Proteins: Structure, Function, and Bioinformatics

WILEY

PROTEINS

RESEARCH ARTICLE

Unraveling GPCRs Allosteric Modulation. Cannabinoid 1 Receptor as a Case Study

Alejandro Cruz | Arieh Warshel

Department of Chemistry, University of Southern California, Los Angeles, California, USA

Correspondence: Arieh Warshel (warshel@usc.edu)

Received: 20 August 2024 | **Revised:** 14 October 2024 | **Accepted:** 18 October 2024

Funding: This work was supported by National Institutes of Health (R35 GM122472), National Science Foundation (MCB 1707167).

Keywords: binding free energy calculations | cannabinoid type-1 receptor | cannabinoid type-1 receptor positive allosteric modulation | cooperativity factor | dissociation constant | folding free energy calculations | GPCRs allosteric modulation

ABSTRACT

G-protein-coupled receptors (GPCRs) constitute one of the most prominent families of integral membrane receptor proteins that mediate most transmembrane signaling processes. Malfunction of these signal transduction processes is one of the underlying causes of many human pathologies (Parkinson's, Huntington's, heart diseases, etc), provoking that GPCRs are the largest family of druggable proteins. However, these receptors have been targeted traditionally by orthosteric ligands, which usually causes side effects due to the simultaneous targeting of homologous receptor subtypes. Allosteric modulation offers a promising alternative approach to circumvent this problematic and, thus, comprehending its details is a most important task. Here we use the Cannabinoid type-1 receptor (CB1R) in trying to shed light on this issue, focusing on positive allosteric modulation. This is done by using the protein-dipole Langevin-dipole (PDLD) within the linear response approximation (LRA) framework (PDLD/S-2000) along with our coarse-grained (CG) model of membrane proteins to evaluate the dissociation constants (K_Bs) and cooperativity factors (αs) for a diverse series of CB1R positive allosteric modulators belonging to the 2-phenylindole structural class, considering CP55940 as an agonist. The agreement with the experimental data evinces that significantly populated allosteric modulator:CB1R and allosteric modulator:CP55940:CB1R complexes have been identified and characterized successfully. Analyzing them, it has been determined that CB1R positive allosteric modulation lies in an outwards displacement of transmembrane α helix (TM) 4 extracellular end and in the regulation of the range of motion of a compound TM7 movement for binary and ternary complexes, respectively. In this respect, we achieved a better comprehension of the molecular architecture of CB1R positive allosteric site, identifying Lys192$^{3.28}$ and Gly194$^{3.30}$ as key residues regarding electrostatic interactions inside this cavity, and to rationalize (at both structural and molecular level) the exhibited stereoselectivity in relation to positive allosteric modulation activity by considered CB1R allosteric modulators. Additionally, putative/postulated allosteric binding sites have been screened successfully, identifying the real CB1R positive allosteric site, and most structure–activity relationship (SAR) studies of CB1R 2-phenylindole allosteric modulators have been rationalized. All these findings point out towards the predictive value of the methodology used in the current work, which can be applied to other biophysical systems of interest. The results presented in this study contribute significantly to understand GPCRs allosteric modulation and, hopefully, will encourage a more thorough exploration of the topic.

Abbreviations: 2-AG, 2-arachidonoylglycerol; AEA, anandamide; ago-PAM, agonist and positive allosteric modulator; CB1R, cannabinoid type-1 receptor; CG, coarse-grained; cryo-EM, cryogenic electron microscopy; ECL, extracellular loop; GPCRs, G-protein-coupled receptors; ICL, intracellular loop; LRA, linear response approximation; LRF, local reaction field; MCPT, Monte Carlo Proton Transfer; NAM, negative allosteric modulator; PAM, positive allosteric modulator; PDLD, protein-dipole Langevin-dipole; RMSD, root-mean-square deviation; SAR, structure–activity relationship; THC, Δ^9-tetrahydrocannabinol; TM, transmembrane α helix

1 | Introduction

G-protein-coupled receptors (GPCRs) constitute a superfamily of integral membrane proteins that exhibit a characteristic seven transmembrane α helices (TMs) structure, involved in the mediation of the majority of intracellular responses to external stimuli, such as light and hormones, giving rise to one of the most outstanding receptor families in the human body [1, 2].

GPCRs are not simple on/off switches since there is a preexisting equilibrium between their inactive and active states through intermediate states, which can be associated with distinct conformational ensembles and functions, respectively (see Figure 1). Ligand binding can shift this equilibrium in such a way that agonists stabilize active state conformations, inverse agonists favor inactive state conformations and antagonists do not favor any particular state but compete with both agonists and inverse agonists for their binding site [3]. GPCRs in their active state can trigger signal transduction processes through a variety of G protein-dependent and -independent pathways. In contrast, inactive state conformations are unable to trigger them (may trigger an opposite action). G protein-dependent pathways constitute GPCRs canonical signaling pathways, while G protein-independent pathways, with β-arrestin one being the most prominent [4], are mainly related to GPCRs internalization and adverse-event risk [3, 5–7]. Moreover, as discussed below, GPCRs can be considered as allosteric proteins since binding of orthosteric ligands, allosteric modulators and G proteins is interconnected, modifying their exhibited affinities/efficacies [3].

Malfunction of these signal transduction processes is one of the underlying causes of many human pathologies, pointing out neurological/neurodegenerative and heart diseases, that are not completely solved [8, 9]. The importance of proper signal transduction makes GPCRs the largest family of druggable proteins in the human genome, being targeted by more than 30% of the approved drugs by the Food and Drug Administration (FDA) [10, 11]. However, only a small fraction of the possible druggable GPCRs has been exploited pharmacologically [2]. There are several reasons behind this low ratio of targeted GPCRs, being one of them that the design of new ligands has traditionally targeted the orthosteric site [12], which is highly conserved, causing side effects due to the multiple target of homologous receptor subtypes [13, 14]. A promising alternative approach to circumvent this problematic is allosteric modulation [15]. Allosteric modulators are ligands which bind to a spatially and

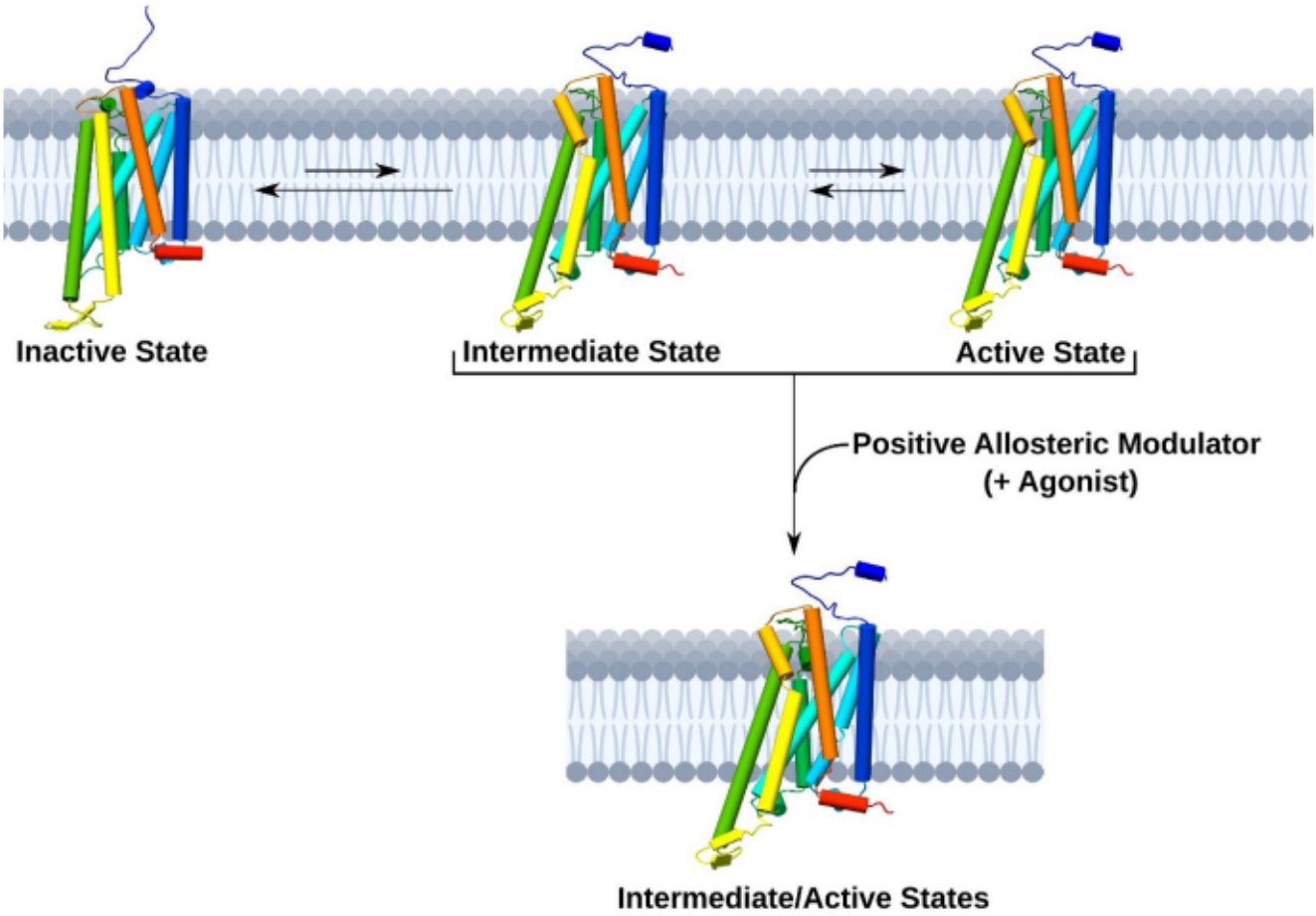

FIGURE 1 | Schematic representation of GPCRs preexisting equilibrium between their inactive, intermediate and active states. This equilibrium is generally shifted towards inactive state when no ligand acts on GPCRs. However, positive allosteric modulators possess the ability of shifting this equilibrium towards intermediate and active states acting alone or together with some agonist. To illustrate different states exhibited by GPCRs, it has been employed the CB1R corresponding ones. This receptor has been represented with pipes and planks employing a rainbow color scale. Similarly to the vast majority of GPCRs, it can be seen that a TM6 (in yellow) outward movement takes place throughout the activation process.

topologically distinct site (the allosteric site) from the orthosteric binding site [16, 17] modulating protein activity. They offer several advantages over orthosteric ligands: (1) Subtype specificity, since allosteric sites tend to be less conserved; (2) Fine-tuning of orthosteric-ligand response, in a time and spatially dependent manner; (3) Signaling bias, favoring beneficial/therapeutic signaling pathways; (4) Probe dependence, since allosteric modulators are orthosteric-ligand sensitive [18, 19]. From a mechanistic point of view, allosteric modulators potentiate or attenuate the orthosteric-ligand affinity/efficacy acting as positive allosteric modulators (PAMs) or negative allosteric modulators (NAMs), respectively. Thus, PAMs shift the preexisting equilibrium towards GPCRs intermediate conformational states that have high affinity/efficacy regarding orthosteric ligands and great structural similarity with regard to active states. In addition, some allosteric modulators with positive activity can exhibit intrinsic efficacy by themselves, acting as pure allosteric agonists, or even, both intrinsic efficacy and potentiation of orthosteric agonists, acting as agonist and positive allosteric modulators (ago-PAMs). It is worth noticing that both pure allosteric agonists and ago-PAMs have the ability of stabilizing active state conformations which differ from the orthosteric-agonist-induced ones since they show different pharmacological and biological profiles [3, 7, 12, 20–23].

To progress in the direction pointed above, it is appealing to achieve a deeper understanding of GPCRs allosteric modulation. This paper aims to shed light on this topic focusing on activation processes, especially in their comprehension at structural and molecular level and the calculation of allosteric modulator binding parameters. In order to carry out this task, it is of utmost importance to be able to distinguish between the real allosteric site and the putative/postulated ones. This will help to establish structural differences between the induced conformational states by orthosteric agonists, and pure allosteric agonist and PAMs, to determine the origin at molecular level of pure PAM and allosteric agonist activities. This should allow to calculate in a reliable way the affinity and cooperativity of allosteric modulators. In other words, to evaluate their dissociation constants (K_Bs) and cooperativity factors (αs).

Cannabinoid type-1 receptor (CB1R) has been thoroughly studied due to its great relevance since this receptor is wide distributed throughout the human body, being the most abundant GPCR in the central nervous system [24, 25], it regulates diverse physiological processes, such as neurotransmission and synapse formation, and it is related to prevalent unsolved medical conditions. These include neuropathic and inflammatory pain, multiple sclerosis and Huntington's and Parkinson's diseases [26]. It is worth noticing that CB1R is best-known for its recreational purposes that arise from consuming its famous exogenous agonist Δ^9-tetrahydrocannabinol (THC), which is the active compound of cannabis. However, CB1R primary endogenous agonists, anandamide (AEA) and 2-arachidonoylglycerol (2-AG), have sparked less interest [24, 25, 27]. Additionally, recreational users of cannabis are familiar with CB1R regulation, possibly allosteric, since they tend to consume THC along with cannabis terpenes due to their synergy. The response is increased 2–10 times in comparison with pure THC [28]. It is important to emphasize that much more wider and deeper knowledge about CB1R regulation than that is available, specifically about its positive allosteric modulation. This makes CB1R an ideal system to study GPCRs positive allosteric modulation for the following reasons: (1) The availability of experimental structures of CB1R complexes, in particular the structure of the ZCZ011(*S*):CP55940:CB1R:G_i complex (PDB code 7WV9) [29] due to the scope of the present study; (2) Previous knowledge about CB1R allosteric sites and modulators [7, 20–23, 29–31]. In the present study we report the results of a computational study using advanced modeling and simulation tools of different ternary or binary complexes of the type allosteric modulator:orthosteric agonist:CB1R and allosteric modulator:CB1R, using the aforementioned quaternary complex as reference of active state. Due to the scope of the present study, only this reference CB1R active state has been considered since all sorts of allosteric modulators with a positive activity induce either active states with subtle structural differences regarding these states induced by orthosteric ligands or intermediate states that have high affinity/efficacy regarding orthosteric ligands and great structural similarity with regard to active states. For the different constructed complexes, CP55940 has been the only orthosteric agonist considered, whereas diverse allosteric modulators including ZCZ011(*S*), ZCZ011(*R*), GAT228, GAT229, the racemic mixture of the two previous compounds (GAT221), both enantiomers of a trifluoro analog of GAT221 (6s), GAT1600, and GAT1601 have been considered (see Figure 2). These compounds have been selected because they are well-characterized experimentally, that is, their K_Bs and αs have been reported, or at least, their half maximal effective concentrations (EC_{50}s) along with their shifts (ΔEC_{50}s) due to cooperativity effects (more detailed in Sections 3.1 and 3.2). It is worth noting that these allosteric modulators are characterized for exhibiting few polar interactions with the receptor, allowing them to explore their different binding modes/poses without involving huge free energy barriers. Therefore, a Boltzmann distribution will govern relative populations of distinct binding modes for each allosteric modulator. Using the protein-dipole Langevin-dipole (PDLD) within the linear response approximation (LRA) framework (PDLD/S-2000) method [32] and assuming a Boltzmann distribution for allosteric-modulators binding modes with state energies given by the total free energy of the corresponding complex (see Sections 3.1 and 3.2), K_Bs and αs have been calculated for all possible binary and ternary complexes derived from the aforementioned compounds, except for both enantiomers of GAT211 trifluoro analog for which only K_B has been calculated. The close agreement obtained for K_Bs and αs evinces that our employed methodology is effective to determine relative stabilities of different binary and ternary complexes in a reliable way and to carry out a quantitative assessment of allosteric-modulators binding free energies (ΔG_{bind}s). This allows us to determine their main binding modes and opening the door to the optimization and/or rational design of allosteric modulators.

