] are now increasingly shedding light on GPCR activation mechanisms [7,61], highlighting an intrinsic dynamicity of this receptor superfamily [62] and common structural hallmarks among different GPCR classes [6,9,62–66]. As shown in Figure 4, the principal structural change involves the intracellular side of the receptor, where TM6 kinks and swings as a result of well-placed conserved proline residues (additionally for class B, conserved glycine residues), moving away (6–14 Å) from the TM bundle and shaping the G-protein-binding site. Also TM5, TM2 and TM7 slightly rearrange, in order to accommodate the intracellular effector, for example, the C-terminal helix of a G protein.

The first direct computational insights into the β2-AR dynamic activation mechanism were proposed in the work by Dror et al. [67], following the first x-ray crystal structure of an active (G-protein or G-protein mimic-bound) GPCR [68,69]. Starting from the agonist-bound active state, with the intracellular G-protein or G-protein mimic nanobody removed, the authors performed extensive simulations observing several spontaneous transitions toward the experimentally determined inactive conformation. The receptor experienced numerous intermediate conformations along the deactivation pathway, but without a strong coupling mechanism between the orthosteric and G-protein-binding sites, giving a putative rationale for the experimental difficulties in fully stabilizing the


\[ \beta_2\text{-AR in the presence of an agonist [79]. In contrast, other MD studies have shown an allosteric communication between the orthosteric site and a conserved polar network of the } \mu\text{OR (opioid receptor } \mu) [71,72].\]

The GPCRs’ conformational heterogeneity, observed upon agonist binding, may be the basis of the different coupling possibilities with intracellular counterparts (i.e., stimulatory G protein – \( G_s \), inhibitory G protein – \( G_i \), and arrestin) [73]. Recently, Kohlhoff et al. [74] performed millisecond timescale MD simulations of the \( \beta_2\text{-AR} \) on the Google’s Exacly cloud, obtaining a detailed description of the overall receptor activation mechanism. Even though different pathways were observed, the model that the author proposed confirmed the displacement of TM6 away from TM3, either before or concurrently with conformational changes at the NPxxY motif (Figure 4) and the connector region (a TM cluster of residues located between the orthosteric site and the G-protein-binding sites).

Li et al. [75] proposed that the direct contacts between the \( A_{2A} \) AR and the agonist could induce a side-chain rearrangement in residues located at the bottom of the orthosteric site, leading to a loss of hydrophobic packing in the core of the receptor. This, in turn, allows TM6 to move away from TM3 due to the break of the ‘ionic lock’ (Figure 4). Further molecular switches of the \( A_{2A} \) AR have been reported in the work by Lee et al. [76].

In the \( M_2\text{-R} \), enhanced sampling MD simulations of the activation mechanism highlighted the concomitant formation of polar interactions between TM5 and TM7, and the outward movement of the intracellular portion of TM6 [77], while TM3 was confirmed as fundamental for the functionality of the receptor [63]. Very recently, a study on the \( M_2\text{-R} \) deactivation (simulated by replacing an agonist with an inverse agonist in the active conformation of the receptor [78] deprived of the G-protein mimetic nanobody [79]), suggested that the opening of the extracellular side of the receptor may occur before the intracellular relaxation of TM6.

About 15 years ago, the concept of biased signaling was first introduced in the GPCR field [80]. According to this novel view, an agonist-bound GPCR can preferentially couple to either one or more of a number of G-proteins or a \( \beta \)-arrestin, generating a particular intracellular signal that can be triggered by any of the three components of the ternary complex (e.g., the agonist, the receptor or the transducers) [81,82]. Even though the structural determinants of the biased signaling are still poorly understood, it is well accepted that small variations in the conformational energy landscape of the receptor lead to different intracellular signaling pathways, in other words, different GPCR active conformations preferentially bind to particular effectors [83,84].

Understanding agonist bias offers the opportunity to develop drugs with an enhanced therapeutic profile, and reduced on-target side effects [84]. Recently, ligand RT and biased agonism have been described as inter-related in the ‘kinetic context’ [85], indicating that the overall binding kinetics can modulate the nature of the receptor signaling [86]. In theory, interactions at any site of the receptor surface that recognize a ligand can shift the conformational equilibrium toward a specific conformation, therefore triggering a particular physiological effect [82]. From this standpoint, MD may represent a useful tool for the investigations of the GPCR conformational landscape, not only in their canonical orthosteric-bound state but also when potential metastable sites are engaged. Schneider et al. [87] investigated the association mechanism and the orthosteric behavior of the \( \mu\text{OR agonist olliceridine (OLINVO)\^TM, TRV-130} \), the first biased opioid to enter Phase III clinical trials [88]. Simulations highlighted marked differences in the structural communication, determined by complex networks of both polar and nonpolar intramolecular interactions between the extracellular side and the intracellular portion of the receptor, depending on whether TRV-130 or the reference unbiased agonist morphine was bound to the orthosteric pocket (mainly at the level of TM3 and TM6). The significance of this resides in the reduced side effects reported for olliceridine in Phase III clinical trials for moderate-to-severe acute pain as a result of reduced \( \beta \)-arrestin signaling (www.trevena.com).

