Application of Mixed-Solvent Molecular Dynamics Simulations for Prediction of Allosteric Sites on G Protein–Coupled Receptors[S]

Wallace K.B. Chan, Heather A. Carlson, and John R. Traynor

Department of Pharmacology and Edward F. Domino Research Center (W.K.B.C., J.R.T.) and Department of Medicinal Chemistry (H.A.C., J.R.T.), University of Michigan, Ann Arbor, Michigan

ABSTRACT

The development of small molecule allosteric modulators acting at G protein–coupled receptors (GPCRs) is becoming increasingly attractive. Such compounds have advantages over traditional drugs acting at orthosteric sites on these receptors, in particular target specificity. However, the number and locations of druggable allosteric sites within most clinically relevant GPCRs are unknown. In the present study, we describe the development and application of a mixed-solvent molecular dynamics (MixMD)-based method for the identification of allosteric sites on GPCRs. The method employs small organic probes with druglike qualities to identify druggable hotspots in multiple replicate short-timescale simulations. As proof of principle, we first applied the method retrospectively to a test set of five GPCRs (cannabinoid receptor type 1, C-C chemokine receptor type 2, M2 muscarinic receptor, P2Y purinoceptor 1, and protease-activated receptor 2) with known allosteric sites in diverse locations. This resulted in the identification of the known allosteric sites on these receptors. We then applied the method to the μ-opioid receptor. Several allosteric modulators for this receptor are known, although the binding sites for these modulators are not known. The MixMD-based method revealed several potential allosteric sites on the mu-opioid receptor. Implementation of the MixMD-based method should aid future efforts in the structure-based drug design of drugs targeting allosteric sites on GPCRs.

SIGNIFICANCE STATEMENT

Allosteric modulation of G protein–coupled receptors (GPCRs) has the potential to provide more selective drugs. However, there are limited structures of GPCRs bound to allosteric modulators, and obtaining such structures is problematic. Current computational methods utilize static structures and therefore may not identify hidden or cryptic sites. Here we describe the use of small organic probes and molecular dynamics to identify druggable allosteric hotspots on GPCRs. The results reinforce the importance of protein dynamics in allosteric site identification.

Introduction

G protein–coupled receptors (GPCRs) are a superfamily of integral membrane proteins responsible for mediating a multitude of physiologic processes. Not surprisingly, they are involved in a wide array of diseases (Römpler et al., 2007; O'Hayre et al., 2013). Consequently, GPCRs are among the most clinically relevant proteins in drug discovery, serving as the target for over 20% of drugs on the market (Chan and Zhang, 2020). The majority of drugs acting at GPCRs bind to and exert their action at the orthosteric site, the location in which endogenous ligands bind. However, modulating allosteric sites—binding pockets located away from the orthosteric site—has increasingly become a viable strategy for achieving more effective therapeutics and offers several advantages over traditional drugs. First, allosteric modulators typically can achieve greater selectivity than orthosteric ligands, as exemplified by the numerous subtype-specific allosteric modulators discovered within the metabotropic glutamate receptor subfamily (Gregory et al., 2013); this is likely due to allosteric sites being less evolutionarily conserved than orthosteric sites (Wenthur et al., 2014). Second, patient overdosing would potentially be reduced or eliminated through treatment with an allosteric drug, as allosteric modulators have no effects or very few by themselves, depend on endogenous (or exogenous) ligands for their function, and the effects reach a maximal limit (Burford et al., 2015b). Lastly, allosteric modulators can provide tissue specific targeting, whereas orthosteric ligands are more promiscuous in their binding (Kenakin, 2017). Overall, development of allosteric

[S] This article has supplemental material available at molpharm.aspetjournals.org.

ABBREVIATIONS: BPTU, N-[2-[2-(1,1-Dimethylethyl)phenoxy]-3-pyridinyl]-N'-[4-(trifluoromethoxy)phenyl]urea; CB1R, cannabinoid receptor type 1; CCR2, C-C chemokine receptor type 2; DAMGO, [D-Ala[2], N-MePhe[4], Gly-ol[5]]-enkephalin; ECL, extracellular loop; GPCR, G protein–coupled receptor; ICL3, intracellular loop 3; MixMD, mixed-solvent molecular dynamics; MOE, Molecular Operating Environment; MOR, μ-opioid receptor; M2R, M2 muscarinic receptor; NAM, negative allosteric modulator; Nb39, nanobody 39; PAM, positive allosteric modulator; PAR2, protease-activated receptor 2; PDB, Protein Data Bank; P2RY1, P2Y purinoceptor 1; RMSD, root-mean-square deviation; TM, transmembrane.

drugs offers clear benefits to identify clinically useful therapeutics.