On the other hand, through analysis of principal binding modes of considered compounds, both the structural basis and the origin at molecular level of CB1R positive allosteric modulation have been determined, which was the main objective of the present study, allowing us to understand how CB1R allosteric modulators belonging to the 2-phenylindole structural class exert their action, to better comprehend the molecular architecture of the CB1R positive allosteric site and to rationalize the stereoselectivity and most structure–activity relationship (SAR)

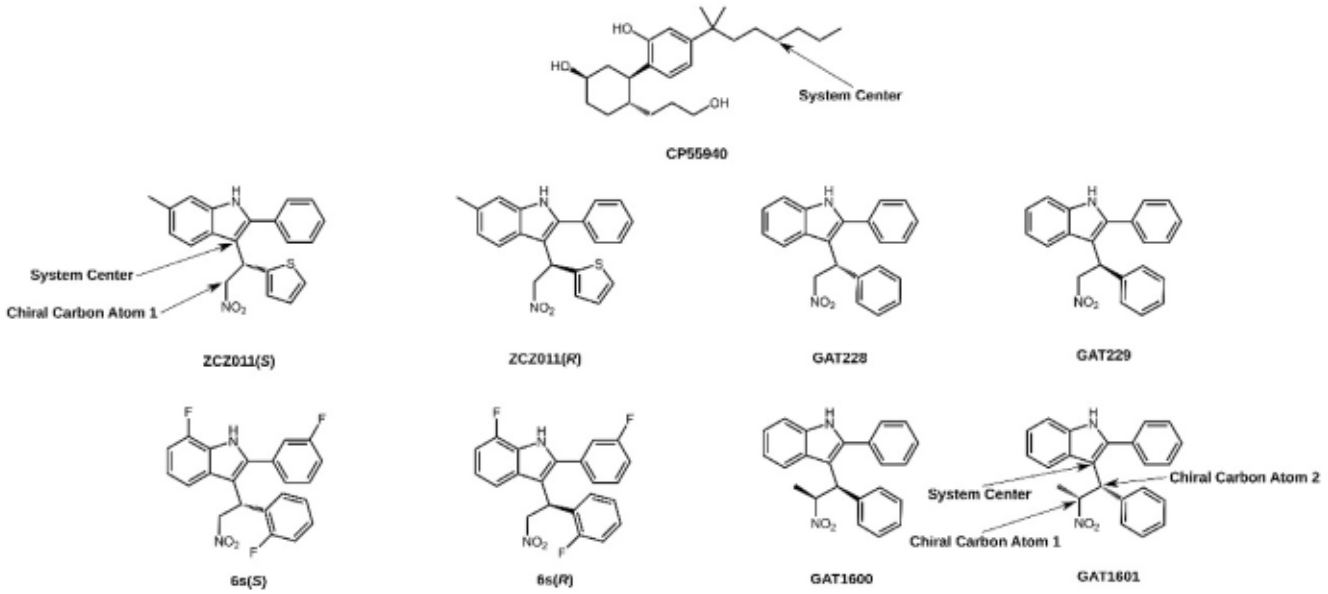

FIGURE 2 | Depiction of agonist (first row) and allosteric modulators (second and third rows) considered in the present study. All allosteric modulators belonging to the 2-phenylindole structural class contain a chiral carbon atom at β position of the NO_2 group (chiral carbon atom 1; see ZCZ011(*S*)), but substitution at α position of the NO_2 group gives rise to an additional one (chiral carbon atom 2: See GAT1601). For both the agonist and allosteric modulators with one and two chiral carbon atoms, the center of the system for relaxation and equilibration MD simulations and PDLD/S-LRA simulations has been indicated. It is worth noticing that ZCZ011, GAT221 and 6s are racemic mixtures of ZCZ011(*S*) and ZCZ011(*R*), GAT228 and GAT229, and 6s(*S*) and 6s(*R*), respectively. Absolute configuration of allosteric-modulators chiral centers: GAT228 (*R*), GAT229 (*S*), GAT1600 (*S*, *R*), and GAT1601 (*R*, *S*) (the rest of compounds has not included in the list because of their name are self-explanatory).

studies of such compounds. Additionally, our methodology also turned out to be effective for screening of putative/postulated CB1R positive allosteric sites [30, 31] since the real one was identified.

Finally, it is worth noticing that due to the presence of cellular membrane, the modeling of such biophysical systems is difficult because of their size and complexity, restricting the usage of all-atoms models for computational reasons. In order to overcome this computational limitation, an implicit model for cellular membrane has been employed [33]. Additionally, we used our developed coarse-grained (CG) model [34, 35] for membrane proteins for evaluating CB1R conformational energies. Both simplified models have been proved to be very effective in investigating other GPCRs [36, 37].

2 | Methodology Section

2.1 | Systems Setup

In order to study the CB1R positive allosteric modulation by allosteric modulators belonging to the 2-phenylindole structural class, we used the cryo-EM structure of the ZCZ011(*S*):CP55940:CB1R:G_i complex (PDB code 7WV9) [29]. This structure corresponds to the unique experimental structure with one of such compounds bound to CB1R in its fully active state. In order to facilitate comparisons with other GPCRs, Ballesteros-Weinstein nomenclature [38] has been employed. It is worth noting that only CB1R active state has been considered given that allosteric modulators with positive activity induce either active states by themselves or intermediate states structurally similar to active ones that exhibit large affinity/efficacy for orthosteric agonist. Regarding considered structure, orthosteric and allosteric ligands along with the heterotrimeric G protein have been removed and a missing gap of intracellular loop (ICL) 3 has been completed employing the corresponding structure to a thermostabilized version of CB1R that retains the ability to bind both ZCZ011 and CP55940 (PDB code 7FEE) [29]. The resulting structure has been protonated with the Monte Carlo Proton Transfer (MCPT) method implemented in Molaris-XG [39] using a pH = 7.0. To ensure convergence of the electrostatic free energy for the folded protein, 10^7 MCPT steps were carried out. From now on, all calculations were carried out using the Molaris-XG package [39] unless otherwise mentioned. The allosteric modulator:CB1R, CP55940:CB1R and allosteric modulator:CP55940:CB1R complexes have been generated by including allosteric and/or orthosteric ligands inside their respective cavities in their corresponding binding modes/poses (see Figure 3). As an orthosteric ligand (agonist) has only been considered CP55940, while as allosteric modulators have been considered ZCZ011(*S*), ZCZ011(*R*), GAT228, GAT229, GAT1600, GAT1601, 6s(*S*), and 6s(*R*) (see Figure 2). In the case of CP55940, only its exhibited pose in the cryo-EM structure of the ZCZ011(*S*):CP55940:CB1R:G_i complex (PDB code 7WV9) [29] has been taken into account, while for allosteric modulators, 3 main distinct binding modes identified by docking calculations have been considered along with the binding mode exhibited by ZCZ011(*S*) in its cryo-EM structure (see Figure 4). For incorporating cellular membrane effects, after inclusion of allosteric and/or orthosteric ligand/s, a 90 × 90 × 16 Å membrane particle grid was added around the Z-axis of CB1R transmembrane bundle whose center of mass shifted 1.6 Å throughout this axis is the origin of that grid. The

partial charges of CP55940 and allosteric modulators were calculated at the CAM-B3LYP/6-31G* level of theory using Gaussian 16 [40]. These resulting charges were fitted to generate the restrained electrostatic potential (RESP)-fitted charges [41] by Ambertools [42]. The ENZYMIX force field [32] was used for the remaining system.

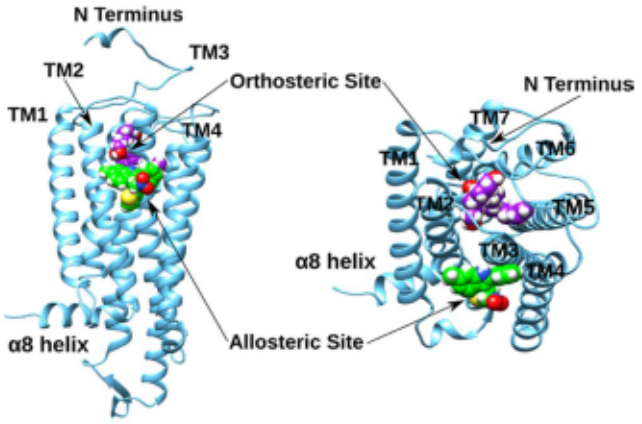

FIGURE 3 | Schematic representation of orthosteric and allosteric CB1R sites through a CPK depiction of the corresponding ligands bound to such binding sites (CP55940 bound to the orthosteric site in purple and ZCZ011(*S*) bound to the CB1R positive allosteric site in green).

2.2 | Docking Calculations

The docking calculations of the allosteric modulators inside the CB1R positive allosteric site were carried out by a two-step process: (1) Determination of main binding modes restraining the allosteric modulator center of mass to be a distance of 4.0 Å from the oxygen backbone atom of Phe191$^{3.27}$ since it is well-known that a hydrogen bond is formed between this atom and the NH group of the 2-phenylindole ring; (2) Optimization of resulting main binding modes performing 1000 Monte Carlo energy minimizations. The receptor remained rigid, but total flexibility was given to allosteric modulators during the conformational exploration.

2.3 | Relaxation and Equilibration Molecular Dynamics (MD) Simulations

All relaxation and equilibration MD simulations followed the same protocol, the only differences between them were the starting structure and the unconstrained system region to be relaxed and equilibrated. The starting structures have been taken as the assembled complexes mentioned in Section 2.1. The center of this unconstrainted region was conveniently chosen as an atom of the ligand (allosteric modulator or CP55940) whose binding free energy (ΔG_{bind}) was assessed

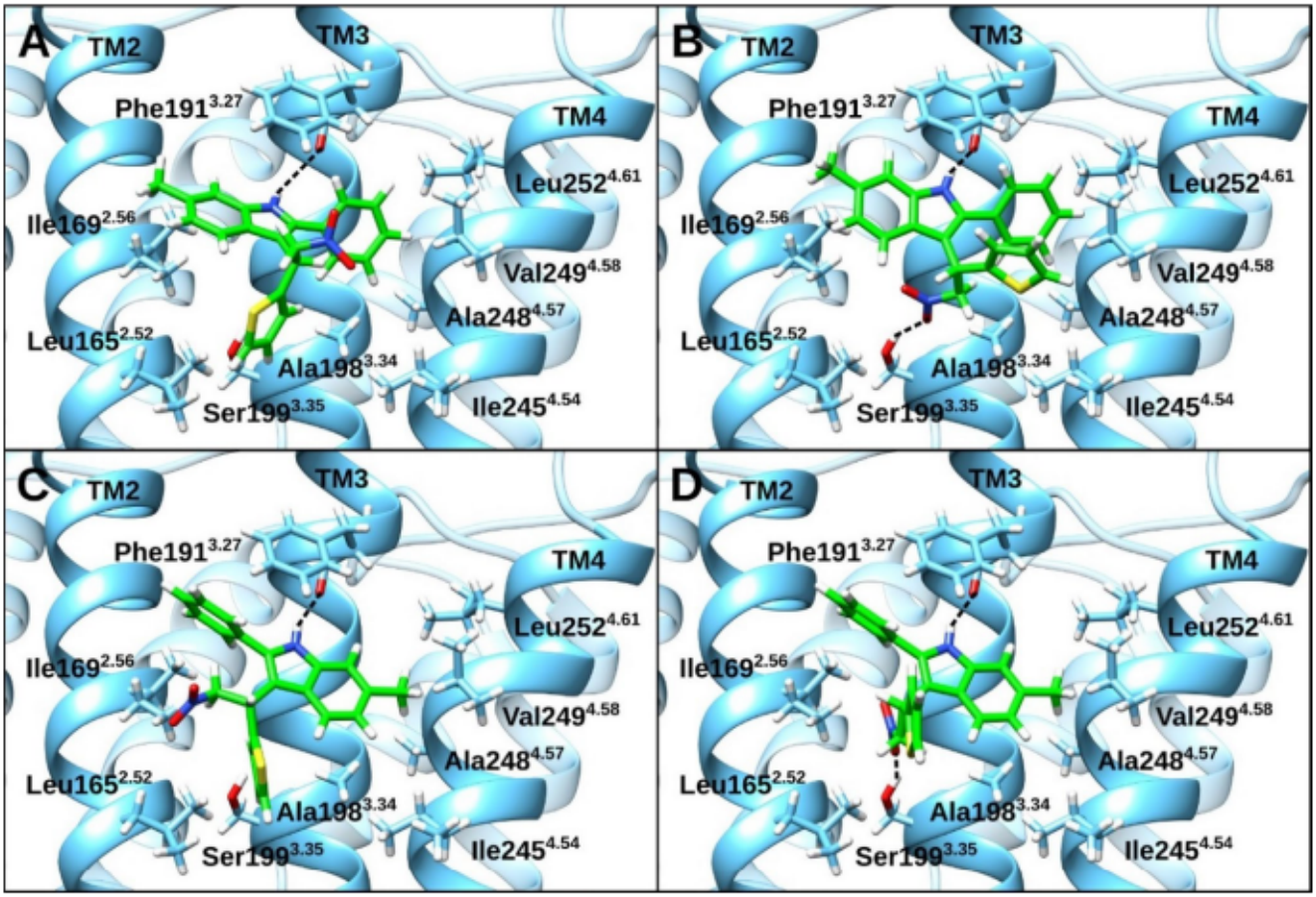

FIGURE 4 | Representation of the main binding modes explored for the considered compounds illustrated by ZCZ011(*S*) (in green). Panels A, B, C, and D correspond to binding modes 1, 2, 3, and 4, respectively. All binding modes exhibit a hydrogen bond between the NH group of the 2-phenylindole ring and the oxygen backbone atom of Phe191$^{3.27}$, while only binding modes 2 and 4 show a hydrogen bond between the NO_2 group and Ser199$^{3.35}$. Hydrogen bonds have been depicted by dash lines.

subsequently. In particular, this center was defined by the substituted indole C_β of allosteric modulator and the central carbon atom of the CP55940 hydrophobic tail (see Figure 2) when allosteric modulators and CP55940 were considered in the evaluation of ΔG_{btnd}, respectively. This choice of system centers was taken to consider allosteric and orthosteric sites simultaneously for ternary complexes without employing an excessively large unconstrainted region of the system. For binary complexes, the system was immersed in a surface-constrainted all-atoms solvent (SCAAS) model [43] water sphere with a radius of 20 Å around the chosen center. This radius was enlarged to 30 Å for ternary complexes to consider both binding sites simultaneously. These resulting spheres constitute the unconstrainted system region to be relaxed and equilibrated, which were surrounded by a 2 Å spherical shell of Langevin dipoles and then a bulk continuum. The water molecules at the boundary of this region were subjected to polarization and radial restraints [43]. The local reaction field (LRF) method [44] was employed to treat the long-range electrostatic interactions, while these interactions from the outside of the unconstrainted region were not included in the energy calculations. Once the system was assembled and solvated, it was submitted to 2000 energy minimizations steps using the steep-descent method to avoid close contacts. Then, the system was gradually heated from 0 to 300 K for a period of 200 ps. Throughout this heating, harmonic restraints were applied to the protein backbone and ligand/water heavy atoms with a force constant of 10 kcal mol^{-1} Å^{-2} that were stepwise released during the first 100 ps of this period, with the exception of the CP55940 ones that were loosened and kept at 3 kcal mol^{-1} Å^{-2} to avoid clashes due to the great flexibility of this compound. Finally, the starting geometry of binary and ternary complexes for subsequent calculations were generated from a 200 ps MD equilibration simulation. It is worth mentioning that the equilibration of a given system is considerable fast employing spherical boundary conditions, which is reflected in the minimal fluctuation of the root-mean-square deviation (RMSD) of heavy atoms observed during this period (see Figures S1–S8).

2.4 | PDLD/S-LRA Simulations

The simulation setup for ΔG_{btnd} calculations was the same as that for relaxation and equilibration MD simulations, with the exception that in the PDLD/S-LRA simulations the unconstrainted system region was divided into two regions. Region 1, for which ΔG_{btnd} was calculated, was defined by the whole allosteric modulator or CP55940, while region 2, which comprises the system region that contributes to that ΔG_{btnd}, was defined by the remaining surrounding system. In the present work, a modified semimacroscopic version of the PDLD method within LRA framework, which is termed PDLD/S-2000 [45], has been employed for each type of considered ligand. This approach, which was introduced in reference [45], evaluates the non-electrostatic term explicitly by a special cycle, while the electrostatic term is evaluated by the LRA approach.

For allosteric modulators, The modified approach involves scaling the non-electrostatic term as follows:

$$\Delta G_{btnd}^{allo} = \Delta G_{btnd}^{elec} + C\Delta G_{btnd}^{non-elec} \quad (1)$$

where ΔG_{btnd}^{allo} refers to the total binding energy of allosteric modulator, ΔG_{btnd}^{elec} and $\Delta G_{btnd}^{non-elec}$ correspond to the electrostatic and non-electrostatic contributions of ΔG_{btnd}^{allo}, respectively, and C is the scaling factor of this last contribution. The scaling factor C is not a universal constant, but is specific to the studied system, influenced by the interplay between different thermodynamic contributions such as Van der Waals interactions, hydrophobic effects, water penetration, and configurational entropy changes [46]. An optimal C value of 0.9 has been determined, which effectively calibrates the calculated ΔG_{btnd}^{allo} to reproduce the experimental K_Bs and αs (see Sections 3.1 and 3.2).

In contrast, for CP55940 we further modified our approach using:

$$\Delta G_{btnd}^{CP55940} = \Delta G_{btnd}^{elec} + \Delta G_{btnd}^{non-elec} + K \quad (2)$$

where $\Delta G_{btnd}^{CP55940}$ is the total binding free energy of CP55940 and K is the corresponding constant which allows to reproduce the experimental value of $\Delta G_{btnd}^{CP55940}$ and to calculate this magnitude in presence of allosteric modulators with enough accuracy to correctly describe the total free energy of ternary complexes (see Section 3.2).

For each previously equilibrated complex, a PDLD/S-LRA simulation was carried out on 4 different replicas. Regarding binary complexes, the 4 replicas had distinct initial configurations obtained through four successive equilibration MD simulations of 10 ps. For binding free energy in protein, 500000 different configurations were considered, while for the reference in water 50000 different configurations were assessed. Due to computational limitations, PDLD/S-LRA simulations for ternary complexes had to be reduced by a factor of 10. It is important to note that even the shorter length of the PDLD/S-LRA simulations employed for ternary complexes is more than enough to achieve a converged ΔG_{btnd} according to previous benchmarks [47].