Very recently, the dynamic recognition process between a GPCR and an intracellular binding partner was reconstructed by means of MD. Miao et al. [89] simulated the intracellular association between the \( M_2\text{-R} \) and a G-protein mimic nanobody. In concordance with mutagenesis experiments, the authors proposed that ICLs play a fundamental role by stabilizing the complex by establishing both hydrophobic and hydrophilic contacts. Moreover, the cytoplasmic end of TM6 favors the association by becoming transiently more disordered as the nanobody approached the receptor. At the extracellular side of the system, the binding pocket evolved from an open conformation (observed in the inactive state) to a closed one (Figure 4), in concordance with the allosteric coupling proposed also for the \( \beta_2\text{-AR} \) [43]. On the other hand, the arrestin activation mechanism was extensively studied by means of simulations in two parallel works by Eichel et al. [90] and Latorraca et al. [91]. Latorraca et al., in particular, rationalized an astonishing amount of previous experimental data. They highlighted three independent sets of GPCR–arrestin interactions that could trigger effector activation, all be it with different affinities and propensities. These are: interactions mediated by both the receptor core and the phosphorylated cytoplasmic tail, interactions involving the phosphorylated cytoplasmic tail only and interactions involving the receptor core only.
Importantly, the capability of the receptor core to induce arrestin stimulation was linked mainly to intracellular loop 2 (ICL2) and ICL3. From this standpoint, compared with the relatively low TM domain flexibility, the higher conformational space explored by ICLs can putatively provide a set of interactions able to engage the effector in a wider number of subtly different conformations, leading to the fine-tuning characteristic of GPCR signaling.

How does molecular dynamics address GPCRs allosterism?

While the GPCR transduction of chemical signals from the extracellular medium to the intracellular one can be considered unidirectional (i.e., external stimuli trigger an intracellular response), the structural mechanism through which this is realized is finely tuned in a bidirectional way. Indeed, the agonist and the intracellular effector reciprocally enhance their stability (and therefore affinity) for the receptor when bound to it [93]. Moreover, as integral membrane proteins, GPCRs communicate with the aqueous and the lipid environments, both contributing to the overall regulation and hetero/homodimerization in a number of different ways through the binding of lipid, cations and metabolites [93–95]. One of the most characterized GPCR endogenous allosteric mechanisms is exerted by sodium binding to a specific site, located in the core of the class A TM bundle below the orthosteric site, where the ion coordinates the side chains of conserved residues, including the Asp of the TM2 LXXXD motif [96], and exerts a negative allosteric effect by stabilizing inactive conformations [97]. Following the first structure resolved with the sodium-bound (the A2A AR [97]), MD simulations were extensively employed with the aim of uncovering the mechanism underlying the sodium allosteric effect [98–104]. Both the active and inactive conformations of the A2A AR were simulated in the presence and absence of bound cation [101]; the outcomes showed that the receptor active state is not able to stabilize the sodium, and that the cation’s first and second sphere of solvation heavily contribute to the overall structural effect by stabilizing the inactive conformation of the receptor. In a very recent work, Selvam et al. [105] computed the free energy profiles of sodium association to a set of 18 GPCRs. By means of extensive MD simulations, they revealed beside a conserved mechanism of binding, the possibility for the ion to access the allosteric site also from the intracellular vestibule of the receptors, at least for the purinoceptor P2Y12 and PAR1. Interestingly, during simulations on the class B GPCR glucagon receptor the sodium visited a putative binding site. However, the short RT observed is in accordance with the lack of experimental evidence (and structural motif) linking this class of GPCRs and ion-mediated allosteric effects. Indeed, while class B GPCRs share a number of structural motifs with class A [106], they lack the LXXXD motif.