High-throughput screening methods have proven effective in discovering novel allosteric modulators, but this method is hampered by its "brute-force" nature and being expensive, time consuming, and laborious. In response, the advent of computer-aided drug design in recent years has allowed for the prioritization of small molecules through virtual screening, utilizing strategies such as molecular docking followed by subsequent experimental validation. This inevitably allows for the decrease in time and cost for the discovery or development of novel drugs (Maia et al., 2020). Docking-based virtual screening is a very common method used in drug discovery but requires the structure of the target molecule to have been solved. Several GPCR structures have been solved in complex with allosteric modulators. These modulators have been observed to bind in a variety of locations from the outer vestibule in the extracellular region to the protein-lipid interface (Lu and Zhang, 2019). However, the location of allosteric sites on numerous medically relevant GPCRs remains unknown.

Computational methods have increasingly been employed to study allostery of proteins, including the characterization of allosteric pathways and pockets (Wagner et al., 2016). Among the vast arsenal of algorithms in this field, computational solvent mapping is among the most promising for the identification of allosteric sites. Yet, most of these methods are limited to using a static crystal structure of a protein for analysis, which neglects protein flexibility. For example, a previous study of the M2 muscarinic receptor (M2R) found a cryptic site through molecular dynamics simulations that was not present in the original crystal structure (Hollingsworth et al., 2019). To remedy this situation, we have been developing and making improvements to the mixed-solvent molecular dynamics (MixMD) method, a molecular dynamics–based strategy that employs small organic probes with druglike qualities to identify druggable hotspots in multiple replicate short-timescale simulations (Ghanakota and Carlson, 2016). Nevertheless, all targets analyzed by MixMD thus far have been soluble proteins. In contrast, GPCRs are transmembrane proteins that require a lipid bilayer in simulations. To the best of our knowledge, there have only been two studies [hERG1 potassium channel (Mousaei et al., 2020) and K-Ras (Prakash et al., 2015)] using techniques similar to MixMD that have been applied to membrane proteins.

In the present study, we report the development and application of MixMD to membrane-bound GPCRs. Using MixMD on a test set of five GPCRs (cannabinoid receptor type 1, C-C chemokine receptor type 2, M2 muscarinic receptor, P2Y purinoceptor 1, and protease-activated receptor 2) with known allosteric sites in diverse locations resulted in identification of confirmed allosteric sites on all of the GPCRs. Additionally, a case study with the μ-opioid receptor (MOR) revealed potential allosteric sites that could be used in structure-based drug design efforts.

Materials and Methods

Structure Preparation. For the test set, five GPCR structures solved in complex with an allosteric modulator (i.e., 'holo') were selected based on the diversity of where the allosteric modulator bound and served as the basis of comparison. Corresponding structures without a bound allosteric modulator (i.e., 'apo') were used as the starting point for simulations. An additional two 'apo' structures for the μ-opioid receptor were used as case studies. A summary of all previously reported structures used in the present study is provided in Supplemental Table 1 and Protein Data Bank (PDB) information provided in Supplemental Table 2. Chain A was used whenever two GPCRs were present in the unit cell. Modifications had to be made to certain GPCRs. Intracellular loop 3 (ICL3) was either modeled with the 'Loop Modeler' module from Molecular Operating Environment (MOE, 2020.09 release) or joined together after removal of the fusion protein where applicable. Moreover, ICL3 from the M2 muscarinic receptor was omitted as done previously (Hollingsworth et al., 2019), and the remaining loop terminal residues were joined together; this loop, consisting of over 100 residues, has been shown earlier with circular dichroism to be intrinsically disordered (Ichiyama et al., 2006). The N-terminal region of MOR, which makes a nonphysiologic contact with the agonist BU72 (PDB: 5C1M) (Huang et al., 2015), was omitted from the structure. All structures were manually stripped of waters, ions, small molecules, and fusion proteins and were then subjected to the 'Structure Preparation' module from MOE. N- and C termini were capped with acetate and N-methyl amide, respectively. Additionally, hydrogens were added, and the protonation state for all residues was set at pH 7.0.

