Negative allosteric modulators of cannabinoid receptor 1: Ternary complexes including CB1, orthosteric CP55940 and allosteric ORG27569

AyoOluwa O. Aderibigbe, Pankaj Pandey, and Robert J. Doerksen

Communicated by Ramaswamy H. Sarma

ABSTRACT
In October 2019, the first X-ray crystal structure of a ternary cannabinoid receptor 1 (CB1) complex (PDB ID: 6KQI) was published, including the well-known orthosteric agonist, CP55940, and the well-studied negative allosteric modulator, ORG27569. Prior to the release of 6KQI, we applied binding pocket analysis and molecular docking to carefully prepared computational models of the ternary CB1 complex, in order to predict the binding site for ORG27569 with the CP55940-bound CB1 receptor. We carefully studied the binding pose of agonist ligands in the CB1 orthosteric pocket, including CP55940. Our computational studies identified the most favorable binding site for ORG27569, in the CP55940–CB1 complex, to be at the intracellular end of the receptor. However, in the 6KQI structure, ORG27569 was found at an extrahelical, intramembrane site on the complex, a site that partially overlaps with the site predicted in our calculations to be second-best. We performed molecular dynamics simulations of the CP55940-bound CB1 complex with ORG27569 at different binding sites. Our analysis of the simulations indicated that ORG27569 bound favorably and stably in each simulation, but, as in the earlier calculations, bound best at the intracellular site, which is different than that found in the crystal structure. These results suggest that the intracellular site might serve as an alternative binding site in CB1. Our studies show that the computational techniques we used are valuable in identifying ligand-binding pockets in proteins, and could be useful for the study of the interaction mode of other allosteric modulators.

Introduction

G-protein coupled receptors (GPCRs) are the largest family of proteins targeted by FDA-approved drugs, with small molecules making up the majority of the probes/agents applied to these targets (Congreve et al., 2020; Santos et al., 2017; Sriram & Insel, 2018). However, the use of small molecules (ligands) to modify receptor function has been...
impacted by certain limitations. First, the binding site of the endogenous ligands of most GPCRs (the orthosteric binding site) is highly conserved, especially within the same family of receptors. Thus, some ligands have been shown to be promiscuous, binding to multiple receptor subtypes. Secondly, highly potent ligands which have better binding affinities and efficacies compared to the endogenous ligands of the receptors produce marked biological effects due to pronounced receptor activation or inhibition. This may result in adverse effects from the amplification of unwanted receptor signaling and limit the development of certain potent ligands as drugs for therapeutic use (Moritz et al., 2018; Roth et al., 2004; Wacker et al., 2017). This may also explain the toxic properties of illicit street drugs, such as the synthetic cannabinoid receptor agonists (Adams et al., 2017; Sachdev et al., 2019). Allosteric modulators, ligands that bind at a site topologically distinct from the orthosteric site, are generally believed to avoid these limitations. This is because allosteric binding sites are less conserved than the orthosteric binding site, and hence ligands that bind to such sites are less likely to be promiscuous. In addition, pure allosteric modulators, which do not alter receptor signaling in the absence of an orthosteric ligand, are considered to exhibit a saturable maximal effect (ceiling effect) in their modulation of the nature and strength of binding and receptor signaling by the co-bound orthosteric ligand (Wang et al., 2009; Wold et al., 2019).

With an increase in the incidence of obesity worldwide, measures other than changes in lifestyle or dietary modifications have been sought to combat this global trend (Obesity and overweight; Narayanaswami & Dwoskin, 2017). The human cannabinoid 1 (CB1) receptor is considered a promising target for the management of obesity. Studies have shown that antagonism of the CB1 receptor confers metabolic benefits, including a reduction in body weight (Jourdan et al., 2010; Trillou et al., 2003; Wei et al., 2018). This is an important target to consider as a step toward increasing the number and kinds of drugs available for treating obesity, especially since some existing anti-obesity drugs have serious limitations (Srivastava & Apovian, 2018). The development of conventional (orthosteric, centrally-acting) antagonists of the CB1 receptor as anti-obesity drugs has been hindered by CNS-mediated neuropsychiatric adverse effects observed in vulnerable populations (Jones, 2008). Different approaches have been sought to avoid the CNS-mediated adverse effects while taking advantage of the beneficial metabolic effects of CB1 antagonists (Stasiulewicz et al., 2020). One approach is to design peripherally-restricted orthosteric antagonists of the CB1 receptor (Chen et al., 2017; Chorvat et al., 2012; Cluny et al., 2010). The design of neutral antagonists, which do not alter the constitutive activity of the receptor but prevent the binding of endocannabinoids, has also been proposed as an approach to avoid these side effects (Gueye et al., 2016). A third approach involves the design of negative allosteric modulators of the CB1 receptor (Alaverdashvili & Laprairie, 2018; Morales & Jagerovic, 2020; Price et al., 2005).

The first set of allosteric modulators of the CB1 receptor was identified by Organon Research. The compounds were observed to increase the binding affinity of a co-bound orthosteric agonist while reducing receptor signaling (Price et al., 2005). Since the discovery of the first set of CB1 negative allosteric modulators (NAMs), other molecules have been identified as NAMs, including PSNCBAM-1 (Horswill et al., 2007), Pepcan-12 (a peptide endocannabinoid) (Bauer et al., 2012), pregnenolone (Vallée et al., 2014), fenofibrate (Priestley et al., 2015), cannabidiol (Laprairie et al., 2015), GAT100 (Kulkarni et al., 2016; Laprairie et al., 2016) and GW405833 (Dhopeshwarkar et al., 2017) (Figure 1). More complete lists of molecules that modulate the CB1 receptor allosterically are available in recent review articles (Alaverdashvili & Laprairie, 2018; Gado et al., 2019; Hryhorowicz et al., 2019; Nguyen et al., 2017). These molecules have been reported to exhibit varying effects that are dependent on the particular allosteric modulator and orthosteric ligand, including modulation of the

![Figure 1. Representative negative allosteric modulators of the cannabinoid CB1 receptor.](image)
binding affinity of the orthosteric ligand or, somewhat independently, of the efficacy of the orthosteric ligand and adjusting the bias of downstream signaling (Khurana et al., 2017). In addition, some CB1 NAMs have been shown to restore agonist binding lost due to mutation (Dopart & Kendall, 2020). Consequently, a wide range of possibilities of outcome and utility are created based on the choice of orthosteric agonist, NAM and receptor signaling pathway of study.

In attempts to understand the observed experimental effects of these allosteric modulators, studies have been performed to elucidate their CB1 binding sites and design other novel analogs (Bian et al., 2019; Hurst et al., 2019). Protein 3D structures obtained through experimental methods, such as those derived from fits to X-ray diffraction data, are widely considered to be the most accurate basis for protein-structure-based drug discovery methods like virtual screening and ligand binding pose prediction (Chen et al., 2019; Loo et al., 2018). For proteins without available X-ray structures, the accuracy of homology models is generally improved when X-ray structures of similar proteins or experimental information (e.g. mutagenesis data) are available (Beuming & Sherman, 2012; Jaiteh et al., 2020). However, the process of crystallization may distort the protein 3D structure obtained, introducing discrepancies relative to the true dynamic protein structure, and in such cases, computational protein models can prove helpful. This argument has been made with the recently published crystal structures of the CB1 receptor (Al-Zoubi et al., 2019). The manner in which known experimental information is interpreted and applied in generating constraints for computational modeling plays a huge role in determining how useful the protein models are (both homology models and experimental protein 3D structures). This is an essential factor to consider in establishing the accuracy of protein models to be used for making predictions in protein-structure-based drug design and for explaining and understanding biological processes at a molecular level.

An interesting case is found with ORG27569. Previous studies contradicted each other regarding the predicted interaction site for ORG27569 within the CP55940-bound complex of the CB1 receptor (Figure 2). CB1 residues that are present within these proposed allosteric sites for ORG27569 are listed in the caption of Figure 2. The allosteric modulator elicits contrasting effects on CB1 co-bound orthosteric ligands, CP55940 (agonist) and rimonabant (inverse agonist). ORG27569 increases CP55940 binding to the CB1 receptor (positive binding cooperativity). However, ORG27569 causes a saturable (incomplete) decrease in the binding of [3H]-rimonabant to the CB1 receptor (limited negative binding cooperativity) (Price et al., 2005).