2.5 | Coarse-Grained (CG) Model of CB1R

Our refined coarse-grained (CG) model has been employed [33, 35] with the aim of determining the CB1R conformational free energy (ΔG_{fold}) for binary and ternary complexes. The total energy of our CG model is given by:

$$\begin{aligned}\Delta G_{fold} &= \Delta G_{main} + \Delta G_{side} + \Delta G_{main-side} \\ &= c_1\Delta G_{side}^{VdW} + c_2\Delta G_{solv}^{CG} + c_3\Delta G_{HB}^{CG} + c_4\Delta G_{mem}^{hyd} + \Delta G_{side}^{hyd} \\ &+ \Delta G_{main-side}^{VdW} + \Delta G_{side}^{elec} + \Delta G_{side}^{polar}\end{aligned} \quad (3)$$

Here the total conformational free energy ΔG_{fold} consists of three terms, the main chain term ΔG_{main}, the sidechain term ΔG_{side} and the interaction term between main and side chain $\Delta G_{main-side}$, which can be broken down as specified in the second raw on Equation (3). The terms on the right of this raw are the sidechain Van der Waals energy, the main chain solvation energy, the main chain hydrogen bond energy, the cellular membrane hydrophobic energy, the sidechain hydrophobic

energy, the main chain/sidechain Van der Waals energy, the sidechain electrostatic energy, and the sidechain polar energy, respectively. The scaling coefficients c_1, c_2, c_3, and c_4 take values of 0.10, 0.25, 0.15, and 1/3.6, respectively. It is worth noting that Equation (3) has successfully been employed to determine ΔG_{fold} in other GPCRs [36, 37]. To determine CB1R ΔG_{fold} for a given complex in a reliable way, a sampling of 10 different conformations collected at every 10 ps of successive gas-phase equilibration MD simulations from the corresponding previously equilibrated complex was considered. Prior to this, a minimization, a heating process, and a relaxation MD simulation have been carried out. Apart from being in gas phase, these simulations have employed the same simulation setup as stated in Section 2.3. The resulting structures have been trimmed into the CG representation, in which the main chain of amino acids was still in all-atom form, but each of their sidechains were reduced into a simplified united atom. Next, another small relaxation has been performed and present ligands have been removed before evaluating ΔG_{fold}. All visualizations and representations have been made with the VMD [48] and UCSF Chimera [49] programs.

3 | Results and Discussion

3.1 | Dissociation Constants and Principal Binding Modes for the Allosteric Modulators of CB1R

From a pharmacological perspective, determining the dissociation constant, both experimentally and computationally, for active compounds in a reliable way is of vital importance. That is why this parameter allows to estimate their affinity, in other words, EC_{50} or IC_{50} depending on whether the considered compound has a positive or negative activity, respectively. In order to determine experimentally the dissociation constant for allosteric modulators (K_B) is commonly employed the allosteric binding model [50, 51] combined with the Black-Leff operational model for agonism (OMAM) [52–55], which describes the allosteric modulation over a given agonist (the orthosteric ligand CP55940 in our case). Apart from K_B, this model also provides other relevant parameters for allosteric modulation. Including the allosteric modulation efficacy (τ_B), which measures the own efficacy of allosteric modulator, the efficacy modulation (β), which measures how efficacy of the considered agonist changes due to the allosteric modulator presence, and the cooperativity factor (α), which measures K_B change due to the agonist presence and vice versa. It is worth noting that all these parameters are interdependent so that incorrect determination of one will affect all the others. In order to apply the aforementioned fitting technique, data of concentration-response curves for agonist at different concentrations of allosteric modulator are needed, which is demanding from an experimental point of view. For this reason, experimental determination of EC_{50} (instead of K_B) is quite common since the approximation $EC_{50} \approx K_B$ is well-accepted, and this determination only requires a single concentration-response curve for allosteric modulator. On the other hand, the calculation of binding free energy (ΔG_{bind}) allows to determine both K_B and α (see Section 3.2). Regarding K_B, its relationship with ΔG_{bind} is given by the following expression:

$$\Delta G_{bind} = RT \ln \left(K_B \right) \quad (4)$$

where R and T are gas constant and temperature, respectively. With the aim of validating our calculation method of ΔG_{bind} for CB1R positive allosteric modulators belonging to the 2-phenylindole structural class (ΔG^{allo}_{bind}). This magnitude has been calculated for compounds GAT228, GAT229, GAT221 (racemic mixture of previous two compounds), GAT1601 and GAT1601 since these are well-characterized experimentally. In particular, the experimental K_Bs, τ_Bs, βs and αs have been determined by the aforementioned fitting technique [55]. In addition, although the experimental ZCZ011 K_B is not available explicitly, this compound has also been considered since ZCZ011 is the only CB1R ago-PAM for which there is an experimental structure available (ZCZ011(*S*):CP55940:CB1:G_i complex; PDB code 7WV9) [29]. For this particular compound, its K_B on CB1R has been estimated from its EC_{50} on such receptor obtained by a G protein dissociation assay, which theoretically should be the most accurate way [56] (see Supporting Information for justification of this choice).

Considering ZCZ011(*S*) experimental binding mode/pose on the ZCZ011(*S*):CP55940:CB1R:G_i complex, the NO_2 group seems not to be relevant since it does not exhibit any interaction. However, for CB1R allosteric modulators belonging to the 2-phenylindole structural class, the NO_2 group turns out to be crucial, since its replacement by other polar/charged groups reduces, or even abolishes, the compound activity [57, 58]. Additionally, it has been reported that $Ser199^{3.35}$ plays an important role in the ZCZ011 allosteric effect on CB1R [29] and the unique group that can interact with this residue is its NO_2 group. Taking everything into account, relevant binding modes for allosteric activity in which the NO_2 group and $Ser199^{3.35}$ interact with each other must exist. In order to explore them, docking calculations have been performed for the aforementioned compounds on the CB1R positive allosteric site. Interestingly, apart from these poses, these docking calculations also suggest the existence of stable binding modes in which the 2-phenylindole ring is placed at TM3 right side, taking as a reference system a frontal view of CB1R in which its extracellular domain is placed above the intracellular one (this reference system is being employed from now on). For this reason, for each considered compound, it has been inspected 4 main binding modes/poses. Specifically, binding mode 1 (Panel A of Figure 4) corresponds to the experimental pose of ZCZ011(*S*) in which the 2-phenylindole ring is placed at TM3 left side while the NO_2 group points towards the cellular membrane, while binding modes 2, 3, and 4 (Panel B, C, and D of Figure 4) arise from flipping the NO_2 group, the 2-phenylindole ring and these two groups at the same time, respectively.

Regarding the calculation of K_B, given that the CB1R positive allosteric site is superficial and allosteric modulator binding is mainly due to hydrophobic interactions, a thermodynamic equilibrium will be established among different allosteric modulator poses since these compounds will have the ability of unbinding from the allosteric site and explore all their possible binding modes without involving huge free energy barriers. For each compound, a Boltzmann distribution was assumed to govern the thermodynamic equilibrium among allosteric modulator:CB1R complexes given rise by different poses of the considered

allosteric modulator. Therefore, K_B for a given compound can be calculated from ΔG^{allo}_{btnd} for its different poses ($\Delta G^{allo,i}_{btnd}$) as follows:

$$K_B = e^{-\Delta G^{allo}_{btnd}/RT}; \Delta G^{allo}_{btnd} = \sum_{i=1}^{N} \frac{e^{-\Delta G^{allo,i}_{total}/RT}}{\sum_{j=1}^{N} e^{-\Delta G^{allo,j}_{total}/RT}} \Delta G^{allo,i}_{btnd} \quad (5)$$

where $\Delta G^{allo,i}_{btnd}$ is ΔG^{allo}_{btnd} of the allosteric modulator pose i, $\Delta G^{allo,i}_{total}$ is the total free energy of allosteric modulator:CB1R complex derived from its pose i and indexes of summation i and j extends to all allosteric modulator poses. According to this thermodynamic equilibrium, only poses which induce binary complexes with the lowest total free energies will be significantly populated and, therefore, will contribute to K_B. It is important to take into account that K_B for racemic compounds, such as ZCZ011 and GAT211, is given by K_Bs associated with each of their enantiomers whose ponderations depend on relative stabilities of their arisen complexes and relative proportions both between enantiomers with each other and between enantiomers and receptor (see Supporting Information for more details). This fact can cause that response-concentration curves can undergo a change of behavior when the enantiomer concentration summation exceeds the receptor's one providing these enantiomers possess different allosteric modulator properties. Given that the receptor concentration is unknown for concentration-response curves, K_B for racemic compounds has been calculated by averaging enantiomers K_Bs.

At this point, one can envisage two approaches to determine the total free energy of allosteric modulator:CB1R complexes (ΔG^{allo}_{total}s): (1) Different CB1R conformations exhibit a similar stability, so that the complex total free energy (ΔG^{allo}_{total}) is given by ΔG^{allo}_{btnd}; (2) Different CB1R conformations does not exhibit a similar stability, so that the conformational free energy (ΔG^{allo}_{fold}) must be determined explicitly and $\Delta G^{allo}_{total} = \Delta G^{allo}_{btnd} + \Delta G^{allo}_{fold}$. It is worth noticing that the assumption made on first approach does not have to be fulfilled. However, this approach is interesting because it is simpler and less computational demanding by not having to calculate ΔG^{allo}_{fold}.

Table 1 contains ΔG^{allo}_{btnd}, ΔG^{allo}_{fold}, ΔG^{allo}_{total} and the relative populations according to approaches 1 and 2 of significantly populated allosteric modulator:CB1R complexes for compounds considered in our validation in their different poses (see Table S1 for all binary complexes), while Table 2 contains the experimental pK_Bs, ΔG^{allo}_{btnd}s associated with them ($\Delta G^{exp,allo}_{btnd}$s) and, furthermore, ΔG^{allo}_{btnd}s calculated according to approaches 1 and 2.

Table 1 indicates that for a given compound, there are allosteric modulator:CB1R complexes of similar ΔG^{allo}_{total} which can come from binding modes with different ΔG^{allo}_{btnd} (see compounds GAT228, GAT229, GAT1600, and GAT1601 in Tables 1 and S1). Thus, the assumption of the first approach is not fulfilled. Although, looking at Table 2 (columns 4 and 5), this fact affects only the value of ΔG^{allo}_{btnd} in a significant way when these complexes exhibit practically identical ΔG^{allo}_{total}, but considerably distinct ΔG^{allo}_{btnd} (see compounds GAT1600 and GAT1601 in Tables 1 and 2). Comparing $\Delta G^{exp,allo}_{btnd}$ with ΔG^{allo}_{btnd} obtained by approaches 1 and 2, it can be observed that both approaches achieve to reproduce experimental data. However, approach 2 fits better to these data and, additionally, allows to order considered allosteric modulators unambiguously by their affinities (see compounds GAT1600 and GAT1601 in Table 2), which is very useful from a pharmacological point of view. All this stands out the necessity of calculating ΔG^{allo}_{fold} for evaluating ΔG^{allo}_{total} correctly, since otherwise ΔG^{allo}_{btnd} can be underestimated or overestimated, being able to cause an erroneous affinity ordering of compounds. This wrong ordering can lead to erroneous conclusions in structure–activity relationship (SAR) studies, provoking a dead-end road for the rational design of active compounds employing both a computational approach and a computationally assisted experimental approach. That is why only approach 2 is going to be used in the following sections.

TABLE 1 | ΔG^{allo}_{btnd}, ΔG^{allo}_{fold}, ΔG^{allo}_{total} and the relative populations according to approaches 1 and 2 of significantly populated allosteric modulator:CB1R complexes for compounds considered in our validation in their different poses. A given binary complex has been considered as significantly populated if its relative population exceeds 0.1000. Racemic compounds have not been included since their values can be derived from their enantiomer values. All free energies have been expressed in kcal mol^{-1}.

Compound	Pose	ΔG^{allo}_{bind}	ΔG^{allo}_{fold}	ΔG^{allo}_{total}	Relative population approach 1	Relative population approach 2
ZCZ011(*S*)	2	−8.50	−613.25	−621.75	0.6391	8.2547×10^{-7}
	3	−8.14	−621.95	−630.09	0.3492	0.9999
ZCZ011(*R*)	2	−10.67	−615.05	−625.72	0.9285	1.0000
GAT228	2	−10.14	−616.63	−626.77	0.9652	0.9851
GAT229	2	−9.38	−611.77	−621.15	0.7575	0.9493
	3	−8.56	−608.93	−617.49	0.1912	0.0021
GAT1600	1	−9.38	−618.16	−627.54	0.0081	0.1904
	2	−12.24	−617.63	−629.87	0.9891	0.8095
GAT1601	2	−8.83	−623.10	−631.93	0.0037	0.2504
	3	−12.17	−620.42	−632.59	0.9954	0.7496

TABLE 2 | The experimental pK_Bs, ΔG^{allo}_{btnd}s associated with them ($\Delta G^{exp,allo}_{btnd}$s) and ΔG^{allo}_{btnd}s calculated according to approaches 1 and 2 for considered allosteric modulators in our validation. The experimental pK_Bs were collected from previous papers [7, 20, 21] $\Delta G^{exp,allo}_{btnd}$s have been calculated from pK_Bs employing Equation (4) and the corresponding errors have been obtained propagating pK_Bs standard deviations according to the uncertainty propagation law. All binding energies have been expressed in kcal mol^{-1}.

Compound	Experimental pK_B	$\Delta G^{exp,allo}_{bind}$	ΔG^{allo}_{bind} according to approach 1	ΔG^{allo}_{bind} according to approach 2
ZCZ011	6.11 ± 0.12	−8.39 ± 0.16	−8.74	−8.54
GAT221	6.76 ± 0.08	−9.28 ± 0.11	−9.43	−9.51
GAT228	7.40 ± 0.12	−10.16 ± 0.25	−10.06	−10.07
GAT229	6.91 ± 0.27	−9.48 ± 0.37	−9.13	−9.23
GAT1600	8.78 ± 0.33	−12.05 ± 0.45	−12.21	−11.70
GAT1601	8.29 ± 0.39	−11.38 ± 0.54	−12.15	−11.33

On the other hand, as introduced above in this section, it is worth noting that small discrepancies for racemic compounds may exist due to enantiomers competition when their concentration summation ($[B_S]+[B_R]$) is bigger than the receptor concentration ($[R]$). Regarding ZCZ011, though its R enantiomer in pose 2 ($\Delta G^{allo}_{total} = -625.72$ kcal/mol; $\Delta G^{allo}_{btnd} = -10.67$ kcal/mol) is shifted by its S enantiomer in pose 3 ($\Delta G^{allo}_{total} = -630.09$ kcal/mol; $\Delta G^{allo}_{btnd} = -8.14$ kcal/mol), there is no noticeable K_B mismatch because ZCZ011 racemic mixture exhibits a highly similar ΔG^{allo}_{btnd} in comparison with ZCZ011(S) in its pose 3 (−8.54 and −8.14 kcal/mol, respectively). In contrast, this is not the case for GAT221, however, its justification is somehow more complex since its K_B and α have been obtained together applying OMAM [55] and the global stability of binary and ternary complexes of GAT228 and GAT229 is the underlying cause of this mismatch. For this reason, explaining this justification in Section 3.2 has been considered more adequate. Additionally, it is important to note that GAT221 pK_B reported experimentally would not be compatible with the calculated pK_B from this magnitude for its enantiomers (−7.09 ± 0.21), assuming that there is no enantiomers competition. This fact highlights the point that response-concentration curves employed to obtain the GAT221 parameters underwent a change of behavior and the quality of the fit must not be excessively good since an amplifying effect due to the concurrent presence of GAT228 and GAT229 would be unlikely given the existence of a single CB1R positive allosteric site (see Section 3.5).

Finally, it is important to note that binding mode 2 and, to a lesser extent, 3 are the main poses that allosteric modulators adopt in their binary complexes with CB1R. This prevalence of binding mode 2 explains the importance in CB1R positive allosteric modulation of the aforementioned Ser199$^{3.35}$ and NO_2 group, while furthermore, the prevalence of binding mode 3 stands out the fact that placing of the 2-phenylindole ring at TM3 right side is feasible from an energetic point of view (see Section 3.3 for molecular details of such binding modes).

3.2 | Cooperativity Between Allosteric and Orthosteric Sites

Cooperativity is a phenomenon which consists in the affinity change of one binding site regarding a ligand upon binding to another one by this or other ligand. GPCRs are allosteric proteins and cooperativity between their orthosteric and G protein sites is well-known [59, 60]. Additionally, for some GPCRs, as is the case of CB1R, this cooperativity has also been observed between orthosteric and allosteric sites [7, 20–23, 29].

The cooperativity factor (α) is usually used to assess the aforementioned property, quantifying the degree of variation in dissociation constants due to ligands cobinding. As stated above, this factor is commonly determined by OMAM [55] as has been the case for our considered compounds, with the exception of ZCZ011 for which CP55940 EC_{50} ratio in presence and absence of ZCZ011 is available, an equivalent magnitude considering α definition and the law of microscopic reversibility of thermodynamics [55]. According to this model [55], the agonist (A) and allosteric modulator (B) bind to the receptor (R) in separated binding sites with dissociation constants K_A and K_B, respectively. The cobinding changes in a reciprocal way the affinity of binding, in other words, the dissociation constant, of each by a factor α, so that now, dissociation constants will be given by K_A/α and K_B/α. According to this definition, α can be calculated as follows:

$$\Delta\Delta G^{\alpha}_{btnd} = \Delta\Delta G^{allo}_{btnd} = \Delta\Delta G^{CP55940}_{btnd} = \Delta G^{allo(CP55940)}_{btnd} - \Delta G^{allo}_{btnd} = \Delta G^{CP55940(allo)}_{btnd} - \Delta G^{CP55940(allo)}_{btnd} = RT\ln(\alpha) \quad (6)$$

where $\Delta\Delta G^{\alpha}_{btnd}$ indicates the binding free energy change associated with α, $\Delta G^{allo(CP55940)}_{btnd}$ / $\Delta G^{CP55940(allo)}_{btnd}$ stands for the binding free energy of allosteric modulator/CP55940 in presence of its corresponding partner and $\Delta\Delta G^{allo}_{btnd}$ / $\Delta\Delta G^{CP55940}_{btnd}$ denotes the change of binding free energy for allosteric modulator/CP55940 due to the presence of its corresponding partner. Therefore, cooperativity factor α for allosteric modulator:CP55940:CB1R complexes can be calculated in two equivalent but distinct ways: (1) Calculating $\Delta G^{allo(CP55940)}_{btnd}$ to determine $\Delta\Delta G^{allo}_{btnd}$; (2) Calculating $\Delta G^{CP55940(allo)}_{btnd}$ and $\Delta G^{CP55940}_{btnd}$ (not calculated previously) to determine $\Delta\Delta G^{CP55940}_{btnd}$.