Besides the sodium, it is well known that bivalent cations are able to modulate the binding of GPCRs. A positive allosteric effect has been observed for the ORs (opioid receptors) [107], the M2-R [108] and the A2A AR [109,110]. This latter was investigated in depth by Ye et al. [111], employing MD simulations to rationalize experimental data. Several putative binding sites for cations (Na\(^{+}\), K\(^{+}\), Ca\(^{2+}\), Mg\(^{2+}\) and Zn\(^{2+}\)) were located by considering acidic residues at the extracellular vestibule as well as residues involved in the switch from the active to the inactive state (like the class A GPCR ‘ionic lock’ – Figure 4).

Increasing attention has been focused on the development of GPCR allosteric modulators as potential therapeutics [112]. The reasons for this lie in the possibility of enhancing the response of an orthosteric drug or the natural ligand by selectively targeting the usually less conserved allosteric sites, and of reducing on-target side effects, by altering the signal bias [95,113]. Understanding the structural basis of allosteric modulation may enable the rational design of new agents. From this standpoint, the work by Dror et al. [114] represents a milestone. The authors reconstructed the putative binding of both positive allosteric modulators (PAMs) and negative allosteric modulators of the M2-R, showing that despite the structural diversity, the ligands were able to form similar interactions at the same site on the extracellular vestibule of the receptor. The occupation of this spot affected the shape of the orthosteric site and vice versa, thus explaining the cooperative behavior observed experimentally.

Allosteric sites were proposed by means of MD studies in the proximity of other GPCR orthosteric sites, as in the δOR (the PAM BMS-986187 putatively binds in a spot delimited by TM2, TM6 and TM7 [115]), or overlapping the canonical orthosteric one. According to this latter hypothesis, some PAMs could reach the orthosteric site and interact with the bound agonist further stabilizing it, as proposed for the A3 AR (adenosine receptor A3) [116] and, more recently, for the CB1R (cannabinoid receptor 1) [117].

MD simulations represent a powerful tool for investigating allosteric communications and how structural information propagates from allosteric sites to the interfaces with intracellular GPCR effectors [118]. From this standpoint, some recent approaches focused on correlated residue motions [119,120] and mutual information metrics [121]. In the GPCR field, Bhattacharya et al. mapped the allosteric communication pipelines responsible for the transmission of the activation signal from the orthosteric site to the intracellular interface of the β2-AR [122]. Interestingly, results
showed a higher dynamicty of the agonist-bound intermediate state of the receptor compared with the inactive or fully active one (the agonist-bound intermediate state is a conformation that, despite a TM structural organization resembling the inactive state, is characterized by side-chain conformational rearrangements observable in the fully active conformation). Moreover, upon binding of an agonist, the extracellular side of the receptor experienced a decoupling from the intracellular one. A further computational approach, able to detect protein pockets and the communication between adjacent pockets during MD simulations, has been proposed by La Sala et al. [123] and applied to the A2A AR. Their findings showed dynamic communication, between the orthosteric site and the sodium allosteric binding site, ascribable to residues located at the interface between the two cavities, able to exchange from one site to the other one.

**Solvation patterns in GPCRs**

The presence of conserved hydrated patterns along the GPCR TM domain should suggest a major role for water molecules in promoting conformational transitions. Active and inactive states are characterized by different spatial organizations of residues located in the core of the receptors and, in turns, by different stable networks of water-mediated interactions [124]. In this scenario, the binding of the orthosteric ligand and the intracellular effector may affect the water molecules accessibility to the GPCR core, besides triggering direct conformational changes responsible for activation/deactivation. It has been proposed that water molecules, flowing in the GPCR TM core from both membrane sides, contribute to activation of the β2AR (β2-adrenergic receptor) [125], the A2A AR [126,127], the β2-AR [126] and Rho (rhodopsin) [126]. Indeed, changes in the rotameric state of conserved amino acids upon activation putatively allow the formation of a hydrated channel that spans and stabilizes the receptors.

In drug design, the role played by water molecules in the whole ligand-receptor recognition event is increasingly considered. Different computational tools based on MD simulations have recently been developed with the aim of characterizing hydrated spots in druggable protein-binding sites [128–133]. WaterMap [128] was employed by Higgs et al. [134] for estimating the free energy profile of water molecules in the unoccupied orthosteric site of the A2A AR. This allowed the authors to infer a link between the structure–activity relationships (SARs) of the triazolopyrimidine antagonists and the extent to which the antagonists overlap several hydrated spots that were energetically classified as unfavorable. Among these, two sites occupied by ligand atoms were directly involved in ligand–protein hydrogen bonds. Effects on the solvation of the A2A AR binding site, produced upon binding, were investigated also for a set of triazine inhibitors [135]. In this study, Bortolato et al. proposed that the RT could be increased by displacing unstable water molecules that are trapped in the first hydration shell of the bound ligand. The same triazine – A2A AR complexes were also employed for the validation of AquaMMapS [132], a recent method developed to locate hydration spots in protein cavities by monitoring the solvent kinetic state [132]. The AquaMMapS protocol detects protein sites more prone to be occupied by water molecules with low root-mean-square fluctuation, distinguishing volumes that can be favorably occupied by a ligand’s hydrophilic moieties from the volumes where water molecules can be favorably displaced by hydrophobic substituents.