In an early attempt to determine the binding site of ORG27569 within the CB1 receptor, Shore et al. (2014) interpreted the experiments such that ORG27569 caused a partial displacement of rimonabant from its binding site (Shore et al., 2014). Based on this interpretation, constraints were placed in their computational design such that the potential allosteric sites that were probed were located at the extracellular end of the CB1 receptor and in close proximity to the orthosteric binding site. Using a combination of mutagenesis studies and computational modeling, the ORG27569 binding site was predicted to have an overlap with the orthosteric site, but to extend toward the extracellular end of the protein, between transmembrane helices (TMHs) 3, 6 and 7 (Site 2 in Figure 2) (Shore et al., 2014). Prior research suggests that ORG27569 induces an intermediate conformation in the CB1 receptor that hampers G-protein signaling and simultaneously enhances agonist binding (Fay & Farrens, 2012). An alternative interpretation does not require an overlap between the orthosteric and allosteric ligand binding sites. A second study, by Stornaiuolo et al. (2015), that used mutagenesis of CB1 residues other than those used by Shore et al. (2014) as well as mass spectrometry and computational modeling, adopted this alternative interpretation and concluded that the ORG27569 allosteric site was toward the intracellular end of CB1 (Site 2 in Figure 2) (Stornaiuolo et al., 2015). This prediction was supported by the fact that a synthesized covalent analog of ORG27569 was determined to bind to a peptide mass that corresponded to the GSL tripeptide sequence, which is only present at the intracellular end of the receptor, at Site 2.
A third study, by Sabatucci et al. (2018), applying computational methods only, determined that ORG27569 could be docked into pockets at either the intracellular or extracellular end in the inactive-state crystal structure (STGZ) of the CB1 receptor (Sabatucci et al., 2018). However, the docking studies revealed that the binding pocket on the extracellular end (Site 3 in Figure 2, which is not the same site described by Shore et al. (2014)) was preferred.

In this work, we applied molecular modeling techniques to predict possible binding sites for ORG27569 within recently published X-ray crystal structures of the CB1 receptor to help resolve the conflicts between previous reports. We allowed protein flexibility in the early stages of the computational protocol, including in the allosteric site identification, and then performed ligand docking to identify the best pocket for ORG27569. We evaluated the results from our computational protocol through comparison with the recently released X-ray crystal structure for the CB1–CP55940–ORG27569 complex. We also performed extensive molecular dynamics simulations to test the robustness of our predictions obtained from the docking.

**Results and discussion**

**Generation of protein 3D structures**

We built four models of the CB1 receptor, one based on each of the first four CB1 X-ray crystal structures available (PDB IDs: 5U09, STGZ, 5XRA and 5XR8), none of which contained an allosteric modulator. The X-ray structures were first refined by the inclusion of additional CB1 residues at the N- and C-termini, as well as for intracellular loop 3 (ICL3), in order to include features of the CB1 receptor that have been reported to influence ligand binding in the native 3D structures (Sabatucci et al., 2018). This was done in order to provide a sufficient linker between TMHs 5 and 6 while eliminating the challenges of predicting the positioning of the full loop. The C-terminal of the CB1 receptor was extended far enough to include C415<sup>ter</sup>, since palmitoylation of that residue has been reported to influence receptor internalization (Oddi et al., 2017). Finally, all mutations used to aid crystallization were reverted back to the residues found in the wild type receptor.

**Ligand docking within the orthosteric site**

Five known orthosteric ligands of the CB1 receptor (taranabant, AM6538, AM11542, AM841 and CP55940) were docked into each of the four receptor models, with the receptor site defined by the centroid of the co-crystallized ligand for each X-ray crystal structure using Glide, Schrödinger (Glide, 2018). Following docking, ligands that were docked within their native protein 3D structures (self-docking) were observed to possess the best GlideScores, with the exception of the 5XRA model in which the agonist AM841 had the best score. The self-docked ligands were observed to adopt docking poses highly similar to the binding poses found within the X-ray crystal structures, as shown by the low RMSDs (Table 1). We also observed that ligands docked better within protein models whose conformational state matched the ligands' CB1 functional activities. Thus, CB1 antagonists/inverse agonists (taranabant and AM6538) docked better within the inactive-state structures, 5U09 and 5TGZ (Hua et al., 2016; Shao et al., 2016), while CB1 agonists (AM11542, AM841 and CP55940) docked better within the active-state structures, 5XRA and 5XR8 (Hua et al., 2017). These results validated our docking protocol and revealed that the docking protocol was able to distinguish between ligands that stabilize different activation states of the protein.

| Ligand       | Ki (CB1) | SU09  | STGZ   | SXRA   | 5XR8  |
|--------------|----------|-------|--------|--------|-------|
| Taranabant   | 0.13 nM  | −13.315 (0.56<sup>a</sup>) | −9.335 | −9.310 | −9.374 |
| AM6538       | 3.4 nM   | −10.697 | −10.224 (1.26<sup>b</sup>) | NA<sup>b</sup> | NA<sup>b</sup> |
| AM11542      | 0.11 nM  | −9.392  | −7.691  | −12.345 (0.34<sup>b</sup>) | −12.046 |
| AM841        | 1.14 nM  | −9.961  | −8.428  | −13.293 | −12.64 (0.63<sup>b</sup>) |
| CP55940      | 0.6 nM   | −9.309  | −7.909  | −11.206 | −11.477 |

<sup>a</sup>Root mean square deviation (RMSD) (Å) of heavy atoms of the native ligand compared with the docked pose.

<sup>b</sup>NA: No docking poses obtained for ligand.

<sup>c</sup>Study by Thakur et al. (2005).
The binding of ORG27569, a cannabinoid agonist, to the CB1 receptor in its active and inactive states was studied. The computational docking and SiteMap analysis were performed on the four rigid complexes of the CB1 receptor bound to CP55940. SiteMap analysis revealed multiple probable allosteric sites located in various parts of the receptor. The binding site scores (SiteScore) and druggability scores (Dscore) were used to rank the allosteric sites, with a few exceptions.

K192 in CB1 signaling and served as a basis to set a constraint to maintain hydrogen bond interaction between this residue and ORG27569 in their computational studies. However, in all experimental 3D structures of the CB1 receptor obtained by X-ray diffraction and cryo-EM (PDB ID: 6N4B), the orientation of the side chain of K192 does not favor the presence of a hydrogen bond with the bound orthosteric ligands. This lends credence to the argument that K192 has an indirect effect on ligand binding by modifying the orthosteric binding site rather than through binding directly to ligands.

In the docking we did, CP55940 adopted a similar pose to that of the agonists co-crystallized in the active-state CB1 structures (Figure 3), considering it necessary to reassess the docking pose of CP55940 within the CB1 orthosteric site prior to the identification of allosteric sites within the CB1–CP55940 complex.

When docked in either the active-state or inactive-state models, the hydrophobic alkyl tail of CP55940 adopted a similar orientation, lying almost perpendicular to the transmembrane (TM) bundle and facing TMH5. The dimethyl group was observed to have receptor activation state-dependent hydrophobic contacts with the rotameric toggle switch residues, F200 and W356. In the active-state models, hydrophobic interactions were found with F200 and W356, but not with W356. However, in the inactive-state models, hydrophobic contacts were found with W356, but not with F200. In the inactive-state models, the phenyl and cyclohexyl rings dock lower within the orthosteric binding pocket towards TMHs 1, 2 and 7. Consequently, the π–π interaction with F268 found in the active-state complex was not observed. The larger volume of the orthosteric binding pocket in the inactive state, compared to in the active state, may account for the ability of CP55940 to adopt these alternate docking poses.

**Figure 3.** Structures of bound ligands found within the active-state CB1 structures (respective PDB structures are indicated) and of CP55940, the cannabinoid agonist used in this study (SAH: Southern Aliphatic Hydroxyl; NAH: Northern Aliphatic Hydroxyl).
For each of the four models, the top-ranked sites had very similar scores.

A significant number of the allosteric sites identified by SiteMap were found at the surface of the lower half (intracellular side) of the CB1 receptor as a result of clefts formed by the position of the helices within the transmembrane bundle. Amino acid sequences that are known to characterize cholesterol binding to proteins include the cholesterol consensus motif (CCM), cholesterol recognition/interaction amino acid consensus (CRAC) and cholesterol recognition motif (CARC) (Epand, 2006; Fantini & Barrantes, 2013; Hanson et al., 2008). Cholesterol is known to be an integral component of the membrane and can also act as an endogenous allosteric modulator of CB1, modifying the affinity of CB1 inverse agonists and influencing the distribution of CB1 in the plasma membrane (Stornaiuolo et al., 2015). The CB1 receptor does

Figure 4. Docked poses of CP55940 (C magenta) superimposed with the native ligands within X-ray crystal structures of the CB1 receptor for the inactive-state (A) SU09 and (B) STGZ; and active-state (C) SXR A and (D) SXR8. Multiple views of these poses are shown in Figure S1 (Supplementary material).