In order to check if our methodology is also effective in calculations of α, we calculated this quantity in these two equivalent manners for the previously considered compound.

Considering the results of the previous section, in order to calculate $\Delta G_{btnd}^{allo(CP55940)}$ and $\Delta G_{btnd}^{CP55940(allo)}$ correctly is necessary to assign populations to different allosteric modulator:CP55940:CB1R complexes according to their total free energy ($\Delta G_{total}^{allo/CP55940}$), which is now given by:

$$\Delta G_{total}^{allo/CP55940} = \Delta G_{btnd}^{allo(CP55940)} + \Delta G_{btnd}^{CP55940(allo)} + \Delta G_{fold}^{allo/CP55940} \quad (7)$$

where $\Delta G_{fold}^{allo/CP55940}$ is the conformational free energy of allosteric modulator:CP55940:CB1R complexes.

In Table 3 appears $\Delta G_{btnd}^{allo(CP55940)}$, $\Delta G_{btnd}^{CP55940(allo)}$, $\Delta G_{fold}^{allo/CP55940}$, $\Delta G_{total}^{allo/CP55940}$, and the associated relative populations of significantly populated allosteric modulator:CP55940:CB1R complexes for the compounds considered in their different poses (see Table S2 for all ternary complexes), while Table 4 provides the experimental αs, $\Delta\Delta G_{btnd}^{\alpha}$s associated with them ($\Delta\Delta G_{btnd}^{exp,\alpha}$s) and calculated $\Delta\Delta G_{btnd}^{allo(CP55940)}$s and $\Delta\Delta G_{btnd}^{CP55940(allo)}$s.

In Table 3, it can be seen that the most stable ZCZ011 ternary complexes and, therefore the most populated, derived from ZCZ011(*S*) in its pose 1 and ZCZ011(*R*) in its pose 3, being the resulting one of *S* enantiomer slightly more stable. This result agrees with the experimental binding mode of ZCZ011(*S*) observed in the cryo-EM structure of the ZCZ011(*S*):CP55940:CB1R:G_i complex (PDB code 7WV9 [29]; see Section 3.3 for a more detailed explanation). It is worth noting that, unlike binary CB1R complexes, binding modes 1 and 3 are the main poses that allosteric modulators adopt in allosteric modulator:CP55940:CB1R complexes, suggesting that Ser199$^{3.35}$ is crucial to trigger positive allosteric modulation activity in absence of the orthosteric agonist.

For its part, Table 4 stands out that our methodology is also capable of reproducing the experimental α of considered compounds by calculation of $\Delta\Delta G_{bind}^{allo(CP55940)}$ (this approach achieves a better agreement with the experimental data than the $\Delta\Delta G_{bind}^{CP55940(allo)}$'s one). Additionally, it can be noticed that $\Delta\Delta G_{btnd}^{allo(CP55940)}$ and $\Delta\Delta G_{bind}^{CP55940(allo)}$ do not match, which must not be like that according to the law of microscopic reversibility of thermodynamics [55], being the latter slightly biased towards more negative values. This result for $\Delta\Delta G_{btnd}^{CP55940(allo)}$ can be attributed to CP55940 movement restraint inside the orthosteric site since this could slightly bias CP55940 sampling, favoring in a greater extend CP55940 poses when an allosteric modulator is present. As stated in the methodology section, this restraint had to be applied with the aim of avoiding clashes between CB1R and CP55940 due to the great flexibility of CP55940. However, despite of $\Delta\Delta G_{btnd}^{CP55940(allo)}$ bias, it is important to note that $\Delta G_{btnd}^{CP55940(allo)}$ for different ternary complexes has enough accuracy to describe this contribution in $\Delta G_{total}^{allo/CP55940}$ since, in principle, no bias exists for CP55940 poses when an allosteric modulator is present.

On the other hand, $\Delta\Delta G_{btnd}^{allo(CP55940)}$ discrepancy for GAT221 can be attributed to the fact that response-concentration curves of CP55940 when $[B_S]+[B_R]>[R]$ reflect to a greater extend GAT228 properties than these for the racemic mixture of GAT228 and GAT229 since both binary and ternary GAT228 complexes are substantially more stable than GAT229 ones (see Tables 1 and 3). Apart from a goodness of fit not excessively good (which has not been reported in the corresponding paper), this change of the behavior in such curves gives rise to a α overestimation. This leads to an underestimation of K_B because of both parameters are interdependent, since GAT228 exhibits a greater affinity (smaller K_B) than GAT228 and GAT229 racemic mixture shifting these curves towards smaller agonist concentrations, which is associated with a bigger α. Overestimation of the experimentally reported GAT221 α is quite clear when this is compared with GAT221 α calculated from the experimental data of GAT228 and GAT229 (1.78 ± 0.65). It is worth noting that ZCZ011 racemic mixture does not provoke a α mismatch between experimental and calculated values given that a specific experiment has been carried out to determine this magnitude. This removes the interdependence with K_B, and there is not a significant change of behavior in $\Delta G_{bind}^{allo(CP55940)}$ because it is highly similar when both there is no competition between enantiomers (−9.34 kcal/mol) and this competition exists (−9.50 kcal/mol). This no significant change of $\Delta G_{btnd}^{allo(CP55940)}$ is due to the fact that relative populations of ZCZ011 ternary complexes do not change to a

TABLE 3 | $\Delta G_{btnd}^{allo(CP55940)}$, $\Delta G_{btnd}^{CP55940(allo)}$, $\Delta G_{fold}^{allo/CP55940}$, $\Delta G_{total}^{allo/CP55940}$ and the associated relative populations of significantly populated allosteric modulator:CP55940:CB1R complexes for considered compounds in their different poses. A given ternary complex has been considered as significantly populated if its relative population exceeds 0.1000. Racemic compounds have not been included since their values can be derived from their enantiomer values. All free energies have been expressed in kcal mol^{-1}.

Compound	Pose	$\Delta G_{bind}^{allo(CP55940)}$	$\Delta G_{bind}^{CP55940(allo)}$	$\Delta G_{fold}^{allo/CP55940}$	$\Delta G_{total}^{allo/CP55940}$	Relative population $\Delta G_{total}^{allo/CP55940}$
ZCZ011(*S*)	1	−10.01	−13.39	−614.29	−637.69	0.8046
	3	−7.56	−12.71	−616.50	−636.77	0.1735
ZCZ011(*R*)	3	−9.22	−11.62	−615.74	−637.57	1.0000
GAT228	1	−9.44	−12.28	−614.41	−637.75	0.4632
	3	−10.84	−14.38	−614.50	−637.72	0.4434
GAT229	1	−10.17	−14.71	−608.10	−632.98	0.1910
	2	−9.73	−12.90	−611.21	−633.84	0.8090
GAT1600	3	−12.85	−13.39	−615.29	−641.58	0.9969
GAT1601	1	−13.12	−13.28	−619.72	−647.11	1.000

TABLE 4 | The experimental αs, $\Delta\Delta G^{\alpha}_{bind}$s associated with them ($\Delta\Delta G^{exp,\alpha}_{bind}$s) and calculated $\Delta\Delta G^{allo(CP55940)}_{bind}$s and $\Delta\Delta G^{CP55940(allo)}_{bind}$s taking into account relative populations of allosteric modulator:CP55940:CB1R complexes according to $\Delta G^{allo/CP55940}_{total}$s. The experimental αs were collected from previous papers [7, 21, 29] $\Delta\Delta G^{exp,\alpha}_{bind}$s have been calculated from αs employing Equation (6) and the corresponding errors have been obtained propagating αs standard deviations according to the uncertainty propagation law. All differences in free energies are expressed in kcal mol^{-1}.

Compound	Experimental α	Experimental $\Delta\Delta G^{exp,\alpha}_{bind}$	$\Delta\Delta G^{allo(CP55940)}_{bind}$	$\Delta\Delta G^{CP55940(allo)}_{bind}$
ZCZ011	4.99 ± 1.24	−0.96 ± 0.15	−0.80	−1.16
GAT221	21 ± 2	−1.81 ± 0.16	−0.94	−1.58
GAT228	1.08 ± 0.34	−0.05 ± 0.35	−0.09	−1.59
GAT229	2.25 ± 1.26	−0.48 ± 0.54	−0.58	−1.57
GAT1600	38.0 ± 6.1	−2.17 ± 0.65	−1.14	−1.76
GAT1601	51.0 ± 2.4	−2.34 ± 0.76	−1.79	−2.59

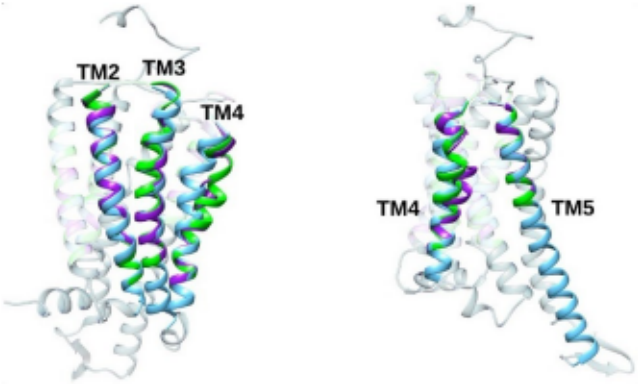

FIGURE 5 | The overlap of the CB1R active reference structure (in blue), the CB1R reference structure associated with pure allosteric agonist activity (in green; GAT228 binary complex in its pose 2) and the CB1R reference structure associated with pure PAM activity (in purple; GAT229 binary complex in its pose 2). It can be noticed that regarding allosteric agonist activity, TM4 distances from TM2, TM3 and TM5, while in the case of PAM activity, TM4 distances from TM2 and TM3, but approaches to TM5. Additionally, the loss of α helix content is more pronounced for PAM activity.

large extent whether there is competition between enantiomers or not (Without competition: 0.4023 for ZCZ011(*S*) in its pose 1, 0.0868 for ZCZ011(*S*) in its pose 3 and 0.5000 for ZCZ011(*R*) in its pose 3; With competition: 0.4906 for ZCZ011(*S*) in its pose 1, 0.1058 for ZCZ011(*S*) in its pose 3 and 0.4036 for ZCZ011(*R*) in its pose 3). It can be concluded from the above information that our methodology is effective to calculate α by $\Delta\Delta G^{allo(CP55940)}_{bind}$ and to determine relative populations of allosteric modulator:CP55940:CB1R complexes.

3.3 | Structural Insights Regarding Allosteric Modulator:CB1R and Allosteric Modulator:CP55940:CB1R Complexes

CB1R allosteric modulators belonging to the 2-phenylindole structural class may behave as allosteric agonists, positive allosteric modulators (PAMs) and agonist and positive allosteric modulators (ago-PAMs), though the exhibited degree of allosteric agonist and PAM activity is highly variable. Nevertheless, GAT228 and GAT229 act as a CB1R pure allosteric agonist and PAM, respectively [21]. For this reason, these two compounds are used as probe compounds to study these two well-differentiated activities. In order to explore the structural basis and the origin at molecular level of the two aforementioned activities, the structures of GAT228 and GAT229 complexes resulting from their significantly populated binding modes were compared with each other and with the CB1R active reference structure, which is represented by the experimental structure of the ZCZ011(*S*):CP55940:CB1R:G_i complex (PDB code 7WV9) [29]. For binary complexes, it is important to note that it is only necessary to consider GAT228 and GAT229 complexes derived from their binding modes 2 since they are the only ones significantly populated for these two compounds (see compounds GAT228 and GAT229 in Table 1). Considering the CB1R active reference structure, although small structural differences can be observed in TM2 and TM3, both the allosteric agonist activity and the PAM activity seem to lie in an outward movement of the TM4 extracellular end. For both activities, this displacement occurs in a similar extension, however, its direction changes. In relation with allosteric agonist activity, the aforementioned outward movement distances TM4 extracellular end from TM2, TM3 and TM5, however, in the case of PAM activity, this movement also moves away TM4 extracellular end from TM2 and TM3, but approaches it to TM5. In addition, this displacement entails a decrease of α helix content of TM4 extracellular end, which is more pronounced for PAM activity (see Figure 5 and Figure S9). Taking into account the structural features that have just been exposed, it is worth noticing that the other considered allosteric modulators in our method validation give rise to significantly populated allosteric modulator:CB1R complexes whose structures are compatible with their experimental activity reported (see Supporting Information for more details, Figure S9 for an overlap of binary and ternary complexes and Figures S10–S12 for a frontal view with increased zoom).

At this point, it is worth noticing that the stereochemistry of the two possible chiral carbon atoms for the type of considered compounds is relevant from the point of view of exhibited activity,

especially for chiral center 1. When only this chiral center is present, taking into account the only resolved enantiomers (GAT228 and GAT229) [21], the spatial arrangement of substituents around this center as in GAT228 and GAT229 is equivalent to allosteric agonist and PAM activities, respectively. Additionally, our calculations suggest that this relationship between positive allosteric modulation activities and spatial arrangement of substituents around the chiral carbon atom 1 is general for the 2-phenylindole allosteric modulators that contain only this chiral center (see Section 3.6 for 6s). Nevertheless, the insertion of a second chiral center allows to incorporate allosteric agonist activity to previous enantiomers with only PAM activity. The spatial arrangement of substituents around the chiral carbon atom 1 as in GAT228 and GAT229 is now equivalent to allosteric agonist and ago-PAM activities, respectively, regardless of the absolute configuration of the second chiral center [7].

Although the considered compounds have different structures, their significantly populated binding modes vary and they exhibit different allosteric activities, some interactions may be conserved between allosteric modulators which exhibit the same type of activity. Considering GAT228 and GAT229 as probe compounds for CB1R pure allosteric agonist and PAM activities, respectively, the origin at molecular level of PAM activity seems to lie in a deeper penetration of the ring that is found at TM3 right side into the narrow cavity that exists between this helix and TM4. In general, this deeper compound penetration provokes a favorable intense hydrophobic interaction with TM4, normally, by means of Leu252$^{4.61}$, other weaker with TM3 by Leu190$^{3.26}$ and a better π-π stacking interaction with Phe191$^{3.27}$ (see Figure 6). On the other hand, allosteric agonist activity seems to be linked to a more favorable intense interaction with residues Leu165$^{2.52}$, Ile169$^{2.56}$, Ala198$^{3.34}$, and Ser199$^{3.35}$ whose character depends on which allosteric modulator group is interacting with them (polar character when the NO_2 group is pointing to this tetrad, while hydrophobic character when a ring is interacting; see Figure 6). All remaining considered compounds follow the above interaction-activity relationships (see Figures S10–S12 for representation of their binding modes and Supporting Information for justification of GAT1600 ago-PAM activity from its interactions).

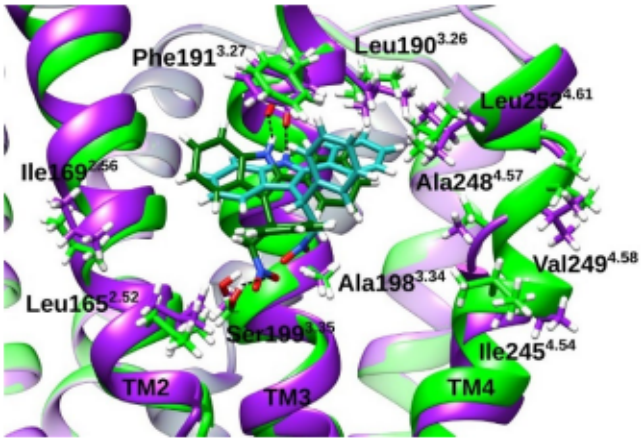

FIGURE 6 | Binding modes of CB1R reference structures regarding allosteric agonist (in green with GAT228 in its pose 2 depicted in dark green) and PAM (in purple with GAT229 in its pose 2 depicted in cyan) activities for binary complexes. The origin at molecular level of PAM activity seems to lie in a deeper penetration into the narrow cavity that exists between TM3 and TM4 by allosteric modulator ring placed in this region, while allosteric agonist activity seems to be linked to a more intense favorable interaction with the tetrad Leu165$^{2.52}$, Ile169$^{2.56}$, Ala198$^{3.34}$, and Ser199$^{3.35}$. Hydrogen bonds have been indicated by dashed lines.