System Setup. All GPCR structures were embedded into a lipid bilayer using the Membrane Builder module from CHARMM-GUI (Lee et al., 2016). Membrane orientation of the GPCRs was derived from the Orientations of Proteins in Membranes (OPM) database (Lomize et al., 2012). The lipid bilayer consisted of approximately 200 phosphatidylcholine (POPC) molecules and 20 cholesterol molecules per system. Waters were added to each face of the bilayer to a thickness of ~40 Å. Lipid entries in the resulting PDB file (step5_assembly.pdb) were converted to an AMBER18-readable format with charmmlipid2amber.py from AmberTools18. In our standard MixMD protocol (Ghanakota and Carlson, 2016), a protein is surrounded with small organic probes (pyrimidine, acetonitrile, or isopropyl alcohol). However, this procedure is not compatible with membrane proteins, as the probes would clash with lipids in the membrane. One could remove the lipids in question, but that would likely prevent the formation of proper protein-lipid interactions. Therefore, a custom Perl script was written to only allow probes in the aqueous phase for the start of the simulation. In brief, the GPCR was surrounded with probes at a low concentration using the tleap program in AMBER18 (Case et al., 2018), where waters that clashed with the probes were removed. Probes within the lipid bilayer were removed, and the overall probe concentration was checked with a custom script. In an iterative fashion, the probe concentration was increased until 5% v/v was attained (Ghanakota and Carlson, 2016). Subsequently, a final round of tleap was employed to generate the final simulation input files. All relevant disulfide bonds were set, and the system was neutralized with either Na^+ or Cl^- ions. The GPCR and lipids were parameterized with the ff14SB (Maier et al., 2015) and lipid17 (Case et al., 2021) force fields, respectively. Additionally, TIP3P waters were used (Jorgensen et al., 1983), whereas parameters for each of the probes were from our previous work (Lexa et al., 2014).

Mixed-Solvent Molecular Dynamics Simulations. Molecular dynamic simulations were run using the GPU implementation of PMEMD from AMBER18 (Case et al., 2018). A first round of minimization was performed on the systems with a 10-kcal/mol·Å^2 harmonic restraint on the protein and lipids, which consisted of 2500 steps of steepest descent followed by 2500 steps of conjugate gradient. The second round of minimization reduced the harmonic restraint to 10 kcal/mol·Å^2, whereas the final round removed the restraint completely. Subsequently, the systems were gradually heated in the NVT ensemble from 0 K to 100 K for 12.5 picoseconds (ps), then from 100 K to 310.15 K in the NPT ensemble for 125 ps at 1 bar. A 10-kcal/mol·Å^2 harmonic restraint was applied to the GPCR and lipids. Upon reaching the target temperature, the systems were equilibrated in the NPT ensemble at 310.15 K and 1 bar. Beginning

with a 5-kcal/mol·Å^2 harmonic restraint on the protein, the restraint was reduced by 1 kcal/mol·Å^2 every 2 ns for a total of 10 ns, then by 0.1 kcal/mol·Å^2 every 2 ns for a total of 20 ns of equilibration. Production simulations had no constraints and were run in the NPT ensemble at 310.15 K and 1 bar for 100 ns except for the cannabinoid receptor type 1 (CB1R), which was run for 150 ns; an additional 50 ns was required for CB1R for the probes to fully map its known allosteric site. All bond lengths concerning hydrogen were constrained with the SHAKE algorithm. A 1-femtosecond (fs) timestep was used for all simulations. Nonbonded interactions were set to cut off at 9.0 Å, whereas Particle Mesh Ewald summation was used for long-range electrostatics. The Langevin thermostat and Monte Carlo barostat were used where applicable. Semi-isotropic pressure coupling and constant surface tension were employed. Periodic boundary conditions were applied to all heating, equilibration, and production runs. Ten independent production runs were carried out for each probe, culminating in 22.5 μs of simulation time for all systems.