Figure 5. Maps of allosteric sites predicted using SiteMap within CP55940-bound complexes of the CB1 receptor, based on X-ray structures of the inactive-state (A) SU09 and (B) STGZ; and active-state (C) SXR A and (D) SXR8. CP55940 (C magenta; ball-and-stick model) is shown within the orthosteric site. Maps of particular sites are shown with the same color across different models if there is a significant overlap between them.
Table 2. Binding site scores (SiteScore) and druggability scores (Dscore) of allosteric sites predicted by SiteMap within the CB1–CP55940 complexes derived from four available X-ray structures (indicated by PDB code).  

| Site | SiteScore | Dscore | Site | SiteScore | Dscore | Site | SiteScore | Dscore | Site | SiteScore | Dscore |
|------|-----------|--------|------|-----------|--------|------|-----------|--------|------|-----------|--------|
| S2   | 1.110     | 1.232  | S151 | 1.146     | 1.235  | S151 | 1.055     | 1.101  | S33   | 1.178     | 1.283  |
| S1   | 1.001     | 1.006  | S252 | 1.058     | 1.092  | S252 | 1.006     | 1.075  | S151  | 1.039     | 1.080  |
| S3   | 0.996     | 1.063  | S353 | 0.884     | 0.939  | S454 | 0.947     | 1.033  | S454  | 0.981     | 1.053  |
| S7   | 0.829     | 0.599  | S454 | 0.834     | 0.884  | S858 | 0.839     | 0.271  | S858  | 0.948     | 0.986  |
| S8   | 0.822     | 0.866  | S555 | 0.834     | 0.876  | S666 | 0.876     | 0.949  | S666  | 0.876     | 0.949  |
| S5   | 0.784     | 0.808  | S555 | 0.824     | 0.888  | S555 | 0.781     | 0.796  | S555  | 0.781     | 0.796  |
| S6   | 0.777     | 0.785  | S555 | 0.806     | 0.846  | S757 | 0.749     | 0.551  | S757  | 0.564     | 0.493  |
| S4   | 0.732     | 0.732  | S666 | 0.749     | 0.748  |       |           |        |       |           |        |
| S9   | 0.513     | 0.454  | S1051| 0.697     | 0.748  | S1051| 0.697     | 0.748  |       |           |        |
| S7   |           |        | S757 | 0.564     | 0.493  |       |           |        |       |           |        |

Table 3. GlideScores and binding free energies for ORG27569 within the CB1–CP55940 complexes.  

| PDB Code | GlideScore (kcal/mol) | Binding Free Energy (kcal/mol) | Site | Site Volume (Å³) |
|----------|-----------------------|-------------------------------|------|------------------|
| SU09     | -5.579                | -47.13                        | S8   | 79.233           |
| STGZ     | -6.611                | -60.29                        | S2   | 395.136          |
| 5XRA     | -5.878                | -66.39                        | S1   | 406.455          |
| 5XR8     | -5.498                | -48.02                        | S1   | 151.949          |

SiteMap analysis following molecular dynamic (MD) simulations  

An X-ray crystal structure presents only one conformation of the protein, in a state that is well adapted for binding to the ligand with which it is co-crystallized. Thus, computational studies that use a rigid protein model including one ligand to identify the binding site for a different ligand may provide inaccurate information. To account for dynamic protein flexibility for the geometry of allosteric binding sites, a 50 ns MD simulation was performed on a CP55940-bound inactive-state CB1 model, using Desmond, Schrödinger. An RMSD plot of the protein and ligand is provided in Figure S2 (Supplementary material). The starting CB1 model for MD simulations was built from the CP55940-bound SU09 protein model. An inactive-state structure was selected for MD simulation since the functional activity of ORG27569 (an intrinsic inverse agonist and negative allosteric modulator) suggests that the molecule will drive the CB1 conformational ensemble towards the inactive state, even when CB1 is bound to the agonist, CP55940. In addition, we reasoned that a short MD simulation would be adequate to perturb the shape and size of potential allosteric sites within the complex, but insufficient to induce a transition of the CB1 receptor from inactive-state to active-state.

SiteMap analysis was performed on 101 frames extracted from the MD trajectory. The script used for this bulk analysis ignored the presence of CP55940 and hence could choose one or more sites that overlapped with the orthosteric site. The sites created from each of the extracted frames were grouped based on similarity, to yield site clusters. Of the 30 top-ranked site clusters obtained (Table 4), we observed that over half of them were present in only one frame, suggesting there is a low probability of finding these sites in the receptor conformational ensemble. Three site clusters, SC1, SC3 and SC6, were found in virtually all the frames analyzed. This suggests a very high probability of finding these binding domains within the receptor conformational ensemble. Three other site clusters (SC5, SC8 and SC13) were found in...
approximately half of the frames analyzed. These site clusters may indicate cryptic binding domains – binding sites that are formed spontaneously and are stabilized by induction as an allosteric modulator binds. The formation of a cryptic pocket has been proposed to be responsible for the receptor subtype selectivity exhibited by allosteric modulators at the muscarinic receptors (Hollingsworth et al., 2019). This shows the value of performing the simulation and using multiple snapshots to assess the durability of binding sites. For further analysis, we considered it more reliable to use sites that were found in multiple snapshots.

We performed InducedFit docking (IFD) for ORG27569 into CB1–CP55940 within all site clusters that were present in more than 40 trajectory frames (Table 5, Figure 7 and Supplementary material Figure S3). The IFD protocol was used to allow for protein flexibility at the predicted ORG27569 docking sites (Sherman et al., 2006). A docking site for IFD was defined by all amino acid residues that made up a particular site cluster (cf. Methods section). We also performed binding free energy calculations on the complexes obtained from the six main site clusters. ORG27569 had similar binding free energies (average ≈ −65.5 kcal/mol) at three site clusters, SC5, SC8 and SC13, suggesting that NAM binding at these sites may confer the same degree of stabilization to the complex. The other three site clusters, SC1, SC3 and SC6, which appeared consistently throughout

Table 4. Performance of site clusters obtained in the CB1–CP55940 complex following MD simulations.

| Site Cluster | Average SiteScore | Average Dscore | Average Site Volume (Å³) | Number of Site/Frame Pairs in Cluster
|--------------|------------------|----------------|--------------------------|----------------------------------------|
| 1            | 1.19             | 1.26           | 529.29                   | 101                                    |
| 2            | 1.07             | 1.11           | 55.91                    | 1                                      |
| 3            | 1.00             | 1.04           | 280.75                   | 214                                    |
| 4            | 0.94             | 0.91           | 57.62                    | 1                                      |
| 5            | 0.87             | 0.90           | 106.11                   | 57                                     |
| 6            | 0.85             | 0.91           | 119.54                   | 100                                    |
| 7            | 0.84             | 0.90           | 94.87                    | 5                                      |
| 8            | 0.82             | 0.86           | 82.84                    | 43                                     |
| 9            | 0.80             | 0.86           | 74.50                    | 16                                     |
| 10           | 0.75             | 0.43           | 36.70                    | 1                                      |
| 11           | 0.73             | 0.71           | 96.04                    | 1                                      |
| 12           | 0.71             | 0.68           | 203.74                   | 1                                      |
| 13           | 0.71             | 0.65           | 101.31                   | 46                                     |
| 14           | 0.69             | 0.33           | 91.24                    | 1                                      |
| 15           | 0.69             | 0.69           | 36.70                    | 1                                      |
| 16           | 0.66             | 0.61           | 86.09                    | 2                                      |
| 17           | 0.65             | 0.47           | 110.79                   | 1                                      |
| 18           | 0.63             | 0.45           | 56.59                    | 1                                      |
| 19           | 0.63             | 0.53           | 89.87                    | 1                                      |
| 20           | 0.63             | 0.50           | 49.05                    | 1                                      |
| 21           | 0.63             | 0.55           | 117.31                   | 1                                      |
| 22           | 0.63             | 0.59           | 100.16                   | 1                                      |
| 23           | 0.62             | 0.50           | 87.35                    | 3                                      |
| 24           | 0.62             | 0.43           | 66.64                    | 22                                     |
| 25           | 0.59             | 0.46           | 67.57                    | 1                                      |
| 26           | 0.58             | 0.25           | 32.24                    | 1                                      |
| 27           | 0.57             | 0.42           | 50.76                    | 2                                      |
| 28           | 0.56             | 0.43           | 52.48                    | 1                                      |
| 29           | 0.56             | 0.39           | 50.42                    | 1                                      |
| 30           | 0.54             | 0.36           | 34.81                    | 8                                      |

*More than one site makes up the site cluster in certain frames.
extracellular end to avoid clashing with the co-bound modulator binding. Of these, only site clusters SC1 and SC3 either in CB1 or in other class A GPCRs as sites for allosteric site clusters SC1, were present in all the frames examined and possessed average SiteScores > 1 and Dscores > 1 (Table 4).