Regarding the allosteric modulator:CP55940:CB1R complexes, proceeding similarly to binary complexes, it can be noted that these ternary complexes are similar between each other, though their exhibited degree of similarity is reduced in comparison with binary complexes (see Figure S9), suggesting that the conformational space associated with ternary complexes is wider. From a structural point of view, structural differences of ternary complexes accumulate mainly in extracellular ends of TM2, TM3, TM4, TM6, and TM7, being the most prominent the TM7 ones (see Figure 7). Unlike binary complexes, slight outwards displacement of TM4 extracellular end is observed for allosteric modulator:CP55940:CB1R complexes, however, these ternary complexes show a compound TM7 movement in which its extracellular end moves outwards whereas its intracellular end moves slightly inwards. Additionally, this TM7 movement is partially accompanied by TM6 extracellular end and α8 helix, where the slight inwards movement of TM7 intracellular end along with the accompanying α8 helix tightens the G protein binding site, enhancing the interaction between this cavity and G protein α5 helix, which is the α helix that is introduced into GPCRs. It is worth noticing that this compound TM7 movement depends on allosteric modulators and also takes place in the CP55940:CB1R complex, being more pronounced regarding the TM7 extracellular end. Therefore, the orthosteric agonist causes given movement while allosteric modulators regulate its range by means of structural differences in TM2, TM3, and TM4 induced by their distinct binding modes and interactions, which are subsequently propagated to TM6 and TM7.

Regarding the allosteric modulators binding modes and interactions, despite the fact that significantly populated binding modes for considered allosteric modulators have changed from binary to ternary complexes, being binding mode 2 along with a non-neglectable contribution of binding mode 3 the most relevant for the former complexes, while binding modes 1 and 3 are the most relevant for the latter ones, interactions trends are preserved to a considerable extent. However, differences between allosteric modulators with distinct activity are now less pronounced, which is due to the subtle displacement of TM4 extracellular end. Similarly to binary complexes, allosteric modulators associated with PAM activity exhibit a slightly deeper penetration of the ring placed between TM3 and TM4 into the existing cavity between both transmembrane α helixes. This observed more superficial penetration in ternary complexes is due to the aforementioned slight TM4 displacement, which shrinks given cavity substantially. Again, this penetration is related to a better π-π stacking interaction with Phe191$^{3.27}$. In contrast, hydrophobic interactions between TM4 and the ring placed between this transmembrane helix and TM3 equalizes, taking place now through Ile245$^{4.54}$ and Val249$^{4.58}$. In relation to the allosteric modulators associated with allosteric agonist activity, they show a favorable stronger interaction with residues Leu165$^{2.52}$, Ile169$^{2.56}$, Ala198$^{3.34}$, Ser199$^{3.35}$ (like in binary complexes) and Trp241$^{4.50}$. The subtle displacement of TM4 brings this tryptophan closer to the previous tetrad, allowing its

interaction with allosteric modulators. Obviously, Trp241$^{4.50}$ establishes non-directional hydrophobic interaction, however, also establishes a π-π stacking interaction with allosteric modulators which show PAM activity (see Figure S13).

Interestingly, the CB1R reference active structure (the ZCZ011(*S*):CP55940:CB1:G_i complex) does not exhibit TM7 extracellular end movement, though this does show the slight inwards movement of TM7 intracellular end and α8 helix (blue complex in Figure 7). This apparent discrepancy between structures could be explained by an allosteric effect of G protein, which is present in this experimental structure but absent in our ternary complexes. On the other hand, the CP55940:CB1R complex has also been reported experimentally (PDB code 7V3Z) [61], though a fusion protein and cholesterol have been employed to stabilize it. In this case, neither the compound TM7 movement nor its associated movements of TM6 extracellular end and α8 helix are exhibited (orange complex in see Figure 7). This time, this structural discrepancy can arise from two different factors: (1) The introduction of a fusion protein, which could stabilize conformations of the CP55940:CB1R complex in which the aforementioned movement does not occur; (2) Cholesterol acting as a CB1R NAM [62], inducing CB1R states not fully active (91% of degree activity) [61]. Regarding other CB1R active forms derived from other agonists [63–66], the compound TM7 movement along with its associated displacements are not observed either. This fact can be justified similarly to CP55940, further calculations are necessary to prove that this conformational change is a common feature for the orthosteric agonists since may these displacements be orthosteric-ligand dependent.

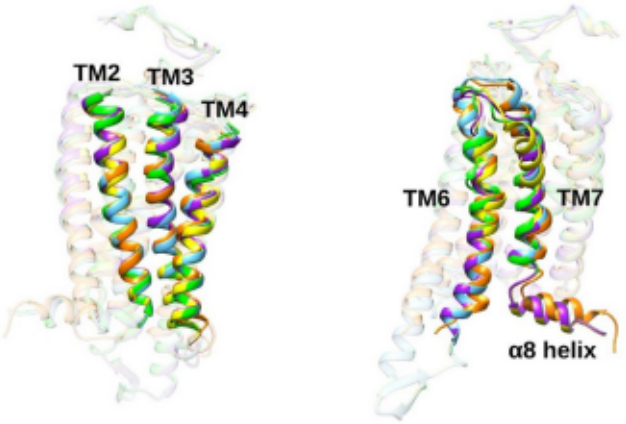

FIGURE 7 | The overlap of the CB1R active reference structure (in blue), the CB1R reference structures for allosteric agonist (in green and derived from pose 1 of GAT228) and PAM (in purple and derived from pose 2 of GAT229) activities, the CP55940:CB1R complex resulting from our calculations (in yellow) and the experimental structure of this complex (in orange; PDB code 7V3Z) [61]. Although there is another significantly populated ternary complex for GAT228, only the main one has been depicted since these are highly similar with each other from a structural point of view. It can be appreciated that structural differences in TM2, TM3, and TM4 induced by allosteric modulators are subsequently propagated to TM6 and TM7, regulating the range of motion of TM7 extracellular end (compare complexes in blue, green, purple, and yellow).

3.4 | Understanding the ZCZ011 Low Resolution

As previously said, the cryo-EM structure of the ZCZ011(*S*):CP55940:CB1R:G_i complex (PDB code 7WV9) [29] is available, whose resolution (2.7 Å) is comparable to the X-ray crystallography ones. However, based on ZCZ011 2Fo-Fc map, this structure shows a quite blurred ZCZ011(*S*) binding mode since it is not clear in which TM3 side lies its 2-phenylindole ring, the corresponding protuberance of the 2-phenylindole CH_3 substituent is very faint, and there is a clear protuberance around the chiral carbon atom 1 that is not compatible with the *S* enantiomer for the proposed binding mode. Additionally, the corresponding region to the 2-phenylindole ring along with its indole ring phenyl substituent is quite smooth, which must not be the case since both rings do not have the same size and a cleft corresponding to the C-C bond joining them must be observed.

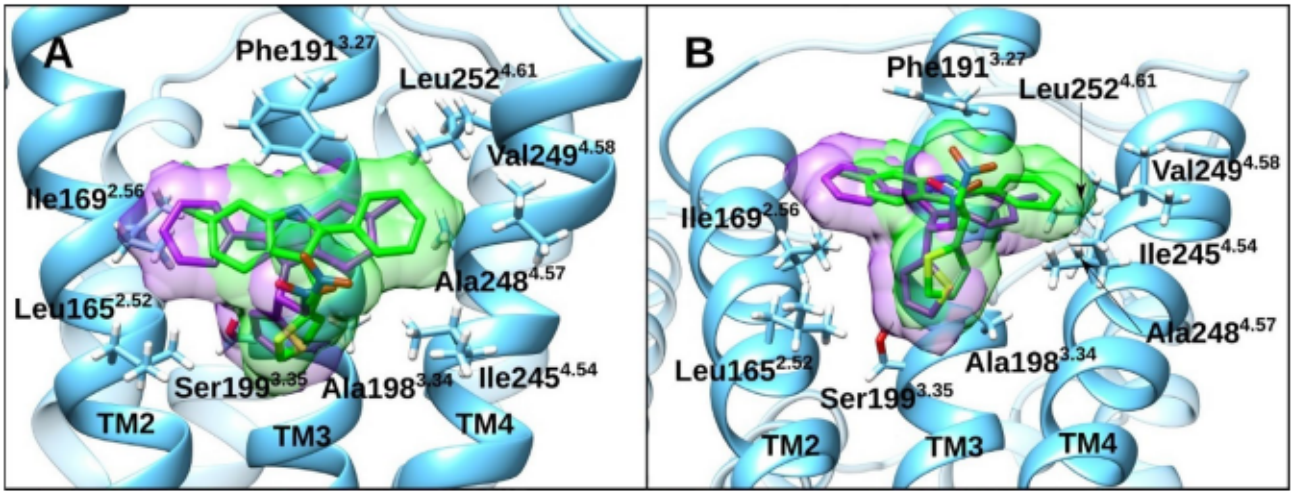

FIGURE 8 | The overlap of ZCZ011(*S*) in its binding mode 1 (in green) and ZCZ011(*R*) in its binding mode 3 (in purple) from a vertical and frontal perspectives regarding ZCZ011 (panels A and B, respectively). Contour surfaces for these ZCZ011 poses have also been depicted in the same color. For the sake of clarity, ZCZ011 hydrogen atoms have not been represented and TM2, TM3, and TM4 have been highlighted.

In contrast, spatial arrangement and orientation of its NO_2 group and thiophene ring are well resolved.

In order to understand this fact, it is important to note that given cryo-EM structure was obtained employing a ZCZ011 racemic mixture whose concentration was sufficiently large, in comparison with the receptor one, to shift the equilibrium of ternary complexes towards the most stable ones. In other words, both ZCZ011 enantiomers in their distinct binding modes have the opportunity to compete for the allosteric site. According to our results for ZCZ011 ternary complexes, this apparent poor resolution for ZCZ011(*S*) is because of particles, from which the aforementioned cryo-EM structure has been obtained, contain ternary complexes that derived from both ZCZ011(*S*) in its binding mode 1 and ZCZ011(*R*) in its binding mode 3 since these exhibit very similar stabilities (−637.69 and −637.57 kcal/mol, respectively). Considering relative stabilities of these complexes, the former one would be slightly more prevalent, as observed experimentally since such ZCZ011(*S*) binding mode gives rise to a better fit of 2Fo-Fc map (for this reason this is reported in the cryo-EM structure). The faint protuberance of the 2-phenylindole CH_3 substituent can be distinguished at TM3 left side, indicating a higher weight of given ZCZ011(*S*) pose. From the overlap of such ZCZ011 binding modes (see Figure 8), the experimental lack of sharpness of ZCZ011(*S*) binding mode can be explained. The faint protuberance for the 2-phenylindole CH_3 substituent is due to the fact that the 2-phenylindole ring along with its indole ring phenyl substituent do not completely match in these ZCZ011 binding modes, placing the indole ring phenyl substituent of one enantiomer over the 2-phenylindole CH_3 substituent of the other. Additionally, this arrangement of rings for both enantiomers justifies the smoothness and the lack of a cleft of ZCZ011 2Fo-Fc map in the region containing these two rings. On the other hand, the presence of ZCZ011(*R*) is what gives rise to the protuberance around the chiral carbon atom 1 that is incompatible with the *S* enantiomer.

3.5 | Putative Alternative Positive Allosteric Binding Sites on CB1R

To progress in the development of allosteric modulators, it is crucial to be able to distinguish the actual allosteric binding site/s from all putative ones, since once determined this/these, the molecular architecture of allosteric site/s can be understood, allowing the refinement of allosteric modulators known so far or, even, the rational design of new ones. As it could be established for CB1R, our methodology to determine stability of binary and ternary complexes turns out to be effective for this purpose as well.

Prior to the cryo-EM structure of the ZCZ011(*S*):CP55940:CB1:G_i complex (PDB code 7WV9) [29] was reported and mutagenesis studies to determine the CB1R positive allosteric modulator binding sites were carried out [31]. Several putative CB1R positive allosteric binding sites were postulated for the 2-phenylindole compounds, being two that differ from the real one the most relevant (see Figure 9): (1) For allosteric agonist activity (using GAT228 as a probe and with the orthosteric site empty), it has been postulated a binding site formed by intracellular ends of

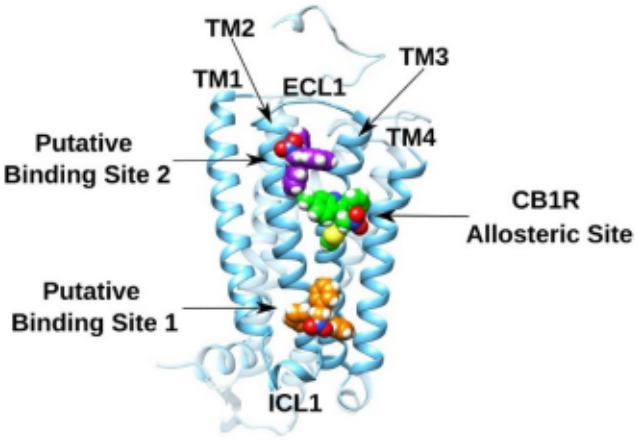

FIGURE 9 | Schematic representation of CB1R putative binding sites along with its actual one through a CPK depiction of the corresponding allosteric modulators bound to such binding sites (ZCZ011(*S*) bound to the CB1R allosteric sites in green; GAT228 bound to the putative binding site 1 in orange; GAT229 bound to the putative binding site 2 in purple).

TM1, TM2 and TM4 along with the intracellular loop (ICL) 1 (putative binding site 1); (2) For PAM activity (using GAT229 as a probe and with CP55940 occupying the orthosteric site), it has been postulated a binding site formed by extracellular ends of TM2 and TM3 along with the extracellular loop (ECL) 1 (putative binding site 2).

These relevant postulated sites were identified by using the Force-Biased Metropolis Monte Carlo simulated annealing program (MMC) [67, 68] over a CB1 receptor model assembled from two CB1R active states with cholesterol as a stabilizing agent (PDB codes 5XRA and 6N4B) [64, 65]. It is worth noticing that the MMC method divides the molecule of interest (GAT228 and GAT229 in this case) into smaller fragments and finds sites of the system where these fragments persist, indicating high affinity between them. Therefore, the MMC method identifies receptor regions with high affinity for the considered fragments, in other words, does not determine neither the binding mode nor the binding site for the molecule of interest explicitly. Additionally, taking into account that cholesterol acts a CB1R NAM [62], the employed CB1 receptor model could represent a CB1R state not fully active. All this would explain the mismatch between putative and real binding sites, pointing out the necessity of having a reliable method to screen allosteric binding sites.

Apart from GAT228 and GAT229 putative binding sites, their binding modes have also been postulated in these sites. It is important to note that allosteric modulators binding modes are affected by CB1R tertiary structure, which does not match totally between the employed CB1 receptor model for coming across such postulated binding sites/modes and the considered CB1R structure in the present work. Although the corresponding region for the putative binding site 1 is quite similar between both structures, the corresponding one for the putative binding site 2 exhibits significant differences. In particular, TM2 has undergone a downward rotation along the helical axis, while TM3 has experienced an upward rearrangement, reinforcing the idea that the employed CB1 receptor model does not represent

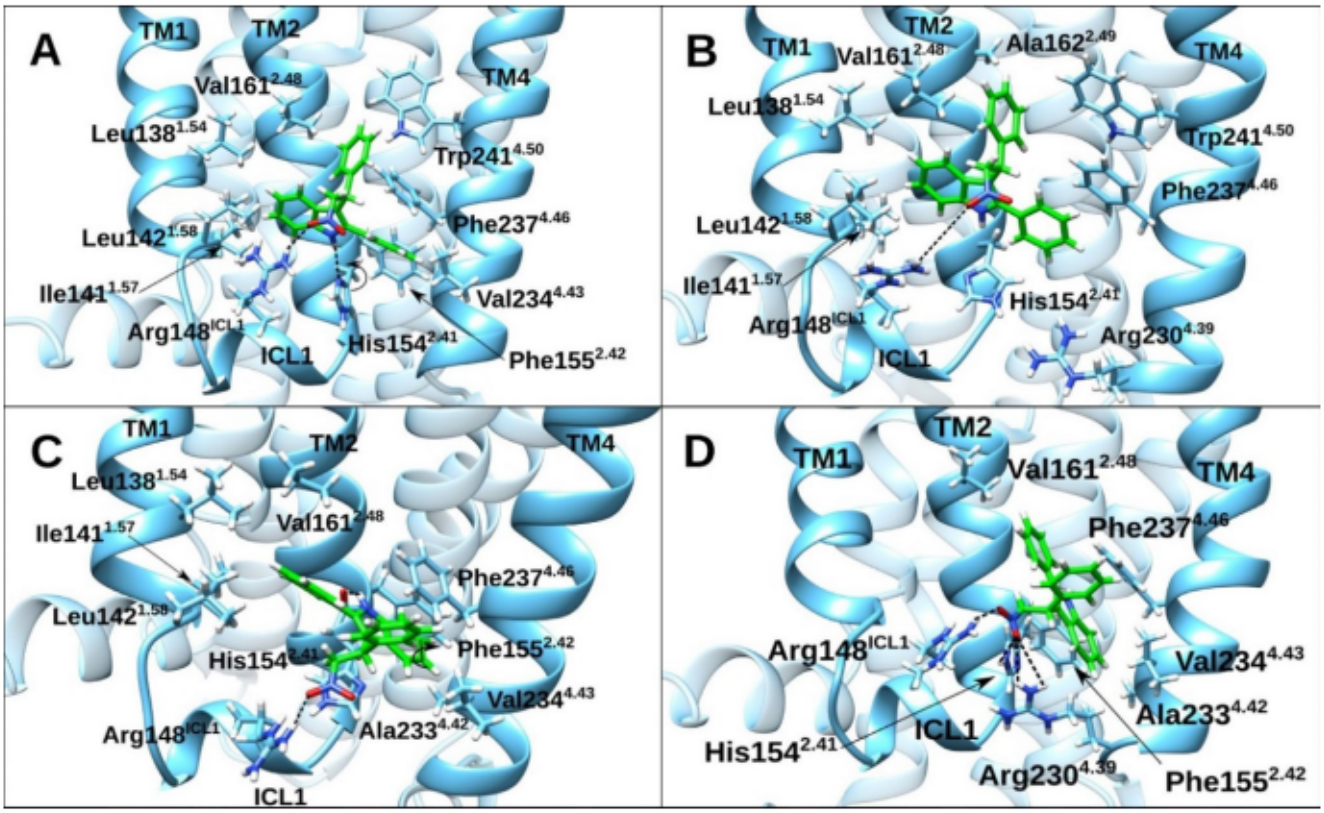

FIGURE 10 | Depiction of GAT228 postulated binding mode (Panel A) along with its best docking solution (Panel C) inside the putative binding site 1 where the considered compound (in green) and the interacting residues have been indicated. After relaxation and equilibration MD simulations, the postulated binding mode remains unaltered practically and the most noticeable change is a counterclockwise 90° rotation of $His154^{2.41}$ along its CB-CG bond (compare Panels A and B). In contrast, the docking binding mode changes significantly after given simulations (Panel D) since GAT228 undergoes a rotation of ~30° along the axis defined by the two 2-phenylindole carbon atoms closest to TM4. Hydrogen bonds have been indicated by dash lines, while rotations along with their directions have been depicted by bent arrows. For the sake of clarity, the rotation axis of GAT228 has been indicated.