Probe Occupancy Profile Calculations. From AmberTools18, CPPTRAJ was employed to calculate the probe occupancies in Z-scores (below) of different components of the systems. Using the final 10 ns from each independent simulation, the frames were centered on the lipid bilayer while all components were imaged and centered. Slabs 0.25 Å thick were taken along the Z axis, which could be conceptualized as analyzing downward from top to bottom from Fig. 1B. The GPCR, lipids, water, and probes were all examined separately.

Probe Occupancy Calculations. CPPTRAJ in AmberTools18 was used to center, image, and align the trajectories for each probe. The locations of the probe atoms from the last 10 ns of each independent simulation were combined and binned onto a grid with 0.5 Å-spacing. To compare across conditions, Z-scores were then calculated using the following equation to normalize the occupancies:

$$z_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

where x_i is the occupancy at grid point i, μ is the mean occupancy of all grid points, and σ is the standard deviation of occupancy at all grid points. The normalized occupancy maps for each probe can then be visualized like electron densities with the PyMOL Molecular Graphics System (version 2.4.1; Schrödinger, LLC).

Assessment of Predicted Sites. MixMD ProbeView analysis (Graham et al., 2018) was used to identify and rank the predicted sites from the occupancy data for each probe in PyMOL. The parameters for DBSCAN clustering were set at occupancy cutoff = 0.1, ε = 3, and minimum number of points = 50. The 'apo' and 'holo' structures were aligned with the starting structure to see whether the orthosteric or allosteric sites were mapped by our method.

Molecular Docking of BMS986122 to MOR. The active-state crystal structure of MOR (PDB: 5C1M) was prepared with the 'Structure Preparation' module from MOE. Meanwhile, BMS986122 was imported into a MOE database and prepared with the 'Wash' module from MOE, generating a minimized 3D conformation with its dominant form at pH 7.4. The allosteric site was specified using residues within 4.5 Å of the allosteric modulators from the 'holo' structures of CB1R (PDB: 6KQI) and protease-activated receptor 2 (PAR2) (PDB: 5NDZ) after aligning them onto MOR. Docking was performed using Triangle Matcher placement with the London dG scoring function while Induced Fit refinement with the GBVI/WSA dG scoring function and tethered residues was used; default settings were retained where relevant. Finally, the top-scoring pose was kept for analysis.

Results

MixMD Performance on a Test Set of Receptors. To determine the accuracy of the MixMD simulations, we used a test set of five GPCRs, namely cannabinoid receptor type 1 (CB1R), C-C chemokine receptor type 2 (CCR2), M2

muscarinic receptor (M2R), P2Y purinoceptor 1 (P2RY1), and protease-activated receptor 2 (PAR2) with structurally confirmed allosteric binding sites in a variety of locations. We employed the allosteric modulator–free versions of the receptors or 'apo' structures (Supplemental Table 1) to reduce any bias arising from the allosteric bound or 'holo' conformations and to emulate a situation in which the allosteric site is not known. Slight differences were observed in the allosteric sites between the 'apo' and 'holo' states for each GPCR (Supplemental Fig. 1), which lends credence to this strategy.

In traditional MixMD simulations (Ghanakota and Carlson, 2016), a protein is surrounded with a layer of one of three organic probes (pyrimidine, acetonitrile, or isopropyl alcohol) prior to simulation. To date, this method has been applied only to soluble proteins, whereas GPCRs are transmembrane proteins that require a lipid bilayer, which would clash with the probes. To remedy this, we modified the original approach to remove any organic probes overlapping with the lipid bilayer (Fig. 1A). In addition, to allow time for the organic probes to traverse the lipid bilayer or move along the GPCR transmembrane domains facing the lipids, we employed longer simulation timescales (100–150 ns) than the conventional MixMD method (20–50 ns). An example system setup for CCR2 probed with pyrimidine is shown in Fig. 1B. Production simulations for all GPCRs and probe combinations were stable as determined by root-mean-square deviation (RMSD) to the 'apo' structure (Supplemental Fig. 2A). The trajectories involving the CB1R were slightly noisier than the others, but this could be explained because we had to model the entire third intracellular loop (ICL3) de novo; nonetheless, during the simulation the structure did not deviate greater than 5 RMSD from the starting structure (Supplemental Fig. 2A). Electron density profiles of the different components of the systems reveal that the GPCR, lipids, and water remain in the proper locations over the course of the simulations; conversely, and as expected with such a complex system, the organic probes were not completely evenly distributed through the entire system, though they did not all partition into one phase over another (Supplemental Fig. 3A).