Besides the hydration patterns, the membrane environment also has an active role in the dynamic behavior of GPCRs [136]. Indeed, the protein-membrane interface can regulate the binding kinetics of lipophilic orthosteric ligands [49] as well as the overall receptor activation and signaling [137]. The effect that membrane lipids exert on the prototypic GPCR Rho has been extensively investigated [138]. The polyunsaturated docosahexaenoic acid is abundant in the photoreceptor cells and influences the Rho functionality by specifically bind to localized sites on the protein surface [139] that are more likely present when the Rho is in an inactive-like conformation [140]. Cholesterol was probably the first membrane component to be computationally addressed by means of MD simulations. On Rho, its action mechanism is still unclear. It has been proposed either to regulate the membrane order and thickness [139] or to specifically interact with Rho TM helices [141]. For the β2-AR, seven different cholesterol-binding sites have been reported on the receptor surface, spread on both the intracellular and the extracellular leaflets [142]. One spot in particular, which is located at the extracellular ends of TM1, TM2 and TM7 has been proposed to be involved in positive allosteric regulation of ligand binding. Moreover a distinct site, formed by TM1 and helix 8 (H8), putatively facilitates receptor dimerization, a phenomenon underlying receptor maturation and regulation [143] that is likely to involve TM4 in several class A GPCRs [144]. Cholesterol-binding sites were also located on the 5HT2A R (5-hydroxytryptamine receptor 2A) [145] and the A2A AR [146], where the lipid showed an increased simulated RT corresponding to three distinct spots without adopting a well-defined binding mode. The effects produced on the β2-AR activation mechanism by protein–lipid interactions was investigated by Neale et al. [147] and, more recently, by Bruzese et al. [148]. Both these studies proposed that the phospholipids putatively
regulate the activation or deactivation mechanism of the receptor, influencing the coupling to intracellular effectors. During MD simulations, negatively charged lipids were able to stabilize an active state of β2-AR, while neutral (zwitterionic) lipids generally favored the inactivation of the receptor, with an overall kinetics that was affected by the lipid headgroup charge distribution [148]. A detailed study of the positive effects exerted by the negatively charged phosphatidylinositol-4,5-bisphosphate on Gs activation by the A2a AR, β1-AR and the neurotensin receptor 1 has been reported recently [149].

**Class B CPCRs**

The small class B GPCR subfamily is composed of 15 receptors for endogenous peptides, structurally characterized by a large N-terminal extracellular domain (ECD – formed by up to 200 amino acids) connected to TM1 through a stalk region (Figure 5). The pharmacological profile of class B GPCRs is strongly influenced by the capability, for some of them, to heterodimerize with receptor activity-modifying proteins (RAMPs) [150], which are able to

---

**Figure 5.** The generic structure of a class B G-protein-coupled receptor bound to a G protein. The figure shows the GLP-1R in an active G-protein-bound state (Protein Data Bank code 6B3J). The N terminus of the agonist peptide binds in the TM bundle, while the C terminus interacts with the receptor ECD. The flexible stalk region and the ICL3 were not resolved in the cryo-electro microscopy map.

ECD: Extracellular domain; ICL: Intracellular loop; TM: Transmembrane.
modulate the selectivity as well as the intracellular signaling profile. The peptide binding is proposed to take place through a two-step mechanism [151]: the ligand C-terminus first binds the ECD, favoring the reorientation and the subsequent association of the peptide N-terminus into the receptor TM bundle (Figure 5). Since the early 2000s, several published structures of isolated ECDs (first), then isolated TMs, and more recently full-length receptors, have shed light on the different steps of this process [64, 152], giving a new impulse to drug design efforts [153, 154]. In this scenario, MD has been a helpful tool for rationalizing the wide amount of experimental data produced. The ECD of the apo GCGR (glucagon receptor) was observed in two preferred different orientations during simulations performed by Yang et al. [155]. According to their findings, the endogenous peptide stabilizes the open form of the receptor, therefore acting through a conformational selection mechanism. Simulations of the CGRPR (calcitonin-gene-related peptide receptor) bound to adrenomedullin [156] highlighted the possibility that conformational changes occurring at the ECLs upon agonist binding could be transmitted to the extracellular ends of TM5 and TM6, triggering, in turn, a reorganization of both hydrophobic and hydrophilic networks between TM3, TM5 and TM6. From this standpoint, RAMPs probably modulate the extracellular side of the TM bundle in a ligand-specific way. In the work from Bower et al. [157], the first step of amylin binding to the ECD of the CTR (calcitonin receptor), both alone and in complex with RAMP1, was simulated by means of SuMD simulations. The authors proposed that the allosteric regulation exerted by the RAMP1 comprises modification in the electrostatics of the receptor and in the binding path of the peptide C-terminus.