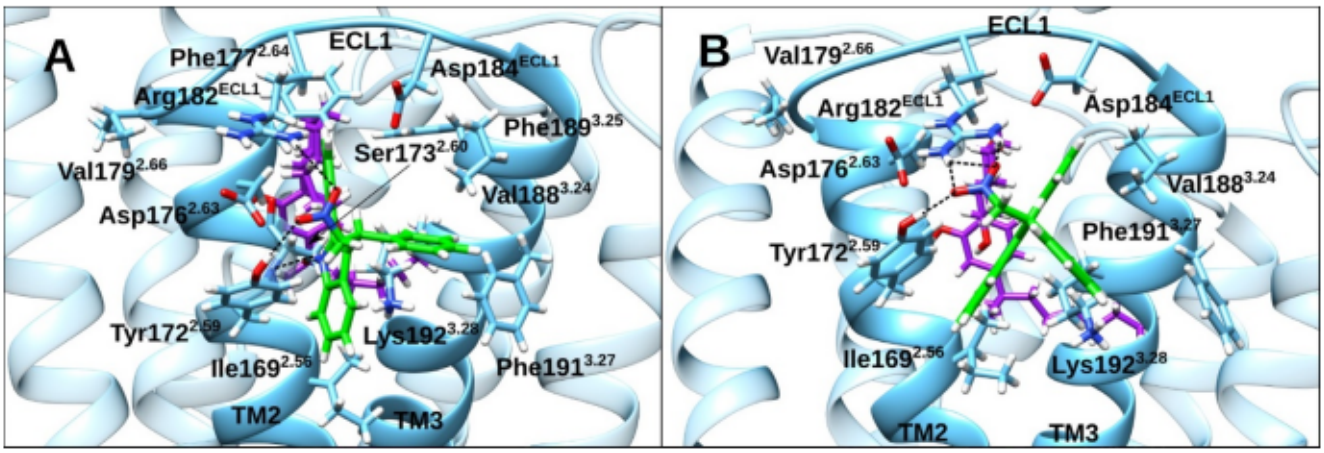

FIGURE 11 | Depiction of GAT229 docking binding mode inside the putative binding site 2 (Panel A). Both the considered compound (in green) and the interacting residues have been indicated. In this case, the orthosteric site is occupied by CP55940 (in purple). Due to the narrowness of this binding site, GAT229 is partially ejected from the cavity during relaxation and equilibration MD simulations, changing its interactions totally (panel B). Hydrogen bonds have been indicated by dash lines.

a fully-active CB1R state since these TM2 and TM3 reorganizations are inherent to CB1R active states [29]. By means of docking calculations, GAT228 postulated pose inside its putative binding site (see panel A of Figure 10) could be reproduced after modifying $Arg148^{ICL1}$ and $His154^{2.41}$ rotamers. However, this was impossible for GAT229 due to the aforementioned

structural differences between models. For this reason, both the postulated binding mode (see panel A of Figure 10) and the best docking solution (see panel C of Figure 10) have been considered for GAT228, while only the latter (see panel A of Figure 11) has been considered for GAT229. The main interactions of these binding modes change, to a greater or lesser extent, upon relaxation and equilibration MD simulations. For this reason, for the sake of brevity, only such interactions for the corresponding systems totally relaxed and equilibrated are going to be described (see Supporting Information for main interactions in initial structures).

Regarding GAT228 binding modes inside the putative binding site 1, it can be seen that the main differences between postulated and docking binding modes lie in the position of the 2-phenylindole ring inside this cavity (see panels A and C of Figure 10, respectively), affecting the rest of its substituents. For the postulated pose, this ring is located slightly on the left side of TM2 intracellular end, allowing its hydrogen bond with $His154^{2.41}$ side chain, whereas for the docking pose, the 2-phenylindole ring is located on the right of TM2 intracellular end, forcing its hydrogen bond with the oxygen backbone atom of $His154^{2.41}$. Although in both cases the NO_2 group points outwards the cavity, forming a hydrogen bond with $Arg148^{ICL1}$, due to a distinct spatial arrangement of the 2-phenylindole ring, the phenyl ring at β position of the NO_2 group can only be introduced in the hydrophobic region between TM2 and TM4 in the postulated pose. Regarding the indole ring phenyl substituent, it is placed in a hydrophobic region in both poses, however, the identity of its neighboring residues and their interaction with this ring are totally different since the given ring is placed in a distinct region in each binding mode. This ring point towards TM4 intracellular end for the postulated pose, while it points towards TM1 in ascending direction for the docking pose.

Obviously, the principal GAT228 interactions inside its putative binding site change depending on the considered pose since its spatial arrangement is totally different in each case. In relation to its postulated pose, this is quite stable since after relaxation and equilibration MD simulations, most of its interactions are preserved to a large extent and GAT228 arrangement inside its putative binding site hardly varies. This can be seen by comparing panels A and B of Figure 10. The most noticeable change is that the hydrogen bond between the 2-phenylindole NH group and $His154^{2.41}$ side chain is broken due to a counterclockwise 90° rotation of $His154^{2.41}$ ring, giving rise to a NH-π interaction between these two groups. The principal GAT228 interactions exhibited once the system is relaxed and equilibrated are the following ones: (1) Hydrogen bond between the NO_2 group and the $Arg148^{ICL1}$ side chain (d(GAT228-O1-HH21-$Arg148^{ICL1}$) = 4.18 Å); (2) Hydrophobic interactions between the 2-phenylindole six-membered ring and residues $Leu138^{1.54}$, $Ile141^{1.57}$, $Leu142^{1.58}$, and $Val161^{2.48}$; (3) Hydrophobic interactions between the phenyl ring at β position of the NO_2 group and residues $Val161^{2.48}$, $Ala162^{2.49}$, $Phe237^{4.46}$, and $Trp241^{4.50}$; (4) Cation-π interaction between the indole ring phenyl substituent and $Arg230^{4.39}$ (d(indole ring phenyl substituent centroid-CZ-$Arg230^{4.39}$) = 7.7 Å); (5) NH-π interaction between the 2-phenylindole NH group and $His154^{2.41}$ side chain (centroid-centroid distance = 5.9 Å); (6) π-π parallel stacking interaction between the indole ring phenyl substituent and $Phe237^{4.46}$ (centroid-centroid distance 4.4 Å). It is worth noting that the GAT228 postulated binding mode [30] has been reproduced satisfactorily as the vast majority of its interactions can also be observed in our binding mode, excluding the aromatic interactions that involve $Trp241^{4.50}$.

Regarding the docking pose, it has turned out to be much less stable since during relaxation and equilibration MD simulations GAT228 undergoes a reorganization. In particular, a clockwise rotation of ~30° around the axis defined by the two 2-phenylindole carbon atoms closest to TM4, modifying notably the initial GAT228 arrangement and interactions (compare Panels C and D of Figure 10). The most significant changes are that hydrophobic and aromatic interactions of the indole ring phenyl substituent and the phenyl ring at β position of the NO_2 group are considerably depleted and initial interactions with $His154^{2.41}$ are broken. In this case, the main GAT228 interactions after relaxation and equilibration MD simulations are: (1) Hydrogen bond between the NO_2 group and the $Arg148^{ICL1}$ side chain (d(GAT228-O1-HH21-$Arg148^{ICL1}$) = 1.89 Å); (2) Doble hydrogen bond between the NO_2 group and the $Arg230^{4.30}$ side chain (d(GAT228-O2-HH11-$Arg230^{4.30}$) = 2.63 Å; d(GAT228-O2-HH12-$Arg230^{4.30}$) = 3.21 Å); (3) Hydrogen bond between the NO_2 group and the $His154^{2.41}$ side chain (d(GAT228-O2-HE2-$His154^{2.41}$) = 3.28 Å); (4) Hydrophobic interactions between the indole ring phenyl substituent and $Val161^{2.48}$; (5) Hydrophobic interactions between the 2-phenylindole six-membered ring and residues $Phe155^{2.42}$, $Ala233^{4.42}$, $Val234^{4.43}$, and $Phe237^{4.46}$; (6) Hydrophobic interactions between the phenyl ring at β position of the NO_2 group and residues $Val234^{4.43}$ and $Phe237^{4.46}$; (7) π-π stacking interaction between the 2-phenylindole ring and $Phe237^{4.46}$ (d(centroid of the 2-phenylindole six-membered ring-$Phe237^{4.46}$ centroid) = 5.0 Å; d(centroid of the 2-phenylindole five-membered ring-$Phe237^{4.46}$ centroid) = 4.3 Å).

Regarding the docking in the GAT229 binding mode inside the putative binding site 2, due to the narrowness of such cavity, GAT229 is partially ejected from the cavity during relaxation and equilibration in MD simulations placing it in a quite superficial way. This entails a total reorganization of its interactions. The main GAT229 interactions after relaxation and equilibration MD simulations are: (1) Triple hydrogen bond between the NO_2 group and the $Arg182^{ECL1}$ side chain (d(GAT229-O1-HH11-$Arg182^{ECL1}$) = 1.96 Å; d(GAT229-O1-HH21-$Arg182^{ECL1}$) = 2.30 Å; d(GAT229-O2-HH21-$Arg182^{ECL1}$) = 1.95 Å); (2) Hydrogen bond between the NO_2 group and the $Tyr172^{2.59}$ side chain (d(GAT229-O2-HOH-$Tyr172^{2.59}$) = 1.95 Å); (3) Hydrophobic interactions between the 2-phenylindole ring and residues $Ile169^{2.56}$ and $Tyr172^{2.59}$; (4) Hydrophobic interactions between the indole ring phenyl substituent and $Val188^{3.24}$; (5) Cation-π interaction between the 2-phenylindole ring and $Lys192^{3.28}$ (d(centroid of the 2-phenylindole six-membered ring-NZ-$Lys192^{3.28}$) = 5.3 Å; d(centroid of the 2-phenylindole five-membered ring-NZ-$Lys192^{3.28}$) = 5.3 Å); (6) π-π parallel stacking interaction between the 2-phenylindole six-membered ring and $Tyr172^{2.59}$ (centroid-centroid distance 4.6 Å); (7) Cation-π interaction between the indole ring phenyl substituent and $Arg182^{ECL1}$ (d(centroid of the indole ring phenyl substituent-CZ-$Arg182^{ECL1}$) = 4.8 Å).

TABLE 5 | Total free energy ($\Delta G_{total}^{allo/allo(CP55940)}$) with its different contributions for binary and ternary complexes derived from GAT228 and GAT229 binding to their respective putative binding sites for their considered binding modes. A dual superscript has been used because allosteric modulator can be in absence (GAT228) or presence (GAT229) of CP55940. All free energies have been expressed in kcal mol^{-1}.

Compound	Pose	$\Delta G_{bind}^{allo/allo(CP55940)}$	$\Delta G_{bind}^{CP55940(allo)}$	$\Delta G_{fold}^{allo/allo(CP55940)}$	$\Delta G_{total}^{allo/allo(CP55940)}$
GAT228	Docking	−7.62	—	−606.53	−614.15
	Postulated	−11.50	—	−607.94	−619.44
GAT229	Docking	−3.41	−12.60	−593.75	−609.77

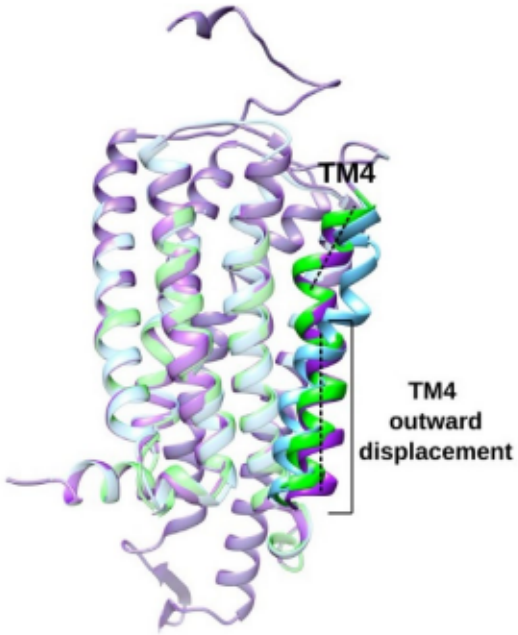

FIGURE 12 | Overlap of GAT228:CB1R complexes derived from binding on the allosteric and postulated binding sites. Most populated GAT228:CB1R complex derived from binding on the allosteric site has been depicted in blue, while these complexes derived from binding on the putative binding site 1 with GAT228 adopting its postulated binding mode and its best docking pose have been depicted in purple and green, respectively. The two axes of the broken TM4 helix have been indicated in dash lines.

As it has been illustrated in Sections 3.1 and 3.2, to assess a given allosteric site, total free energy of the resulting complex derived from binding in such site must be calculated. Table 5 contains this total free energy along with its contributions for complexes derived from GAT228 and GAT229 binding to their respective putative binding sites for their considered binding modes. Although both putative binding sites are considerably less hydrophobic than the real one, the same scaling factor for the non-electrostatic contribution of the binding free energy was employed in order to obtain comparable values, despite the fact that the most rigorous treatment would be not carry out such scaling. In relation to GAT228, both considered binding modes in its putative binding site show a comparable ΔG_{bind}^{allo} to the obtained one for this compound in the allosteric site (compare GAT228 pose 2 in Table 1 with GAT228 in Table 5), so that GAT228 exhibits a similar stability both in the allosteric site and in its putative binding site. In contrast, the stability of induced conformations upon GAT228 binding inside such sites is totally different, being CB1R conformations derived from binding in the allosteric site much more stable. This larger conformational stability is the determining factor for binding inside the experimentally reported allosteric site. From a structural point of view, this observed conformational instability is due to an outwards displacement of TM4 intracellular end (see Figure 12), which is necessary to be able to settle GAT228 in its putative binding site. This displacement causes TM4 to be a "broken" α helix since it exhibits two well-differentiated axes. In contrast, regarding GAT229, both the alternative binding mode inside its postulated binding site and CB1R conformations induced upon such binding are much more instable than the corresponding ones to the experimentally reported allosteric site (compare GAT229 poses 1 and 2 in Table 3 with GAT229 in Table 5), discarding its postulated binding site as well. The observed binding instability is due to the superficiality of GAT229 binding in given site, while the observed conformational instability is because of an increase of the spacing between TM2 and TM3 extracellular ends induced by GAT229 binding between them. This propagates to TM1 and TM4 that exhibit an outward movement of their extracellular ends. Additionally, this TM2 and TM3 spacing avoids the compound TM7 movement along with its accompanying TM6 extracellular end and α8 helix movements, which are associated with CB1R fully-active states since, as exposed in Section 3.3, the resulting CB1R conformations have a better G-protein binding (see Figure 13). Interestingly, neither GAT228 nor GAT229 binding in their postulated binding sites induce CB1R conformations that exhibit structural features compatible with its active states (see Section 3.3), which, along with the aforementioned instability, discards given postulated binding sites totally.

Finally, it is important to note that mutagenesis studies to determine the CB1R positive allosteric sites discard, among others, the postulated binding sites considered in this work. However, the paper corresponding to such mutagenesis studies [31] stated that allosteric-modulators binding takes place mainly in the experimentally reported allosteric site [29], though the putative binding site 2 cannot be totally dismissed due to a notable decrease in GAT229 activity upon mutation Arg220Ala$^{3.56}$. This latter conclusion is not totally correct since given activity decrease is also observed in a comparable extent for the remaining considered compound in that paper. In particular, for CP55940, ZCZ011, and GAT228, which indicates that this arginine takes part in the global CB1R activation process since this is related with both allosteric activation and orthosteric activation, discarding its associated putative binding site as well.