Potential binding sites were determined and ranked by employing ProbeView analysis (see methods) (Graham et al., 2018) of the most highly occupied grid points from probe occupancies for each receptor in the 'apo' state. Ranked results are shown in Table 1, and hotspots and ProbeView analyses are displayed in Supplemental Fig. 4A. Overall, the number of potential sites found on each receptor, including the orthosteric site, ranged from 6 to 14. The orthosteric sites were ranked within the top three sites for CB1R, CCR2, and M2R and lower for PAR2; the orthosteric site of P2RY1 was not identified. The 'orthosteric' ligands for PAR2 and P2RY1 are actually allosteric modulators that bind to a site adjacent to the orthosteric site (Zhang et al., 2015; Cheng et al., 2017). The P2RY1 orthosteric antagonist MRS2500 is an adenosine nucleotide analog that binds to the receptor via a charged interaction between its phosphate groups and basic residues on the receptor (Zhang et al., 2015). Given that our set of probes are neutral, we did not expect charged regions to be mapped. Furthermore, pyrimidine and acetonitrile hotspots are located close to the adenine ring of MRS2500 (Supplemental Fig. 4A). Therefore, had we employed negatively charged probes, it is possible that their hotspots would join up with the hotspots identified by the neutral probes to form an identifiable
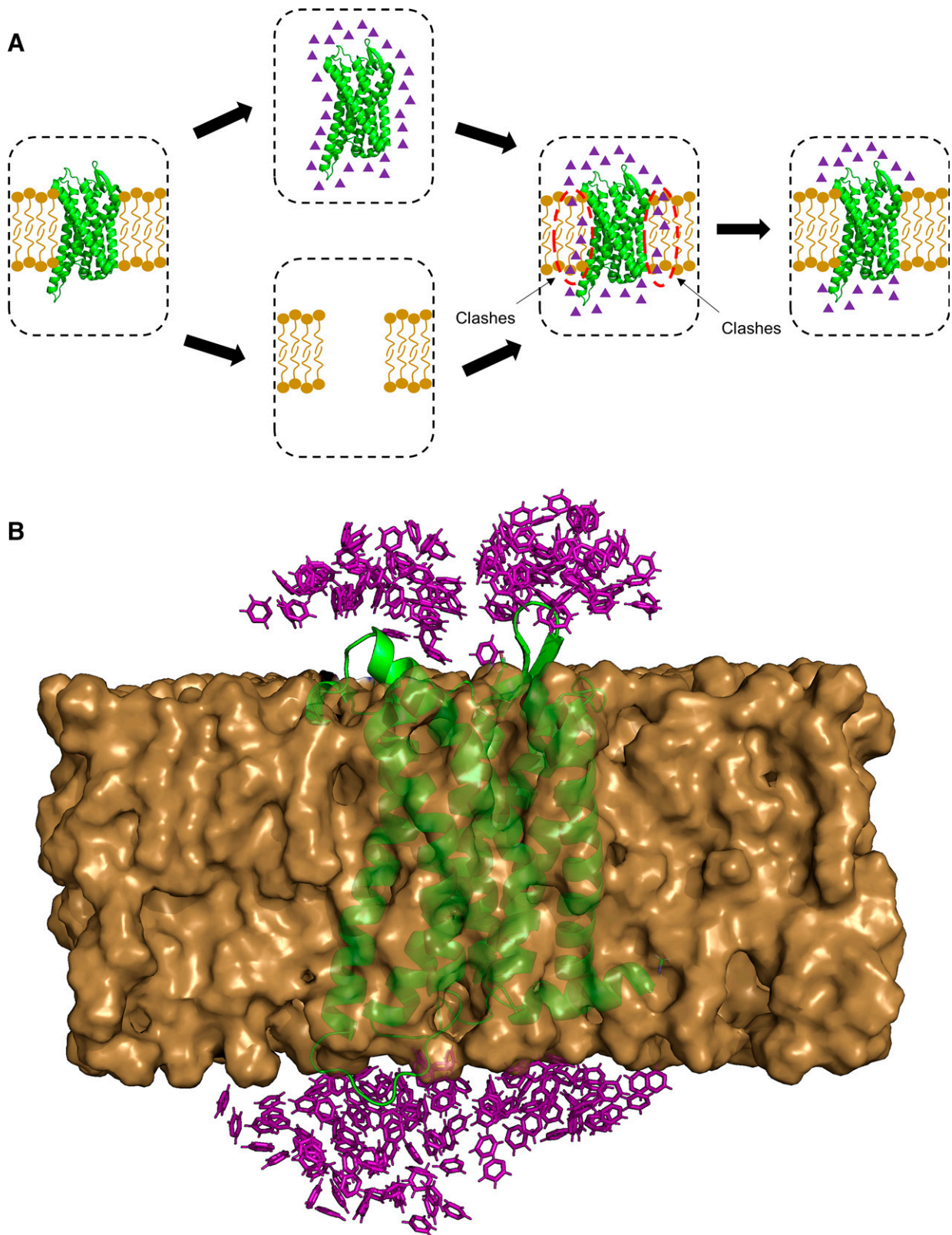
Fig. 1. System setup as exemplified by the C-C chemokine receptor type 2 (CCR2). (A) Setup begins with a receptor (green) embedded in a lipid bilayer (tan). The receptor is then solvated separately from the lipid bilayer with one of the small molecule probes (purple). The receptor is then reintegrated into the lipid bilayer, and probes that reside in the membrane are removed. The resulting system has probes packed near the extracellular- and intracellular-facing regions of the receptor. (B) The resulting system is shown in a 3D-rendering; water molecules have been removed for clarity.