Site Cluster 1 (SC1). SC1 largely overlapped with the CP55940 binding site, extending to the extracellular end of the receptor. Because the bulk processing of sites ignored the presence of the orthosteric ligand CP55940 in the MD simulation, SC1 is extremely large and overlaps both with the orthosteric binding pocket and with the allosteric binding pockets such as S2 found in rigid 5TGZ (shown in Figure 5 and Table 2). For comparison, when we applied SiteMap to individual snapshots from the MD simulation then the presence of CP55940 was recognized and the binding sites detected close to the extracellular side matched very well with sites found from the pre-MD models (those shown in Figure 5 and Table 2).

Table 5. Performance of ORG27569 docking within the main site clusters generated from the CB1–CP55940 complex.

| Site Cluster | GlideScorea | IFD Scorea | Binding Free Energya | Number of Site/Frame Pairs in Clusterb |
|--------------|-------------|------------|----------------------|---------------------------------------|
| 11           | -6.190      | -549.22    | -50.00               | 101                                   |
| 33           | -9.740      | -556.08    | -87.41               | 214                                   |
| 55           | -6.542      | -489.49    | -66.50               | 57                                    |
| 66           | -8.439      | -495.44    | -76.66               | 100                                   |
| 88           | -7.862      | -495.47    | -65.08               | 43                                    |
| 1313         | -6.862      | -493.61    | -64.92               | 46                                    |

aKcal/mol.

bMore than one site makes up the site cluster in certain frames.

The second site cluster present in all the frames, SC3, was located towards the intracellular end of the receptor, largely within the transmembrane bundle. A part of this site cluster was also outside the TMH bundle in a cleft formed by TMHs 3, 4 and 5 in a region where cholesterol is bound in the 6N4B active-state CB1 structure. The ternary complex obtained with ORG27569 at SC3 was more energetically favored (~87.41 kcal/mol) compared to all the other main complexes (Table 5). The best docking pose for ORG27569 was near TMHs 2, 6, 7 and 8 (Supplementary material Figure S4). In this docking pose, ORG27569 was near TMHs 2, 6, 7 and 8 (Supplementary material Figure S4).

No output poses were obtained when ORG27569 was docked into the receptor grid defined by the residues that make up SC1. We applied an offset of 6 Å to move the center of the receptor grid site in the +Z-direction towards the extracellular end to avoid clashing with the co-bound CP55940 (Figure 7 and Supplementary material Figure S3).

Using this setting, we found that ORG27569 docked further towards the extracellular end of the site cluster within the CP55940-bound CB1 complex. Thus, while the SiteMap script used to identify SC1 included space that overlapped with the orthosteric binding pocket, we applied a more restrictive definition of SC1 in our docking protocol, such that there was no overlap between CP55940 and the ORG27569 docking sites. Note that it has been reported that a positive allosteric modulator binds at a site similar to this CB1 SC1 site in the muscarinic acetylcholine receptor (PDB ID: 4MQT) (Kruse et al., 2013). As mentioned in the introduction, sites similar to SC1 were predicted previously as the favorable spot for ORG27569 docking to CB1–CP55940 (Sabatucci et al., 2018; Shore et al., 2014). In addition, it has been proposed that cannabidiol, another CB1 NAM, binds at this extracellular allosteric site (Tham et al., 2019). From our calculations, we observed that SC1 had the worst binding free energies (~50 kcal/mol) of all six site clusters examined. This suggests that this binding site is least likely to form a stable CB1–CP55940–ORG27569 complex.

Site Cluster 3 (SC3). The second site cluster present in all the frames, SC3, was located towards the intracellular end of the receptor, largely within the transmembrane bundle. A part of this site cluster was also outside the TMH bundle in a cleft formed by TMHs 3, 4 and 5 in a region where cholesterol is bound in the 6N4B active-state CB1 structure. The ternary complex obtained with ORG27569 at SC3 was more energetically favored (~87.41 kcal/mol) compared to all the other main complexes (Table 5). The best docking pose for ORG27569 was near TMHs 2, 6, 7 and 8 (Supplementary material Figure S4). In this docking pose, ORG27569 was found to make two hydrogen bond interactions. The ORG27569 indole moiety served as a hydrogen bond donor to D4038.49, while the amide linkage served as a hydrogen bond acceptor to Y1532.40. ORG27569 was also found to make hydrophobic interactions with the canonical ionic lock residues, R2143.50 and D3386.30. The proximity of ORG27569 to the ionic lock could contribute to disruption of receptor signaling when CP55940 is also bound to CB1. We note that while the ORG27569 binding site in the post-MD inactive-state model matches the intracellular binding domain predicted using pre-MD active-state models (S1 in Figure 5C and D), ORG27569 had no interactions with TMH4 in the post-MD binding site but it did have such interactions in the pre-MD active-state predictions.
Some other class A GPCR allosteric antagonists have been determined to bind at a similar site as our predicted ORG27569 binding site. The beta2 adrenergic receptor (β2AR), CC chemokine receptor type 2 (CCR2) and CC chemokine receptor type 3 (CCR3) have each been crystallized with intracellular allosteric antagonists near TMHs 1, 2, 6 and 7 and helix 8 (Liu et al., 2017; Oswald et al., 2016; Zheng et al., 2016). The binding sites of the CCR2 and CCR9 NAMs, CCR2-RA-[R] and vercirron, are slightly displaced further upwards in the TM bundle compared to the binding site of Cmpd-15PA (β2AR NAM) in the β2AR and our ORG27569 docking site in the CB1. While this offset causes CCR2-RA-[R] and vercirron to bind higher than Cmpd-15PA and ORG27569 within the TMH bundle, some common protein–ligand interactions were observed. A hydrogen bond that involves the residue at position 8.49 (Ballesteros–Weinstein numbering) – D403<sup>B.49</sup>, D331<sup>B.49</sup>, K311<sup>B.49</sup> and R323<sup>B.49</sup> – was found for the allosteric modulators in CB1, β2AR, CCR2 and CCR9, respectively. Of these, the SC3 docking site of ORG27569 found in our work was most similar in location to the binding site for Cmpd-15PA, the allosteric antagonist co-crystallized with ORG27569 docking predictions. The Stornaiuolo et al. pose featured the ORG27569 phenylpiperidine end lying against TMH2, while its indole-2-carboxamide moiety pointed towards TMH3 and 4.