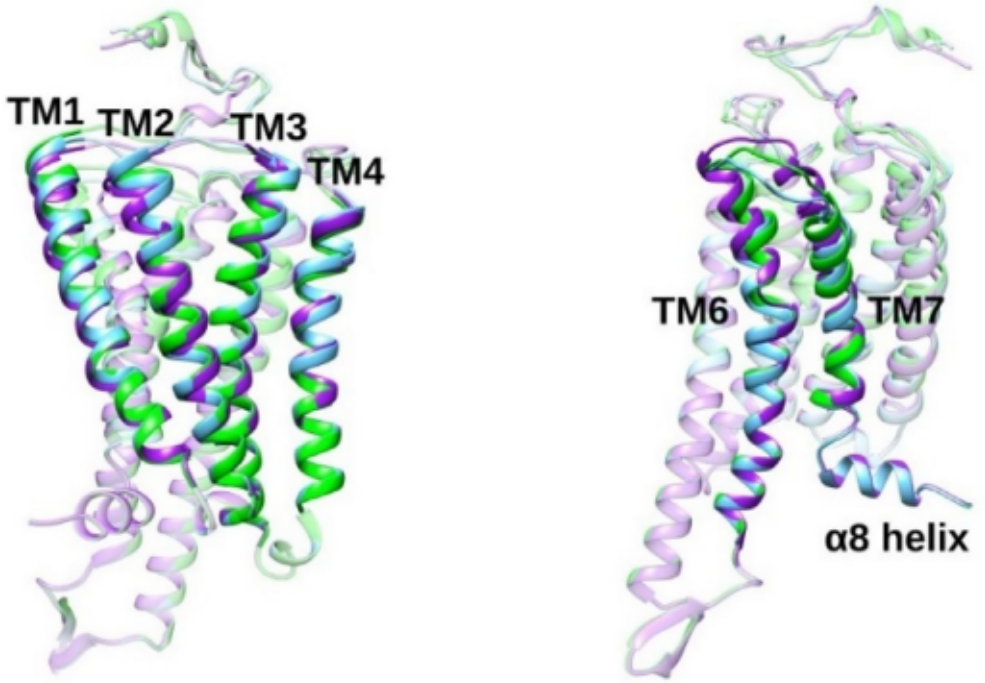

FIGURE 13 | Overlap of GAT229:CB1R complexes derived from binding on the allosteric and postulated binding sites. Significantly populated GAT229:CB1R complexes derived from binding on the allosteric site (pose 1 in green and pose 2 in blue) and on the GAT229 postulated site (in purple) have been depicted. It is worth noticing that GAT229 binding in its putative binding site provokes a spacing between TM2 and TM3 extracellular ends, which is further propagated to TM1 and TM4 that exhibit an outward movement of their extracellular ends (compare purple structure with green and blue ones). All this together avoids the compound TM7 movement along with its accompanying TM6 extracellular end and α8 helix movements, which are associated with CB1R fully-active states.

TABLE 6 | The 6s experimentally reported EC_{50}, its derived pK_B and $\Delta G_{bind}^{exp,allo}$ and ΔG_{bind}^{allo} for the 6s racemic mixture calculated from relative populations of its binary complexes according to ΔG_{total}^{allo} of them. The experimental EC_{50} of 6s was collected from a previous paper [22]. Its corresponding pK_B was calculated assuming that $K_B \approx EC_{50}$ and employing $pK_B = -\log(K_B)$ where its corresponding error has been obtained propagating EC_{50} standard deviation according to the uncertainty propagation law. $\Delta G_{bind}^{exp,allo}$ has been calculated from EC_{50} employing Equation (4) and its corresponding error has been obtained propagating EC_{50} standard deviation according to the uncertainty propagation law.

Compound	Experimental EC_{50}	Experimental pK_B	$\Delta G_{bind}^{exp,allo}$	ΔG_{bind}^{allo}
6s	26 ± 0.13	7.585 ± 0.002	−10.411 ± 0.003	−10.28

3.6 | Importance of Polar/Charged Residues/Groups on the CB1R Allosteric Site

The incorporation of polar groups in CB1R allosteric modulators that belong to the 2-phenylindole structural class can improve their affinity significantly. GAT221 would be the best compound to illustrate this fact since its intrinsic affinity has been increased through fluoridation of its carbon atom skeleton or substitution of its indole ring phenyl substituent by heteroaromatic substituents [22, 23]. All this brings to light that relevant polar/charged residues/groups must be present in the CB1R positive allosteric site. However, apart from Ser199$^{3.35}$, whose role is forming hydrogen bonds with the NO_2 group according to our results (see Section 3.1), no residue of this nature has been reported experimentally for that site [29]. With the aim of shedding light on this issue, the binding modes of 6s in the CB1R positive allosteric site have been studied similarly. 6s arises from including a fluorine atom in each of GAT221 ring systems (see Figure 2). Obviously, the fluorine atom contained in the 2-phenylindole ring is always pointing towards the same direction since the NH group of this ring must form a hydrogen bond with the oxygen backbone atom of Phe191$^{3.27}$. However, this is not the case for the other two fluorine atoms, given that the phenyl rings where these fluorine atoms are placed have total rotational freedom. For this reason, 16 distinct binding modes have been considered for each 6s enantiomer which derived from combining the 4 original binding modes with the two possible orientations, towards the CB1R positive allosteric site or cellular membrane, of the fluorine atoms whose phenyl rings have free rotation. For the sake of clarity, the fluorine atom placed at the indole ring phenyl substituent is called as F_1, while the fluorine atom placed

at the phenyl ring at β position of the NO_2 group is called F_2. Additionally, the orientation towards the CB1R positive allosteric site is referred as an inwards orientation, while the orientation towards cellular membrane is referred as an outwards orientation.

In Table 6 appears the 6s experimentally reported EC_{50}, its derived pK_B and $\Delta G_{btnd}^{exp,allo}$ and ΔG_{btnd}^{allo} for the 6s racemic mixture calculated from relative populations of its binary complexes according to ΔG_{total}^{allo} of them (see Tables S3 and S4 for ΔG_{btnd}^{allo}, ΔG_{fold}^{allo}, ΔG_{total}^{allo} and relative populations of binary complexes derived from both 6s enantiomers in their different poses). As for previous compounds, it has been obtained a highly similar value to $\Delta G_{btnd}^{exp,allo}$, however, due to the tiny standard deviation reported for EC_{50}, a complete match between ΔG_{total}^{allo} and $\Delta G_{btnd}^{exp,allo}$ has not been achieved. It is worth noting that the 6s experimental EC_{50} has been obtained employing a fitting model to determine this magnitude where the standard deviation for three independent fittings has been reported as its error. This way of proceeding can underestimate EC_{50} error since this also depends on the goodness of fit (not reported), which seems to be the case given the reported accuracy for other equivalent measures of the same work (e.g., 6s EC_{50} in presence of CP55940 = 13.7 ± 11.3 nM). This underestimation of EC_{50} error would justify the small mismatched observed between ΔG_{total}^{allo} and $\Delta G_{btnd}^{exp,allo}$. Regarding 6s *S* enantiomer, its significantly populated binary complexes derived from its pose 1 with F_1 pointing outwards while F_2 inwards, its pose 2 with F_1 pointing inwards while F_2 outwards, its pose 4 with F_1 pointing inwards while F_2 outwards and its pose 4 with F_1 and F_2 pointing outwards (relative populations of 0.3585, 0.1762, 0.1339, and 0.3042, respectively). In contrast, 6s *R* enantiomer only exhibits two significantly populated binary complexes that derived from its pose 1 with F_1 pointing outwards while F_2 inwards and its pose 2 with F_1 and F_2 pointing outwards (relative populations of 0.2939 and 0.6958, respectively). It is important to note that fluorine atoms incorporation does not prevent binding modes 2 and 4, poses in which the NO_2 group is hydrogen bonded to Ser199$^{3.35}$, from being the most relevant ones for binary complexes. This result highlights again the key role of both the NO_2 group and Ser199$^{3.35}$ in relation to CB1R positive allosteric modulation, which is in agreement with the experimental observations [29, 57].

On the other hand, additional relevant electrostatic interactions of the CB1R positive allosteric site, in other words, without considering hydrogen bonds between the 2-phenylindole NH group and the oxygen backbone atom of Phe191$^{2.27}$ and between the NO_2 group and the Ser199$^{3.35}$ sidechain, can be unveiled through analysis of 6s fluorine atoms interactions. According to different poses that both 6s enantiomers adopt in the CB1R positive allosteric site for their significantly populated binary complexes, it can be established that their fluorine atoms mainly interact with the sidechain NH_3^+ group of Lys192$^{3.28}$, the backbone NH group of Gly194$^{3.30}$ and the sidechain OH group of Ser199$^{3.35}$ by means of hydrogen bonds (see Figure 14). These interactions justify the better 6s affinity in relation to its parent compound since 6s preserves most GAT221 favorable interactions, whose character is mainly hydrophobic, and, additionally, incorporates the electrostatic interactions coming from its fluorine atoms. According to our results,

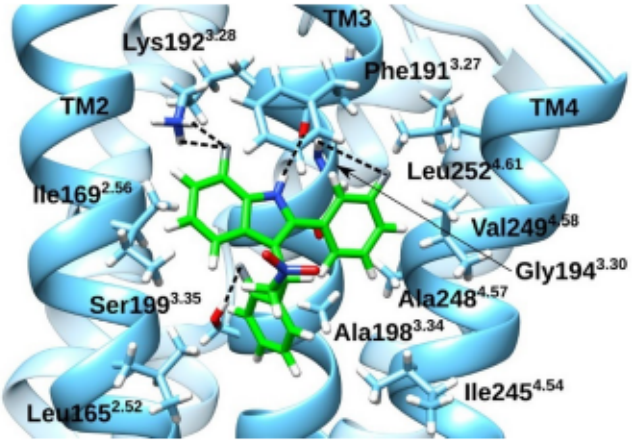

FIGURE 14 | Schematic representation of additional relevant electrostatic interactions for the CB1R positive allosteric site caused by inclusion of extra polar/charged groups to allosteric modulator. 6s *S* enantiomer in its pose 1 with F_1 and F_2 pointing inwards has been depicted since all these interactions take place simultaneously (see hydrogen bonds which involve Lys192$^{3.28}$ and Gly194$^{3.30}$). Hydrogen bonds have been indicated by dash lines.

Lys192$^{3.28}$ and Gly194$^{3.30}$ give rise to additional relevant electrostatic interactions of the CB1R positive allosteric site, opening the door to improve the affinity of mainly hydrophobic CB1R allosteric modulators by targeting them. Additionally, though there may be some other relevant polar/charged groups, electrostatic interactions that arise from Phe191$^{2.27}$, Lys192$^{3.28}$, Gly194$^{3.30}$, and Ser199$^{3.35}$ allow to rationalize most SAR studies of GAT221 analogs derived from fluoridation of its carbon atom skeleton, incorporation of nitrogen atoms inside this skeleton and substitution of its indole ring phenyl substituent by heteroaromatic substituents [22, 23]. In particular, incorporation of hydrogen bond acceptors in the two substitutable positions of the 2-phenylindole six-membered ring adjacent to its NH group is favored due to interactions with the sidechain NH_3^+ group of Lys192$^{3.28}$ (poses 1 and 2) or the backbone NH group of Gly194$^{3.30}$ (poses 3 and 4). However, the CB1R positive allosteric site is highly hydrophobic and access to its hydrogen bond donors that give rise to additional relevant electrostatic interactions is restricted, provoking that hydrogen bond acceptors incorporated to allosteric modulator must have a particular size in order to not cause unfavorable interactions with hydrophobic residues or the oxygen backbone atom of Phe191$^{2.27}$, which explains the opposed effect observed for fluoridation and incorporation of nitrogen atoms in such positions. In relation to incorporation of hydrogen bond acceptors in the ring at β position of the NO_2 group, only the ortho position allows to establish hydrogen bonds with the sidechain OH group of Ser199$^{3.35}$ (poses 2 and 4) without causing unfavorable interactions with hydrophobic residues found in this region or cellular membrane. Again, bias in the incorporated hydrogen bond acceptors is observed. It is worth noting that simultaneous incorporation in both ortho positions of the considered ring would be unfavorable since one of the hydrogen bond acceptors would be pointing to cellular membrane, as observed experimentally. Regarding the modification of the indole ring phenyl substituent, incorporation of a hydrogen bond acceptor is favorable in ortho and meta positions for both

six-membered rings and five-membered rings due to interactions with the sidechain NH_3^+ group of Lys192$^{3.28}$ (poses 3 and 4) or the backbone NH group of Gly194$^{3.30}$ (poses 1 and 2). It is important to note that given incorporation is favorable, providing it does not entail incorporation of a hydrogen bond donor or an additional hydrogen bond acceptor, since otherwise unfavorable interactions with cavity hydrophobic residues would be established. Similarly, bias in the incorporated hydrogen bond acceptors is also observed, indeed, a preferential trend of sulfur atom > oxygen atom > nitrogen atom was established. Surprisingly, incorporation of a hydrogen bond donor in ortho position is tolerated in this particular case due to the possibility of forming a hydrogen bond with the oxygen backbone atom of Phe191$^{2.27}$. It is important to note that most polar/charged residues/groups of the CB1R positive allosteric site with the ability of interacting with allosteric modulators are hydrogen bond donors, that is why incorporation of hydrogen bond acceptors is better tolerated.

Regarding allosteric modulation activity, it has been determined experimentally that 6s racemic mixture behaves as an ago-PAM [22]. Nevertheless, according to results of Section 3.3, each 6s enantiomer must give rise to a single different positive allosteric modulation activity, in particular, its *R* enantiomer must be related to PAM activity, while its *S* enantiomer must be associated with allosteric agonist activity taking into account the Cahn-Ingold-Prelog rules. Considering significantly populated binary complexes for 6s enantiomers, it can be drawn that this relationship between positive allosteric modulation activities and spatial arrangement of substituents around the carbon atom in β position of the NO_2 group (chiral carbon atom 1) is fulfilled since the complexes of each enantiomer exhibit the structural features associated with its corresponding positive allosteric modulation activity according to given relationship. Specifically, it has been observed that 6s *S* enantiomer provokes an outwards displacement of TM4 extracellular end, suffering a slight loss of α helix content, that distances this α helix from TM2, TM3, and TM5, which is consistent with allosteric agonist activity, while 6s *R* enantiomer also causes this displacement, though now TM4 extracellular end undergoes a significant loss of α helix content and this transmembrane helix distances from TM2 and TM3 but approaches to TM5 (see Figure S14). This result reinforces the hypothesis that in the case of the 2-phenylindole compounds with a single chiral carbon atom, each CB1R positive allosteric modulation activity can be associated with a specific spatial arrangement of substituents around this chiral carbon atom.

4 | Conclusions

Employing the PDLD/S-2000 approach along with our refined CG model, it has been unveiled the subtleties of CB1R positive allosteric modulation arisen from its allosteric modulators belonging to the 2-phenylindole structural class.

Our methodology allowed to reproduce the experimental K_Bs for a diverse assortment of this type of compounds well-characterized experimentally considering CP55940 as an agonist. This indicates that we identified and characterized binary and ternary complexes arisen from them in their different poses successfully. In other words, both ΔG_{bind} for allosteric modulators and CP55940 in their different poses in their respective binding sites and ΔG_{fold} for the resulting binary and ternary complexes have been determined correctly. Thus allowing the calculation of the relative populations of these complexes through ΔG_{total}.

Relative populations for binary complexes justify the importance of both Ser199$^{3.35}$ and the NO_2 group of this type of compounds in CB1R positive allosteric modulation. That is why the main binary complexes derived from allosteric modulators in their binding mode 2, in which hydrogen bonds between both are established. In contrast, relative populations for ternary complexes indicate that their significantly populated complexes derived from allosteric modulators in their binding modes 1 and 3, which agrees with the experimentally reported ZCZ011 binding mode.

Analyzing structural features and interactions between allosteric modulators and the CB1R positive allosteric site for significantly populated complexes, both the structural basis and the origin at molecular level of CB1R positive allosteric modulation have been determined. In the case of binary complexes, this modulation lies in an outwards displacement of TM4 extracellular end whose direction and loss of α helix content change according to the type of positive allosteric modulation. Regarding the allosteric modulators-CB1R interactions, PAM activity is associated with a spatial arrangement of substituents around the chiral carbon atom 1 equal to the GAT229's one. We note that it arises from a deeper insertion between TM3 and TM4 of the allosteric modulator ring placed at the right side of this first transmembrane helix, which is linked to a better π-π stacking interaction with Phe191$^{3.27}$ and a more favorable hydrophobic interaction with TM4. On the other hand, the allosteric agonist activity, associated with a spatial arrangement of substituents around the chiral carbon atom 1 equal to the GAT228's one or insertion of chirality in the chiral carbon atom 2, is linked to a more intense favorable interaction with residues Leu165$^{2.52}$, Ile169$^{2.56}$, Ala198$^{3.34}$, and Ser199$^{3.35}$. It is worth noting that this observed relationship between the spatial arrangement of substituents around chiral carbon atoms and the type of CB1R positive allosteric modulation suggests that the stereoselectivity is an inherent property of this type of compounds. In the case of ternary complexes, CB1R positive allosteric modulation lies in the regulation of the range of motion of the compound TM7 movement apparently caused by the orthosteric agonist itself. Although the differences are less noticeable, the observed trends for binary complexes in relation with allosteric modulators-CB1R interactions are largely still valid. Additionally, it has been determined that residues Lys192$^{3.28}$ and Gly194$^{3.30}$ are responsible for additional electrostatic interactions established by allosteric modulators belonging to the 2-phenylindole structural class that incorporate extra polar groups. These two residues along with Phe191$^{3.27}$ and Ser199$^{3.35}$ allow the rationalization of most SAR studies for this type of compounds.