site. Though all 'apo' receptors had an orthosteric ligand bound or allosteric modulator bound in an adjacent pocket, several 'holo' structures also had orthosteric ligands bound and were

identified by ProbeView analysis (Supplemental Fig. 4A). On the other hand, allosteric sites were identified among the predicted sites for all five GPCRs in the 'apo' state within the test

TABLE 1
Ranking of predicted allosteric sites using ProbeView analysis of MixMD simulations

| Name	Abbreviation	Number of Sites Found	ProbeView Site Ranking	
			Orthosteric	Allosteric
CB1 Receptor	CB1R	11	2,3	9
C-C Chemokine Receptor Type 2	CCR2	6	1	4
M2 Muscarinic Receptor	M2R	13	1	13
P2Y1 Receptor	P2Y1	14	Not Found	5
Protease-Activated Receptor 2	PAR2	14	8	2

set, but their ranking varied widely. Next, we detail predicted allosteric sites for each of the receptors in the test set and compare these with solved structures.

Cannabinoid Receptor Type 1. In the 'holo' structure (PDB: 6KQI), CB1R is bound to CP55940 an orthosteric agonist and ORG27569, a negative allosteric modulator (NAM) (Shao et al., 2019). ORG27569 sits in an open pocket formed between transmembrane (TM)2, TM3, and TM4 (Fig. 2A). This site was ranked 9 out of 11 potential predicted sites using ProbeView analysis (Table 1). In fact, the chloro-indole group of ORG27569 is near the predicted site while also corresponding with the pyrimidine and isopropyl occupancy (Fig. 2B). In general, aromatic and hydrophobic residues primarily form the pocket that accommodates ORG27569 (Supplemental Fig. 5A). Mutation of His-154 and Phe-237 was found in a previous study to significantly decrease probe binding sensitivity of ORG27569 (Stornaiuolo et al., 2015). Both His-154 and Phe-237 are in contact with the chloro-indole group of ORG27569, in addition to likely helping stabilize pyrimidine and isopropyl alcohol binding (Supplemental Fig. 5A), suggesting a role in negative allosteric modulation.