Site Cluster 6 (SC6). SC6 was located outside the TMH bundle toward the cytoplasmic end of the receptor, within a cleft formed by TMHs 1–4 and the first intracellular loop (ICL1). SC6 was prevalent in the frames obtained from our MD simulation and, our docking results within SC6 are the second best (after SC3) for ORG27569 using the GlideScore and binding free energies (Table 5). A recent study by Hurst et al., in which a force-based Metropolis Monte Carlo simulated annealing program was applied to ligand fragments, proposed that GAT228, a CB1 allosteric agonist, binds at this site (Garai et al., 2020; Hurst et al., 2019). The pharmacology profiles of both ligands, ORG27569 and GAT228, reveal that they are not ‘clean’ allosteric modulators; both exhibit some intrinsic activity in the absence of an orthosteric ligand (Ahn et al., 2013; Laprairie et al., 2016, 2017). This may indicate that CB1 residues at SC6 can alter CB1 signaling through intrinsic and allosteric control mechanisms upon ligand binding. We aligned an inactive-state model and an active-state model with the new experimental structure (PDB ID: 6KQI) and show overlays of these structures in Figure S6 (Supplementary material). Comparing our docked pose of ORG27569 within SC6 with its experimental pose in the X-ray crystal structure of CB1 bound to CP55940 and ORG27569 (PDB ID: 6KQI) (Shao et al., 2019), the indole-2-carboxamide moieties matched well in their interactions with TMH2 residues, but the phenylpiperidine ends pointed in different directions: in our docked pose the phenylpiperidine pointed toward TMH2 whereas in 6KQI it stretched over to interact with TMH4 residues (Supplementary material Figure S6C). Of the three prior studies mentioned above that predicted the ORG27569 binding site, the ORG27569 binding site revealed in 6KQI agreed best with the second study by Stornaiuolo et al., but not with the studies by Shore et al. or Sabatucci et al. (Sabatucci et al., 2018; Shore et al., 2014; Stornaiuolo et al., 2015). However, the Stornaiuolo et al. binding pose was significantly different from the pose found in 6KQI or in our docking. The Stornaiuolo et al. pose featured the ORG27569 phenylpiperidine end lying against TMH2, while its indole-2-carboxamide pointed towards TMHs 3 and 4. Similar to prior X-ray crystal structures of CB1, 6KQI was crystallized with a fusion protein in lieu of ICL3, with residues H302 – P332 missing. Five mutations (S203K, T210A, E273K, T283V and R340E) were used to make the 6KQI construct, which had an overall melting temperature of >85°C. The resolution of the resultant structure was 3.245 Å. Ligand interactions between CP55940 and CB1 in the 6KQI structure are similar to the interactions we found for CP55940 in the active-state CB1 receptor (Supplementary material Figure S6A), with the exception of the rotameric toggle switch interaction. Similar to our previous docking result in the inactive-state model, CP55940 in 6KQI was found to make hydrophobic contacts with W356<sup>E.48</sup>, but not with F200<sup>E.36</sup>. Similar to our docking results, we found that CP55940 in 6KQI has π–π interactions with F170<sup>E.57</sup> and F268<sup>E.42</sup> and a hydrogen bond interaction between its phenolic hydroxyl group and S383<sup>E.39</sup>. Two additional hydrogen bond interactions observed in the 6KQI structure are between the northern aliphatic hydroxyl (NAH) of CP55940 and the backbone of I267<sup>E.42</sup>, and between the southern aliphatic hydroxyl (SAH) and S173<sup>E.60</sup>. These additional interactions correlate to the increased binding affinity CP55940 has for the ORG27569-bound CB1 relative to the unbound CB1. Also, no hydrogen bond interaction was found between CP55940 and K192<sup>E.28</sup> in the 6KQI structure. The intermediate-state CB1 structure (6KQI) revealed that ORG27569 binds on the membrane interface of CB1 across TMHs 2, 3 and 4. The location of its indole-2-carboxamide moiety overlaps with our previously identified site, SC6 (Figure 7 and Supplementary material Figure S6B).

**Time-dependent stability analysis of CB1CP55940ORG27569 ternary complexes**

We performed two independent MD simulation runs with Desmond, Schrödinger, each lasting 200 ns, of each of three
complexes to assess their stability: the recent crystal structure (PDB ID: 6KQI) and our two top predicted complexes (SC3 and SC6). Binding free energy calculations were performed on frames extracted at an interval of 4 ns from each trajectory. To account for differences related to type of solvent molecule found in the allosteric pockets, three approaches were applied in the preparation of structure files for Molecular Mechanics Generalized Born/Surface Area (MMGBSA) binding free energy approximations (see section on Binding Free Energy Calculations in Materials and Methods). In the first approach, MMGBSA calculations were performed on complexes that had been stripped of all ions, water and membrane molecules. Using this approach, in which the outer dielectric constant at all binding pockets was set to 80, ORG27569 had the best binding free energy when bound in SC3. Variations in the binding free energies calculated in this approach were found to match well with the variations in the binding pose of the ligand that occurred during the simulations (Supplementary material Figure S7A).

In the second approach, the ORG27569 binding sites in 6KQI and SC6 were designated to be low dielectric regions, using the Prime implicit membrane model, in contrast to the SC3 pocket, which was maintained as having a high dielectric constant to represent its exposure to water. For this approach, the ORG27569 binding free energy approximations were found to follow a similar pattern to that of the first approach during the MD simulations (Figure 8). The average ORG27569 binding free energy at SC3 remained the same, while there was an increase (∼2 kcal/mol) in the average ORG27569 binding free energy in either 6KQI or SC6, accounting for a reduced penalty for the loss of hydrophilic interactions in these complexes (Supplementary material Table S2). Using this second approach, a breakdown by residue of the contributions to the free energy was calculated and is shown in Figure 9 for CB1 residues that contributed >10% of the average binding free energy of ORG27569.

In the third approach, the ORG27569 binding free energy was set as the sum of the energetic contributions of CB1, CP55940 and ORG27569 from complexes in which all ions, water and membrane molecules within 7 Å of ORG27569 were retained prior to the binding free energy calculations (Supplementary material Figure S7B). The binding free energy results obtained from this approach did not appear to correlate well with changes in the pose of ORG27569 as observed in the first two approaches (cf. Figure 8 and Supplementary material Figure S7). In all three approaches, on average we found that ORG27569 was energetically favored to bind at SC3 rather than at SC6. The use of the Prime implicit membrane model was considered to provide the more accurate characterization of the dielectrics at the different allosteric sites. Thus, binding free energy approximations reported in the rest of this article were obtained using the second approach.

The 6KQI structure was submitted for MD simulations with ICL3 modeled in as described for the other CB1 models. The protein and ORG27569 were stable during the simulation in both independent runs, as shown from the RMSD plots (Supplementary material Figure S8A). ORG27569 retained the extended conformation observed in the X-ray structure which we used as the starting structure. ORG27569 was held in place largely by non-specific hydrophobic interactions with residues H1542–41, F2374–46, C2384–47, W2414–50 and I2454–54 (Figure 9A). While the indole-2-carboxamide moiety showed very few minor fluctuations, the phenyl piperidine end of the ligand showed greater fluctuations, largely due to rotation of the phenyl and piperidine rings. While the general orientation of ORG27569 was the same across both independent runs, we observed that the average position of ORG27569 in the simulations relative to its position in the 6KQI X-ray crystal structure was further away from the CB1 intracellular loop 1 (ICL1) in Run 1, and closer to ICL1 in Run 2. This slight displacement in the binding position of ORG27569 accounts for the difference of ∼16 kcal/mol observed between the average binding free energies in the two trajectories. The steady average binding free energy of ORG27569, which was −48.26 ± 4.05 kcal/mol in Run 1 and −64.58 ± 4.08 kcal/mol in Run 2 (also cf. Figure 8), suggests the high stability of the ligand pose at this extrahelical site.

Our model of ORG27569 docked into the CP55940-bound 5U09-based CB1 complex within SC6, the binding site predicted to be second best from the GlideScore and binding free energies (Table 5), was also submitted for two independent 200 ns MD simulation runs (Supplementary material Figure S8B). In Run 1, we observed that ORG27569 adopted three distinct binding orientations at SC6. At the start of the simulation, ORG27569 adopted a highly similar pose and similar protein–ligand interactions to those observed for ORG27569 in the MD simulations of 6KQI, with the exception that the indole-2-carboxamide group was rotated in place by ∼170°. In this first phase of the simulation, which lasted for the first ∼120 ns of the simulation, ORG27569 had a relatively favorable average binding free energy of −50.94 ± 6.86 kcal/mol. However, following this stage, the phenyl piperidine end of ORG27569 moved away from CB1, causing a partial loss in interactions, with a cost of ∼20 kcal/mol to the ORG27569 binding free energy. In this second phase, ORG27569 retained its extended conformation and CB1 interactions with the indole-2-carboxamide end of the ligand, including hydrophobic contacts with H1542–41, V2344–43, F2374–46, C2384–47, W2414–50 and I2454–54 (Figure 9B). In the last phase of the simulation, ORG27569 adopted an L-shape that allowed partial restoration of tight binding of ORG27569 to CB1 (average binding free energy: −44.89 ± 2.55 kcal/mol): the L-shape allowed the phenyl piperidine moiety to regain some favorable contacts with the CB1 residues W2414–50, T2424–51 and I2454–54. In the second MD run, ORG27569 adopted an L-shaped bent conformation which was relatively stable during the 200 ns simulation, yielding a good average binding free energy, with relatively little deviation from the average, −50.88 ± 3.45 kcal/mol. While this L-shaped conformation of ORG27569 itself from Run 2 was similar to the bent conformation observed at the end of Run 1, the orientation of the ligand relative to the protein in Run 2 was observed to change consistently
throughout the entire simulation, with it rotating in place – relative to the protein – in the docking site.