On the other hand, our methodology is also effective for screening of putative CB1R positive allosteric sites since the CB1R positive allosteric site has been successfully identified through the calculation of $\Delta G_{total}^{allo/allo(CP55940)}$ of the resulting complexes upon binding in those sites.

Our study offers a deeper insight into GPCRs allosteric modulation focusing on its positive branch as well as a better comprehension of the molecular architecture of the CB1R positive allosteric site. It also sheds light in how allosteric modulators belonging to the 2-phenylindole structural class interact with this site and give rise to their allosteric modulation. All this points out towards the predictive value of the methodology used in the current work, which can be applied to other biophysical systems of interest. We expect that this study will contribute to both encourage a more thorough exploration of GPCRs allosteric modulation and the development of new CB1R positive allosteric modulators with enhanced properties.

Author Contributions

Alejandro Cruz: conceptualization, investigation, writing – original draft, methodology, visualization, formal analysis. **Arieh Warshel:** funding acquisition, writing – review and editing, conceptualization, supervision, methodology, resources, investigation, project administration.

Acknowledgments

This work was supported by the National Institutes of Health (R35 GM122472) and the National Science Foundation (Grant MCB 1707167).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/apl/gateway/wos/peer-review/10.1002/prot.26762.

References

1. K. L. Pierce, R. T. Premont, and R. J. Lefkowitz, "Seven-Transmembrane Receptors," *Nature Reviews. Molecular Cell Biology* 3, no. 9 (2002): 639–650.

2. A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiöth, and D. E. Gloriam, "Trends in GPCR Drug Discovery: New Agents, Targets and Indications," *Nature Reviews. Drug Discovery* 16, no. 12 (2017): 829–842.

3. W. I. Weis and B. K. Kobilka, "The Molecular Basis of G Protein-Coupled Receptor Activation," *Annual Review of Biochemistry* 87 (2018): 897–919.

4. G. Turu and L. Hunyady, "Signal Transduction of the CB1 Cannabinoid Receptor," *Journal of Molecular Endocrinology* 44, no. 2 (2010): 75–85.

5. M. S. Ibsen, M. Connor, and M. Glass, "Cannabinoid CB1 and CB2 Receptor Signaling and Bias," *Cannabis and Cannabinoid Research* 2, no. 1 (2017): 48–60.

6. R. B. Laprairie, A. M. Bagher, M. E. M. Kelly, and E. M. Denovan-Wright, "Biased Type 1 Cannabinoid Receptor Signaling Influences Neuronal Viability in a Cell Culture Model of Huntington Disease," *Molecular Pharmacology* 89, no. 3 (2016): 364–375.

7. S. Garai, L. M. Leo, A. M. Szczesniak, et al., "Discovery of a Biased Allosteric Modulator for Cannabinoid 1 Receptor: Preclinical Anti-Glaucoma Efficacy," *Journal of Medicinal Chemistry* 64, no. 12 (2021): 8104–8126.

8. P. Pacher, S. Bátkai, and G. Kunos, "The Endocannabinoid System as an Emerging Target of Pharmacotherapy," *Pharmacological Reviews* 58, no. 3 (2006): 389–462.

9. J. Wang, C. Garert, and H. A. Rockman, "G-Protein–Coupled Receptors in Heart Disease," *Circulation Research* 123, no. 6 (2018): 716–735.

10. M. Rask-Andersen, M. S. Almén, and H. B. Schiöth, "Trends in the Exploitation of Novel Drug Targets," *Nature Reviews Drug Discovery* 10, no. 8 (2011): 579–590.

11. K. Sriram and P. A. Insel, "G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs?," *Molecular Pharmacology* 93, no. 4 (2018): 251–258.

12. E. A. Wold, J. Chen, K. A. Cunningham, and J. Zhou, "Allosteric Modulation of Class A GPCRs: Targets, Agents, and Emerging Concepts," *Journal of Medicinal Chemistry* 62, no. 1 (2019): 88–127.

13. B. J. Melancon, C. R. Hopkins, M. R. Wood, et al., "Allosteric Modulation of Seven Transmembrane Spanning Receptors: Theory, Practice, and Opportunities for Central Nervous System Drug Discovery," *Journal of Medicinal Chemistry* 55, no. 4 (2012): 1445–1464.

14. E. Jacoby, R. Bouhelal, M. Gerspacher, and K. Seuwen, "The 7TM G-Protein-Coupled Receptor Target Family," *ChemMedChem* 1, no. 8 (2006): 760–782.

15. D. Wootten, A. Christopoulos, and P. M. Sexton, "Emerging Paradigms in GPCR Allostery: Implications for Drug Discovery," *Nature Reviews Drug Discovery* 12, no. 8 (2013): 630–644.

16. E. T. Van Der Westhuizen, C. Valant, P. M. Sexton, and A. Christopoulos, "Endogenous Allosteric Modulators of G Protein–Coupled Receptors," *Journal of Pharmacology and Experimental Therapeutics* 353, no. 2 (2015): 246–260.

17. T. Kenakin, "Allosteric Theory: Taking Therapeutic Advantage of the Malleable Nature of GPCRs," *Current Neuropharmacology* 5, no. 3 (2007): 149–156.

18. D. J. Foster and P. J. Conn, "Allosteric Modulation of GPCRs: New Insights and Potential Utility for Treatment of Schizophrenia and Other CNS Disorders," *Neuron* 94, no. 3 (2017): 431–446.

19. T. P. Kenakin and T. Kenakin, "Biased Signalling and Allosteric Machines: New Vistas and Challenges for Drug Discovery," *British Journal of Pharmacology* 165, no. 6 (2012): 1659–1669.

20. H. M. Green, D. B. Finlay, R. A. Ross, I. R. Greig, S. B. Duffull, and M. Glass, "In Vitro Characterization of 6-Methyl-3-(2-Nitro-1-(Thiophen-2-Yl)Ethyl)-2-Phenyl-1 H-Indole (ZCZ011) at the Type 1 Cannabinoid Receptor: Allosteric Agonist or Allosteric Modulator?," *ACS Pharmacology & Translational Science* 5, no. 12 (2022): 1279–1291.

21. R. B. Laprairie, P. M. Kulkarni, J. R. Deschamps, et al., "Enantiospecific Allosteric Modulation of Cannabinoid 1 Receptor," *ACS Chemical Neuroscience* 8, no. 6 (2017): 1188–1203.

22. S. Garai, P. M. Kulkarni, P. C. Schaffer, et al., "Application of Fluorine- and Nitrogen-Walk Approaches: Defining the Structural and Functional Diversity of 2-Phenylindole Class of Cannabinoid 1 Receptor Positive Allosteric Modulators," *Journal of Medicinal Chemistry* 63, no. 2 (2020): 542–568.

23. S. Garai, P. C. Schaffer, R. B. Laprairie, et al., "Design, Synthesis, and Pharmacological Profiling of Cannabinoid 1 Receptor Allosteric Modulators: Preclinical Efficacy of C2-Group GAT211 Congeners for Reducing Intraocular Pressure," *Bioorganic & Medicinal Chemistry* 50 (2021): 116421.

24. R. G. Pertwee, "Endocannabinoids and Their Pharmacological Actions," in *Handbook of Experimental Pharmacology*, vol. 231, ed. R. G. Pertwee (Cham: Springer, 2015), 1–37.

25. R. G. Pertwee, A. C. Howlett, M. E. Abood, et al., "International Union of Basic and Clinical Pharmacology. LXXIX. Cannabinoid Receptors and Their Ligands: Beyond CB1 and CB2," *Pharmacological Reviews* 62, no. 4 (2010): 588–631.

26. D. Haspula and M. A. Clark, "Cannabinoid Receptors: An Update on Cell Signaling, Pathophysiological Roles and Therapeutic Opportunities in Neurological, Cardiovascular, and Inflammatory Diseases," *International Journal of Molecular Sciences* 21, no. 20 (2020): 7693.

27. R. G. Pertwee, "Ligands That Target Cannabinoid Receptors in the Brain: From THC to Anandamide and Beyond," *Addiction Biology* 13, no. 2 (2008): 147–159.

28. N. Raz, A. M. Eyal, D. B. Zeitouni, et al., "Selected Cannabis Terpenes Synergize With THC to Produce Increased CB1 Receptor Activation," *Biochemical Pharmacology* 212 (2023): 115548.

29. X. Yang, X. Wang, Z. Xu, et al., "Molecular Mechanism of Allosteric Modulation for the Cannabinoid Receptor CB1," *Nature Chemical Biology* 18, no. 8 (2022): 831–840.

30. D. P. Hurst, S. Garai, P. M. Kulkarni, P. C. Schaffer, P. H. Reggio, and G. A. Thakur, "Identification of Cb1 Receptor Allosteric Sites Using Force-Biased MMC Simulated Annealing and Validation by Structure-Activity Relationship Studies," *ACS Medicinal Chemistry Letters* 10, no. 8 (2019): 1216–1221.

31. H. M. Green, D. M. J. Fellner, D. B. Finlay, D. P. Furkert, and M. Glass, "Determination of the Cannabinoid CB1 Receptor's Positive Allosteric Modulator Binding Site Through Mutagenesis Studies," *Pharmaceuticals* 17, no. 2 (2024): 154.

32. F. S. Lee, Z. T. Chu, and A. Warshel, "Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the POLARIS and ENZYMIX Programs," *Journal of Computational Chemistry* 14, no. 2 (1993): 161–185.

33. I. Vorobyov, I. Kim, Z. T. Chu, and A. Warshel, "Refining the Treatment of Membrane Proteins by Coarse-Grained Models," *Proteins: Structure, Function, and Bioinformatics* 84, no. 1 (2016): 92–117.

34. M. Lee, V. Kolev, and A. Warshel, "Validating a Coarse-Grained Voltage Activation Model by Comparing Its Performance to the Results of Monte Carlo Simulations," *Journal of Physical Chemistry B* 121, no. 50 (2017): 11284–11291.

35. S. Vicatos, A. Rychkova, S. Mukherjee, and A. Warshel, "An Effective Coarse-Grained Model for Biological Simulations: Recent Refinements and Validations," *Proteins: Structure, Function, and Bioinformatics* 82, no. 7 (2014): 1168–1185.

36. R. Alhadeff and A. Warshel, "A Free-Energy Landscape for the Glucagon-Like Peptide 1 Receptor GLP1R," *Proteins: Structure, Function, and Bioinformatics* 88, no. 1 (2020): 127–134.

37. R. Alhadeff, I. Vorobyov, H. W. Yoon, and A. Warshel, "Exploring the Free-Energy Landscape of GPCR Activation," *Proceedings of the National Academy of Sciences of the United States of America* 115, no. 41 (2018): 10327–10332.

38. J. A. Ballesteros and H. Weinstein, "Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors," *Methods in Neurosciences* 25, (1995): 366–428.

39. A. Warshel, Z. T. Chu, J. Villa, et al., *Molaris-Xg. Version 9.15* (Los Angeles: University of Southern California, 2012).

40. M. J. Frisch, *Gaussian 16. Revision A03* (Wallingford CT: Gaussian Inc, 2016).

41. C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, "A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model," *Journal of Physical Chemistry* 97, no. 40 (1993): 10269–10280.

42. D. A. Case, T. E. Cheatham, T. Darden, et al., "The Amber Biomolecular Simulation Programs," *Journal of Computational Chemistry* 26, no. 16 (2005): 1668–1688.

43. A. Warshel and G. King, "Polarization Constraints in Molecular Dynamics Simulation of Aqueous Solutions: The Surface Constraint all Atom Solvent (SCAAS) Model," *Chemical Physics Letters* 121, no. 1–2 (1985): 124–129.

44. F. S. Lee and A. Warshel, "A Local Reaction Field Method for Fast Evaluation of Long-Range Electrostatic Interactions in Molecular Simulations," *Journal of Chemical Physics* 97, no. 5 (1992): 3100–3107.

45. N. Singh and A. Warshel, "Absolute Binding Free Energy Calculations: On the Accuracy of Computational Scoring of Protein–Ligand Interactions," *Proteins: Structure, Function, and Bioinformatics* 78, no. 7 (2010): 1705–1723.

46. Y. Y. Sham, Z. T. Chu, H. Tao, and A. Warshel, "Examining Methods for Calculations of Binding Free Energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA Calculations of Ligands Binding to an HIV Protease," *Proteins: Structure, Function, and Bioinformatics* 39, no. 4 (2000): 393–407.

47. A. Nandi, A. Zhang, E. Arad, R. Jelinek, and A. Warshel, "Assessing the Catalytic Role of Native Glucagon Amyloid Fibrils," *ACS Catalysis* 14, no. 7 (2024): 4656–4664.

48. W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual Molecular Dynamics," *Journal of Molecular Graphics* 14, no. 1 (1996): 33–38.

49. E. F. Pettersen, T. D. Goddard, C. C. Huang, et al., "UCSF Chimera – A Visualization System for Exploratory Research and Analysis," *Journal of Computational Chemistry* 25, no. 13 (2004): 1605–1612.

50. J. M. Stockton, N. J. Birdsall, A. S. Burgen, and E. C. Hulme, "Modification of the Binding Properties of Muscarinic Receptors by Gallamine," *Molecular Pharmacology* 23, no. 3 (1983): 551–557.

51. F. J. Ehlert, "Estimation of the Affinities of Allosteric Ligands Using Radioligand Binding and Pharmacological Null Methods," *Molecular Pharmacology* 33, no. 2 (1988): 187–194.

52. M. R. Price, G. L. Baillie, A. Thomas, et al., "Allosteric Modulation of the Cannabinoid CB1 Receptor," *Molecular Pharmacology* 68, no. 5 (2005): 1484–1495.

53. J. W. Black and P. Leff, "Operational Models of Pharmacological Agonism," *Proceedings of the Royal Society of London - Series B: Biological Sciences* 220, no. 1219 (1983): 141–162.

54. T. Kenakin, "New Concepts in Drug Discovery: Collateral Efficacy and Permissive Antagonism," *Nature Reviews Drug Discovery* 4, no. 11 (2005): 919–927.

55. J. Jakubík, A. Randáková, N. Chetverikov, E. E. El-Fakahany, and V. Doležal, "The Operational Model of Allosteric Modulation of Pharmacological Agonism," *Scientific Reports* 10, no. 1 (2020): 1–20.

56. R. Sadana and C. W. Dessauer, "Physiological Roles for G Protein-Regulated Adenylyl Cyclase Isoforms: Insights From Knockout and Overexpression Studies," *Neurosignals* 17, no. 1 (2009): 5–22.

57. D. Lu, S. S. Immadi, Z. Wu, and D. A. Kendall, "Translational Potential of Allosteric Modulators Targeting the Cannabinoid CB1 Receptor," *Acta Pharmacologica Sinica* 40, no. 3 (2018): 324–335.

58. C. C. Tseng, G. Baillie, G. Donvito, et al., "The Trifluoromethyl Group as a Bioisosteric Replacement of the Aliphatic Nitro Group in CB1 Receptor Positive Allosteric Modulators," *Journal of Medicinal Chemistry* 62, no. 10 (2019): 5049–5062.

59. M. E. Maguire, P. M. van Arsdale, and A. G. Gilman, "An Agonist-Specific Effect of Guanine Nucleotides on Binding to the Beta Adrenergic Receptor," *Molecular Pharmacology* 12, no. 2 (1976): 335–339.

60. A. De Lean, J. M. Stadel, and R. J. Lefkowitz, "A Ternary Complex Model Explains the Agonist-Specific Binding Properties of the

Adenylate Cyclase-Coupled Beta-Adrenergic Receptor," *Journal of Biological Chemistry* 255, no. 15 (1980): 7108–7117.

61. X. Wang, D. Liu, L. Shen, et al., "A Genetically Encoded F-19 NMR Probe Reveals the Allosteric Modulation Mechanism of Cannabinoid Receptor 1," *Journal of the American Chemical Society* 143, no. 40 (2021): 16320–16325.

62. M. Bari, A. Paradisi, N. Pasquariello, and M. Maccarrone, "Cholesterol-Dependent Modulation of Type 1 Cannabinoid Receptors in Nerve Cells," *Journal of Neuroscience Research* 81, no. 2 (2005): 275–283.

63. K. Krishna Kumar, M. J. Robertson, E. Thadhani, et al., "Structural Basis for Activation of CB1 by an Endocannabinoid Analog," *Nature Communications* 14, no. 1 (2023): 1–11.

64. K. Krishna Kumar, M. Shalev-Benami, M. J. Robertson, et al., "Structure of a Signaling Cannabinoid Receptor 1-G Protein Complex," *Cell* 176, no. 3 (2019): 448–458.e12.

65. T. Hua, K. Vemuri, S. P. Nikas, et al., "Crystal Structures of Agonist-Bound Human Cannabinoid Receptor CB1," *Nature* 547, no. 7664 (2017): 468–471.

66. T. Hua, X. Li, L. Wu, et al., "Activation and Signaling Mechanism Revealed by Cannabinoid Receptor-Gi Complex Structures," *Cell* 180, no. 4 (2020): 655–665.e18.

67. M. Clark, F. Guarnieri, I. Shkurko, and J. Wiseman, "Grand Canonical Monte Carlo Simulation of Ligand–Protein Binding," *Journal of Chemical Information and Modeling* 46, no. 1 (2005): 231–242.

68. F. Guarnieri and M. Mezei, "Simulated Annealing of Chemical Potential: A General Procedure for Leading Bound Waters. Application to the Study of the Differential Hydration Propensities of the Major and Minor Grooves of DNA," *Journal of the American Chemical Society* 118, no. 35 (1996): 8493–8494.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.