Considering the glucagon-like peptide-1 receptor (GLP-1R), simulation aimed at characterizing the role of the conserved polar residues located in the proximity of the TMs [158] and at rationalizing the peptide-mediated signaling of the GLP-1R, proposing a key role for the polar network located in the TM core [159]. The biased signaling of GLP-1R was later investigated in depth in the work by Wootten et al. [160]. MD simulations of salmon calcitonin (sCT) and hCT (human calcitonin) pointed out the ability of the former to better interact and stabilize the receptor ECL2, explaining the higher affinity of salmon calcitonin [161]. Weaver et al. [162] combined mutagenesis data and MD simulation on homology models to address the molecular basis of selectivity between the parathyroid hormone and tuberoinfundibular peptide 39 peptides binding to parathyroid hormone 2 receptor (PTH2R); an active truncated peptide was used due to uncertainties in the binding mode of the peptide N-terminus.

Conclusion
Since the first GPCR crystal structures were published in the early 2000s, several cutting-edge methodologies have overcome the intrinsic difficulties that characterize membrane protein crystallography. Moreover, thanks to recent advances in cryo-EM, the protein structure can be determined in a more native environment. This is leading to an outbreak of GPCR structural knowledge and, in turn, to a broad application of MD-related methods to the study of this membrane protein superfamily. MD simulations can be considered as the computational method of election for the study of nonequilibrium phenomena, such as drug-receptor (un)binding events, the GPCR activation, and the allosteric structural communications. Moreover, MD simulations allow modeling the protein–solvent interfaces at atomic scale, enabling the study of how water and lipid molecules can influence the GPCRs.

Future perspective
Considering the continuous improvement in the informatics field, and the increasing computational capability, it is plausible that in the next few years MD will be increasingly used to translate the massive amount of static information into the dynamic language of biological systems. Thus, allowing a better description of several biological aspects that are still far from clear. In particular, the structural determinants for both biased signaling and allosterism may represent a more challenging topic in the GPCR field, and therefore the main targets of future computational efforts.

Financial & competing interests disclosure
CAR is a Royal Society Industry Fellow and acknowledges support from the BBSRC (BB/M006883/1). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.
Executive summary

- Recently, both experimental structural determinations and molecular dynamics (MD)-related methods are providing astonishing advances in the G-protein-coupled receptor (GPCR) field. MD allows the time-dependent (dynamic) description of biological systems at an atomic scale.
- The first step in the GPCRs signaling is the extracellular ligand recognition: MD is a valuable tool for investigating this process from the non equilibrium point of view.
- MD simulations are shedding light on the GPCR activation mechanism, taking advantage of recent insights from x-ray crystallography and cryo-electron microscopy techniques.
- Biased agonism, which can potentially lead to groundbreaking improvement in the design of new drugs, is still a poorly understood phenomena; MD simulations may contribute to an explanation of this aspect from a structural point of view.
- To date, different allosteric mechanisms have been suggested by means of MD simulations.
- MD investigations of class B GPCRs are gaining importance since the first experimental full-length active structures have been recently published.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest


Pande VS, Beauchamp K, Bowman GR. Everything you wanted to know about Markov state models but were afraid to ask. Methods 52(1), 99–105 (2010).


**A nonequilibrium point of view on biased agonism.**


**First computational insights on a GPCR-biased agonist.**


Peeking at G-protein-coupled receptors through the molecular dynamics keyhole

Review


A comprehensive investigation on GPCR sodium allostery.


108 Burgmer U, Schulz U, Tränkle C, Mohr K. Interaction of Mg


110 Mazzoni MR, Martini C, Lucachini A. Regulation of agonist binding to A2A adenosine receptors: effects of guanine nucleotides (GDP[S] and GTP[S]) and Mg


Deganutti, Moro & Reynolds


- An exhaustive update on class B GPCR hot topics.
Peeking at G-protein-coupled receptors through the molecular dynamics keyhole