A third set of simulations was run, with our model of ORG27569 docked into SC3 within the CP55940-bound 5U09-based CB1 complex – the site predicted in GlideScore and binding free energy calculations to be the most energetically favorable to bind ORG27569 (Table 5 and Figure 8). Despite the relatively large size of SC3 (see Table 4), ORG27569 was observed to maintain interactions mostly with residues from TMHs 2, 6, 7 and helix 8 in both runs. In Run 1, ORG27569 was found to adopt an L-shaped bent conformation, with a binding free energy of $-53.09 \pm 6.49$ kcal/mol, throughout the simulation. At the start of the first run, ORG27569 was found to interact with residues from TMHs 2, 6 and 7, with the phenyl piperidinyl end close to TMHs 6 and 7 pointing towards the extracellular side, similar to the pose in Figure S4 (Supplementary material). However, the indole carboxamide end lay across TMH 6. This pose was stabilized by a double hydrogen bond interaction between D3386.30 of the CB1 ionic lock and both nitrogen-bound hydrogens in ORG27569. Other interactions observed included hydrophobic interactions with Y1532.40, Y224ICL2 and L3416.33, as well as water-mediated interactions with S1522.39, Y224ICL2 and Q3346.26. Comparing the two runs, the ligand RMSD and binding free energy showed that ORG27569 was more stable and possessed more favorable binding energy in Run 2 compared to in Run 1 (Figure 8 and Supplementary material Figure S8C).

Overall, in the third set of simulations (with ORG27569 at the intracellular site, SC3, using the 5U09-based CB1–CP55940 model), ORG27569 exhibited better average binding free energies compared to in the other two sets of MD simulations in which ORG27569 was at the extrahelical site (SC6 in the 5U09-based CB1–CP55940 model and the ORG27569 binding site in 6KQI). This supports the prediction that the SC3 site serves as an energetically favored binding site for ORG27569 within the ternary complex. It is worth noting that based on the ligand RMSD plots, ORG27569 fluctuations at the SC3 site are much greater than those observed at the extrahelical site (Supplementary material Figure S8). This can be attributed to the larger volume of the SC3 site compared to the extrahelical SC6 site. The higher fluctuations found when it is in SC3 suggest that ORG27569, a relatively non-bulky ligand, is less stably-bound at this site than at the extrahelical site, which is smaller and more solvent-exposed. Thus, we predict that the SC3 site is an alternate binding site for ORG27569. It is important to mention that longer MD simulations are required to fully characterize the CB1 functional effect that ORG27569 will elicit when it binds to the SC3 allosteric site. With the overall position of ORG27569 relative to CB1 being very similar in the two simulations, differences in the conformation of the ligand itself explain much of the variations in the ligand binding free energies between the pairs of simulation. Two main stable conformations were found for ORG27569 during the MD simulations – a straight, extended conformation and a bent,
Figure 9. Per residue energy (kcal/mol) breakdown for CB1 residues with at least 10% contribution to the average ORG27569 binding free energy within the CB1–CP55940–ORG27569 complex based on (A) the 6KQI X-ray crystal structure, (B) our 5U09-based CB1 model with ORG27569 in SC6, and (C) our 5U09-based CB1 model with ORG27569 in SC3.
L-shaped conformation. These two conformations differ primarily due to variations in a torsion angle located in the middle of the ethylene linker in ORG27569. We observe that while both conformations are stable, the straight, extended conformation affords more favorable interactions with CB1 in the different sets of MD simulations. These results show the utility in running multiple independent runs of an MD simulation in analyzing macromolecular interactions.

Materials and methods
Generation of protein 3D structure and ligand preparation

The inactive-state crystal structures of the CB1 receptor, 5TGZ and 5U09 (Hua et al., 2016; Shao et al., 2016), were used as templates to build two inactive-state homology models. Similarly, two agonist-bound crystal structures of the CB1 receptor, 5XR8, and 5XRA, were used to build two active-state homology models. For each of the four models, residues Q97–E416 of the CB1 receptor were generated using the energy-based building method homology modeling tool in Prime, Schrödinger (Prime, 2018). Insertions (for a series of residues found in the target but not in the template) were built and refined based on optimal energy. Template rotamers were retained for conserved residues and the side chains of target residues in insertions were energy-minimized. The modeling was performed such that all ligands co-crystallized with the receptor, including cholesterol if present, were also included in the homology models constructed. Superscripts based on the Ballesteros–Weinstein numbering scheme have been used to supplement the conventional amino acid sequence numbering in most of this article (Ballesteros & Weinstein, 1995).

In all four CB1 crystal structures, the N-terminus is truncated and only begins at E100, G99, D104 and F102 for 5U09, 5TGZ, 5XRA and 5XR8, respectively (Hua et al., 2016, 2017; Shao et al., 2016). We set each model to begin at Q97, so that a disulfide bridge could be built between C98 and C107. Similarly, additional CB1 C-terminus residues, beyond the final residue present in the crystal structures – F412, F412, S414 and S414 for 5U09, 5TGZ, 5XRA and 5XR8, respectively – were built into our homology models, because palmitoylation of C415 has been shown to be important for the functioning of the protein. In addition, in each crystal structure, ICL3 (a particularly lengthy loop with >30 residues) was replaced with a fusion protein to enhance crystallization (Hua et al., 2016, 2017; Shao et al., 2016). During the construction of our homology models, all of those mutations were reverted back to the wild-type CB1 residues. Following the initial homology model building, loops that were not built based on a template were further refined using the Prime loop refinement tool. During the loop refinement, portions of the N-terminus present in the templates were constrained to their initial coordinates with a force constant of 10 kcal/(mol·Å²). All ligands used in the study were prepared at pH 7.4 in LigPrep, Schrödinger (LigPrep, 2018).

Ligand docking within the orthosteric binding site

The centroid of the co-crystallized ligand was used to define the site center for receptor grid generation. Next, five molecules – CP55940 and the four co-crystallized ligands, AM6538 (PDB ID: 5TGZ), tazarotenate (PDB ID: 5U09), AM11542 (PDB ID: 5XRA) and AM841 (PDB ID: 5XR8) – were each docked into the orthosteric site of each of the CB1 receptor models, using rigid receptor docking. The Standard Precision (SP) protocol in Glide, Schrödinger (Glide, 2018) was applied for the docking. For each ligand docked, the structure with the best Glide Emodel score was selected as the best docking pose.

Identification of non-orthosteric binding sites within the models

Sitemap, Schrödinger (Halgren, 2009) was used to identify high-ranking potential binding sites on each CB1–CP55940 receptor complex. The sites obtained were then used to generate receptor grids. ORG27569 was docked into each identified site using the virtual screening workflow in Glide, as follows. Docking was performed first using the Glide Standard Precision mode and subsequently in the Extra Precision mode. Post-docking minimization was performed, after which the binding free energy of the complex was calculated. The CB1–CP55940 complex that gave the best docking score for ORG27569 was obtained as output for each model.

Identification of potential binding sites within the receptor conformational ensemble

MD Simulations

To further assess the multiple allosteric sites obtained in the CP55940-bound inactive-state binary complex (5U09), a short classical MD simulation was performed. The starting structure was selected as the 5U09–CP55940 complex. We applied a similar MD simulation protocol to that described previously (Chaurasiya et al., 2019). Using the OPLS3 (optimized potentials for liquid simulations) force field in Desmond, Schrödinger (Schrödinger Release, 2018), the complex was embedded in a POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) bilayer and solvated with a 10 Å TIP3P water buffer. The system was neutralized using 11 chloride ions, after which 0.15 M NaCl was added to the system. The OPLS3
The force field is an enhancement of the OPLS2005 force field, which implements parameters for both the protein and ligands, and does not require separate ligand force field parameterization. To achieve greater accuracy in ligand parameterization, OPLS3 includes a large number of stretch/bend parameter types, based on DFT calculations at the B3LYP/6-31G* level. The standard truncated Fourier series form of torsional potential was used in fitting the OPLS3 torsion parameters. The van der Waals parameters for OPLS3 were derived from liquid state simulations of small organic molecules. Finally, OPLS3 uses a CM1A-BCC based charge model, which combines CramerTruhlar CM1A charges with specifically fit bond charge correction (BCC) terms. Consequently, predicted relative binding free energy calculations for ligands have shown greater accuracy relative to other force fields when comparing to experimental results (Harder et al., 2016). The system was then energy-minimized using the LBFGS method over 25,000 iterations, with a convergence threshold of 1.0 kcal/(mol·Å). The system was equilibrated using the following protocol. First, the system was simulated for 1 ns using Brownian dynamics in the NVT ensemble at 10 K with the restraint of 50 kcal/mol on solute heavy atoms. Secondly, a 300 ps simulation was run in the NVT ensemble using the Berendsen thermostat (10 K) while retaining the restraint on solute heavy atoms. Thirdly, a 300 ps simulation was run in the NPT ensemble using the Berendsen thermostat (10 K) and barostat (1 atm) while restraints were retained. Over the next 300 ps, the system was gradually heated to 300 K. A final 600 ps simulation was performed in which all restraints were removed. After that, the production run (50 ns) was performed in the NPT ensemble using a timestep of 2 fs. The Nose–Hoover chain thermostat and Martyna–Tobias–Klein barostat were used for the production run.

To assess the stability of CB1–CP55940–ORG27569 ternary complexes of the experimental structure and of our top two predicted complexes, two independent 200 ns classical MD simulation runs were performed for each. The prepared X-ray crystal structure and the docked poses for our predicted complexes were selected as the starting structures for MD simulations. Two independent runs for each of the complexes had the same starting configuration and only differed in the seed velocities applied for the production runs. MD simulations were run using the OPLS2005 force field in the Desmond program, with XSEDE resources (Towns et al., 2014). The OPLS2005 force field was developed to provide larger coverage of organic functionality than prior versions of the OPLS force fields, by refitting torsional parameters to reproduce conformational energetics at the LMP2/cc-pVTZ level. In addition, additional stretch, bend and torsion parameters were fitted to afford a larger organic functionality. Each system was built in a similar manner to that of the CB1–CP55940 binary complex described above; however, a slightly different equilibration protocol was applied. The system was first simulated using Brownian dynamics in NVT at 10 K with a 50 kcal/mol restraint on solute heavy atoms. Over the next 100 ps, at 100 K and using Brownian NPT, a 20 kcal/mol restraint was applied on solute heavy atoms and a 5 kcal/mol restraint was applied on membrane heavy atoms along the Z-axis. Over the next 200 ps, the restraints were reduced to 10 kcal/mol and 2 kcal/mol for solute heavy atoms and the membrane, respectively. Following that, the system was heated to 300 K over 200 ps in the NP/T ensemble. Next, restraints on heavy atoms of the protein backbone and ligand were gradually reduced to 5 kcal/mol and finally 0 kcal/mol in the NVT ensemble over 1 ns. Finally, a 10 ns simulation without any restraints was run in the NPT ensemble prior to the production run. The Langevin thermostat and barostat were used for the 200 ns production run.

**Binding free energy calculations.** Following MD simulations of the ternary complexes, the binding free energies of ORG27569 were computed on frames extracted from the trajectory at an interval of 4 ns using the thermal_mmgbsa script by Schrödinger. Using this script, hetgroups (ions, water and membrane atoms) in the frames were removed prior to the binding free energy calculation on the ternary complex. The binding free energy calculation applied the Molecular Mechanics Generalized Born/Surface Area (MMGBSA) model with a continuum solvation model for water with a dielectric constant set to 80. To assess the effects of solvation at different site clusters, we repeated the binding free energy calculations using two alternate approaches.

In the second approach, hetgroups were removed from the frames and replaced with the Prime implicit membrane model. The Prime implicit membrane defines a region with a low dielectric constant. Using the Prime implicit membrane, the outer dielectric within the region defined by the implicit membrane was set to be 1. The position and thickness of the Prime implicit membrane for each complex was determined according to the structure in the Orientations of Proteins in Membranes (OPM) database (Lomize et al., 2006) for the CB1 model used. For complexes based on SC3, exclusion spheres were placed to cover the ORG27569 binding site, to ensure that the site was defined as a region with high dielectric region, in contrast to the SC6 site.

In the third approach, all ions and hetgroups within 7 Å of ORG27569 were retained. The ORG27569 binding free energy was computed as described using the first method above. Finally, an energy breakdown by residue was performed and the energy contributions from CP55940, ORG27569 and all CB1 residues were summed to remove the contributions of hetgroups from the ORG27569 binding free energy obtained.

**SiteMap analysis**

SiteMap analysis (binding site identification and clustering) was performed on frames extracted from the trajectory at an interval of 0.5 ns using the trajectory_binding_site_volumes script built by Schrödinger. Following the SiteMap analysis, the amino acid residues that made up the six predominant site clusters were used to generate receptor grids. The following residues were considered to be present within Site Cluster 1: Q97, C98, E100–L111, P113, Q116, L117, I119–L124, L126, T130, L159, A162, D163, G166, I169, F170, S173–H178, F180–K183, F189, L190, K192–G194, V196,
After InducedFit docking, the binding free energies of the top 20 structures using the Glide SP protocol were calculated. Next, self-docking was performed to optimize the side chains. Following the first round of docking in the IFD protocol, the receptor grid box was created from the site clusters using the InducedFit docking approach. The receptor grids used for the first round of docking were defined using a cubic grid of length 26 Å positioned at the centroid of residues that made up each site cluster. To accommodate all residues in Site Cluster 1 and Site Cluster 3, the receptor grid box was increased to 31 Å and 30 Å, respectively. Following the first round of docking, the receptor grid box was extended extracellularly, as was the case for the other site clusters. The centroid of the residues in a cluster was used to generate the receptor grid for docking. To accommodate the co-bound ligand, CP55940, within the receptor grid box of Site Cluster 1, an offset of 6 Å in the +z-direction was added to the coordinates of the receptor grid.

**Induced fit docking.** ORG27569 was docked into the receptor grids created from the site clusters using the InducedFit Docking (IFD) protocol in Schrödinger. The input structure for MD simulation was used as the starting structure for IFD. In this protocol, ORG27569 was first docked into flexible receptor sites. The receptor and ligand van der Waals (vdW) scaling factors were set to 0.5. The vDW scaling factor ranges from 0 to 1, where a scaling factor of 0 represents a completely rigid site. The receptor grids used for the first round of docking in the IFD protocol were defined using a cubic box of length 26 Å positioned at the centroid of residues that made up each site cluster. To accommodate all residues in the larger SC1 and SC3, the receptor grid box was increased to 31 Å and 30 Å, respectively. Following the first Glide docking, residues within 5 Å of the docked ligand were refined to optimize the side chains. Next, self-docking was performed for the top 20 structures using the Glide SP protocol. After InducedFit docking, the binding free energies of the complexes obtained were determined using the Prime MM-GBSA tool.

**Conclusions**

Computational prediction of ligand binding sites within proteins is challenging. The manner in which experimental information is interpreted and applied in guiding such computational studies may accidentally exclude certain possibilities in computational predictions. In our work, we followed a systematic approach based on available X-ray structures. We established a good docking pose for the orthosteric agonist CP55940. We predicted multiple possible allosteric binding pockets on 4 different X-ray structure-based CB1 models, two of the active state and two of the inactive state. We found that results with and without allowing the protein models to be flexible led to significant differences, including for the size of the pockets and for the scoring of the binding sites of ORG27569 in the CB1–CP55940 complex. Out of the six main site clusters obtained from binding site analysis following molecular dynamics simulations of the inactive-state model of CB1–CP55940, three were present in virtually all the trajectory frames. These three site clusters, SC1, SC3 and SC6, have been shown to serve as allosteric ligand-binding domains within several class A GPCRs. Of the three, SC1, which overlaps with the orthosteric binding site and extends extracellularly, was ranked the poorest, based on docking scores and binding free energy calculations from our computational studies. However, it has been found that a positive allosteric modulator binds at a site similar to the SC1 site in the muscarinic acetylcholine receptor (PDB ID: 4MQT). SC3 was predicted to be the most energetically-favorable binding domain for ORG27569 by our computational approaches. SC3 is located at the intracellular end of the inactive-state CB1 model near TMHs 2, 6 and 7 and helix 8, and matches the binding site found for some intracellular antagonists of other class A GPCRs. However, the recent X-ray crystal structure of CB1 revealed that the ORG27569 binding site overlaps with SC6, the site predicted in our computational studies – including Glide docking scores, binding free energy calculations, and extended MD simulations – as the second-best to bind ORG27569. From our analysis of 200 ns MD simulations of three ORG27569–CB1–CP55940 ternary complexes, we observed that ORG27569 binds stably in the experimental binding site revealed in 6KQI and in our predicted sites SC6 and SC3, but that the most energetically favored site is SC3. Therefore, we suggest that SC3 may serve as an alternative binding pocket for ORG27569.

**Disclosure statement**

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Funding**

This research was funded by grant numbers P20GM104932 and R15GM119061 from the National Institute of General Medical Sciences (NIGMS), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NIGMS or NIH. This investigation was conducted in part in a facility constructed with support from the Research Facilities Improvements Program (C06RR14503) from the National Institutes of Health (NIH) National Center for Research Resources. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) Bridges at the Pittsburgh Supercomputing Center using the allocation TG-MCB190033, which is supported by National Science Foundation grant number ACI-1548562.
ORCID
Robert J. Doerksen http://orcid.org/0000-0002-3789-1842

References
Adams, A. J., Banister, S. D., Irizarry, L., Trecchi, J., Schwartz, M., & Gerona, R. (2017). “Zombie” outbreak caused by the synthetic cannabinoid AMB-FUBINACA in New York. The New England Journal of Medicine, 376(3), 235–242. https://doi.org/10.1056/NEJMoa1610300

Ahn, K. H., Mahmoud, M. M., Shim, J.-Y., & Kendall, D. A. (2013). Distinct roles of β-Arrestin 1 and β-Arrestin 2 in ORG27569-induced biased signaling and internalization of the cannabinoid receptor 1 (CB1). The Journal of Biological Chemistry, 288(14), 9790–9800. https://doi.org/10.1074/jbc.M112.436804

Alaverdashvili, M., & Laprairie, R. B. (2018). The future of type 1 cannabinoid receptor allosteric ligands. Drug Metabolism Reviews, 50(1), 14–25. https://doi.org/10.1080/0302055X.2018.1428341

Al-Zoubi, R., Hurst, D., & Reggio, P. (2019). Structural insights from recent CB1 X-ray crystal structures. In W. J. Costain & R. B. Laprairie (Eds.), Recent advances in cannabinoid research. IntechOpen. https://doi.org/10.5772/intechopen.80783

Ballesteros, J. A., & Weinstein, H. (1995). Integrated methods for the construction of three-dimensional models and computational probing of structure–function relations in G protein-coupled receptors. In Stuart C. Sealfon (Ed.), Methods in neurosciences (Vol. 25, pp. 366–428). Elsevier. https://doi.org/10.1016/S1043-9471(05)80049-7

Bauer, M., Chicca, A., Tamborrini, M., Eisen, D., Lerner, R., Lutz, B., Poetz, O., Pluschke, G., & Gertsch, J. (2012). Identification and quantification of a new family of peptide endocannabinoids (PePCans) showing negative allosteric modulation at CB1 receptors. Journal of Biological Chemistry, 287(44), 36944–36967. https://doi.org/10.1074/jbc.M112.382481

Beuming, T., & Sherman, W. (2012). Current assessment of docking into GPCR crystal structures and homology models: Successes, challenges, and guidelines. Journal of Chemical Information and Modeling, 52(12), 3263–3277. https://doi.org/10.1021/ci300411b

Bian, Y., Jing, Y., Wang, L., Ma, S., Jun, J. J., & Xie, X.-Q. (2019). Prediction of orthosteric and allosteric regulations on cannabinoid receptors using supervised machine learning classifiers. Molecular Pharmaceutics, 16(6), 2605–2615. https://doi.org/10.1021/acs.molpharmaceut.9b00182

Chaurasiya, N. D., Zhao, J., Pandey, P., Doerksen, R. J., Muhammad, I., & Tekwani, B. L. (2019). Selective inhibition of human monoamine oxidase B by acacetin 7-methyl ether isolated from Damiana. ACS Chemical Neuroscience, 10(8), 1530–1539. https://doi.org/10.1021/acschemneuro.9b00182

Chen, W., Vemuri, V., Chambers, A., Limebeer, C., Bedard, H., Wood, J., Lutz, B., Zimmer, A., Parker, L., Makriyannis, A., & Sharkey, K. (2010). A novel peripherally restricted cannabinoid receptor antagonist, AM6545, reduces food intake and body weight, but does not cause malaise, in rodents. British Journal of Pharmacology, 161(3), 629–642. https://doi.org/10.1111/j.1476-5381.2010.00908.x

Congreve, M., de Graaf, C., Swain, N. A., & Tate, C. G. (2020). Impact of GPCR structures on drug discovery. Cell, 181(1), 81–91. https://doi.org/10.1016/j.cell.2020.03.003

Dhosphewshwarkar, A., Murataeva, N., Makriyannis, A., Straiker, A., & Mackie, K. (2017). Two Janus cannabinoids that are both CB2 agonists and CB1 antagonists. The Journal of Pharmacology and Experimental Therapeutics, 360(2), 300–311. https://doi.org/10.1124/jpet.116.236539

Diaz, O., Dalton, J. A. R., & Giraldo, J. (2019). Revealing the mechanism of agonist-mediated cannabinoid receptor 1 (CB1) activation and phospholipid-mediated allosteric modulation. Journal of Medicinal Chemistry, 62(11), 5638–5654. https://doi.org/10.1021/acs.jmedchem.9b00612

Dopart, R., & Kendall, D. A. (2020). Allosteric modulators restore orthosteric agonist binding to mutated CB1 receptors. The Journal of Pharmacology and Pharmacy, 72(1), 84–91. https://doi.org/10.1111/jphp.13193

Epand, R. M. (2006). Cholesterol and the interaction of proteins with membrane domains. Progress in Lipid Research, 45(4), 279–294. https://doi.org/10.1016/j.plipres.2006.02.001

Fantini, J., & Barrantes, F. (2013). How cholesterol interacts with membrane proteins: An exploration of cholesterol-binding sites including CRAC, CARC, and tilted domains. Frontiers in Physiology, 4, 31. https://doi.org/10.3389/fphys.2013.00031

Fantini, J., Di Scala, C., Baier, C. J., & Barrantes, F. J. (2016). Molecular mechanisms of protein–cholesterol interactions in plasma membranes: Functional distinction between topological (tilted) and consensus (CARC/CRAC) domains. Chemistry and Physics of Lipids, 199, 52–60. https://doi.org/10.1016/js.chemphyslip.2016.02.009

Fay, J. F., & Farrens, D. L. (2012). A key agonist-induced conformational change in the cannabinoid receptor CB1 is blocked by the allosteric ligand Org 27569. The Journal of Biological Chemistry, 287(40), 33873–33882. https://doi.org/10.1074/jbc.M112.352328

Fay, J. F., & Farrens, D. L. (2013). The membrane proximal region of the cannabinoid receptor CB1 N-terminus can allosterically modulate ligand affinity. Biochemistry, 52(46), 8286–8294. https://doi.org/10.1021/bi400842k

Gado, F., Meini, S., Bertini, S., Digiacomo, M., Macchia, M., & Manera, C. (2019). Allosteric modulators targeting cannabinoid CB1 and CB2 receptors: Implications for drug discovery. Future Medicinal Chemistry, 11(15), 2019–2037. https://doi.org/10.4155/fmc-2019-0005

Garai, S., Kulkarni, P. M., Schaffer, P. C., Leo, L. M., Brandt, A. L., Zagzoog, A., Black, T., Lin, X., Hurst, D. P., Janero, D. R., Abood, M. E., Zimmowitch, A., Straiker, A., Pertwee, R. G., Kelly, M., Szczesniak, A.-M., Denovan-Wright, E. M., Mackie, K., Hohmann, A. G., ... Thakur, G. A. (2020). 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(2), 542–568. https://doi.org/10.1021/acs.jmedchem.9b01142

Glade, C. (2018). Schrödinger Release 2018-1; LLC.

Gueye, A. B., Pryslawsky, Y., Trigo, J. M., Pouila, N., Delis, F., Antoniou, K., Loureiro, M., Laviolette, S. R., Vemuri, K., Makriyannis, A., & Le Foll, B. (2016). The CB1 neutral antagonist AM4113 retains the therapeutic efficacy of the inverse agonist rimonabant for nicotine dependence and weight loss with better psychiatric tolerability. International Journal of Neuropsychopharmacology, 19(12), pyw068. https://doi.org/10.1093/ijn/pyw068

Halgren, T. A. (2009). Identifying and characterizing binding sites and assessing druggability. Journal of Chemical Information and Modeling, 49(2), 377–389. https://doi.org/10.1021/ci800324m

Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V.-P., Chien, E. Y., Velasquez, J., Kuhn, P., & Stevens, R. C. (2008). A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. Structure, 16(6), 897–905. https://doi.org/10.1016/j.str.2008.05.001