C-C Chemokine Receptor Type 2. The 'holo' structure of CCR2 (PDB: 5T1A) consists of the receptor bound with the orthosteric agonist BMS-681 and the NAM CCR2-RA-[R] (Zheng et al., 2016). CCR2-RA-[R] occupies a site within the cytosolic portion of the receptor (Fig. 3A) that overlaps with the G protein binding site and ostensibly sterically hinders G protein binding. This site is only partially formed in the 'apo' structure (Supplemental Fig. 1B). However, with MixMD simulations, we were able to open up the pocket more fully, resulting in the presence of all three probes within the site (Fig. 3B). Additionally, this site ranked 4 out of 6 (Table 1). In particular, the 2-fluoro-4-chlorophenyl moiety of CCR2-RA-[R] overlaid with the pyrimidine and acetonitrile occupied spots as well as with the predicted site (Fig. 3B). Alignment of the 'holo' structure with an average structure of the last 10 ns reveals that CCR2-RA-[R] fits into the opened up pocket (Fig. 3B), suggesting that its shape may be orthosteric-ligand dependent. Nonetheless, it was previously established through mutagenesis that Val-244, Lys-311, Phe-312, and Tyr-305 are critical for CCR2-RA-[R] binding (Zweemer et al., 2014). Of these, Phe-312 and Tyr-305 appear to be directly responsible for
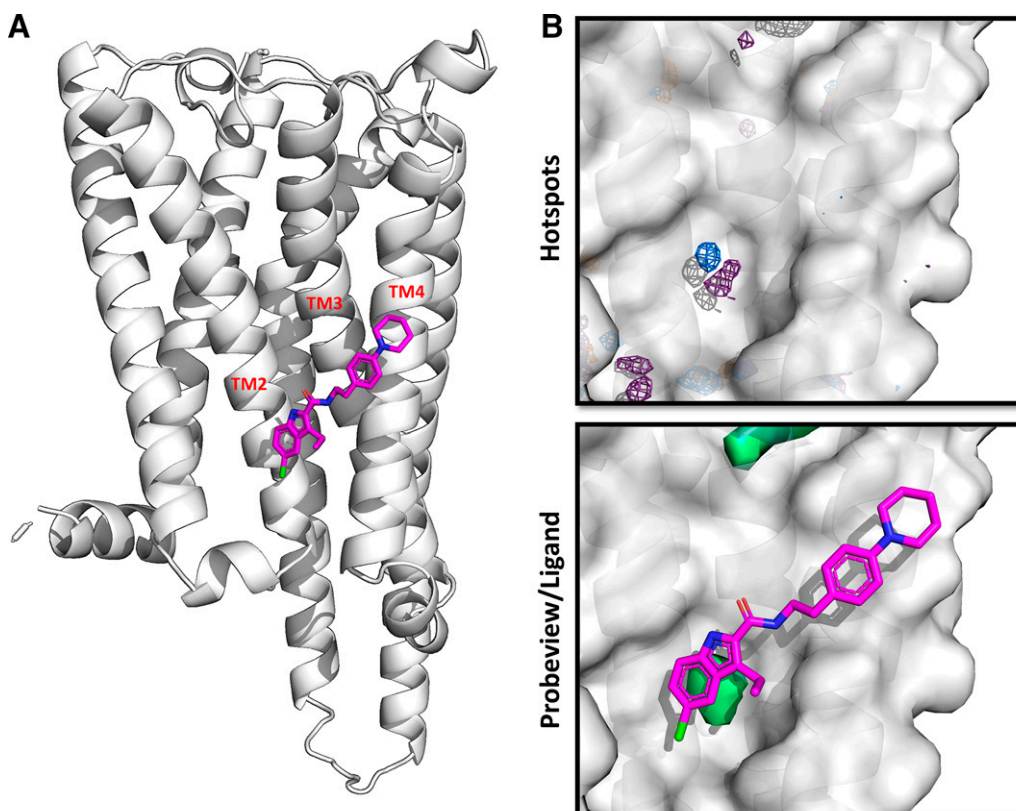
Fig. 2. Allosteric site prediction for cannabinoid receptor type 1 (CB1R). (A) The 'holo' structure (PDB: 6KQI) was aligned onto the averaged structure (white) from the pyrimidine MixMD simulations. The crystal structure of the receptor was removed for clarity while only keeping ORG27569 (magenta). (B) Pyrimidine (purple) and isopropyl alcohol (blue) hotspots were observed within the allosteric site at 50 σ. The site analysis from ProbeView is shown in green.
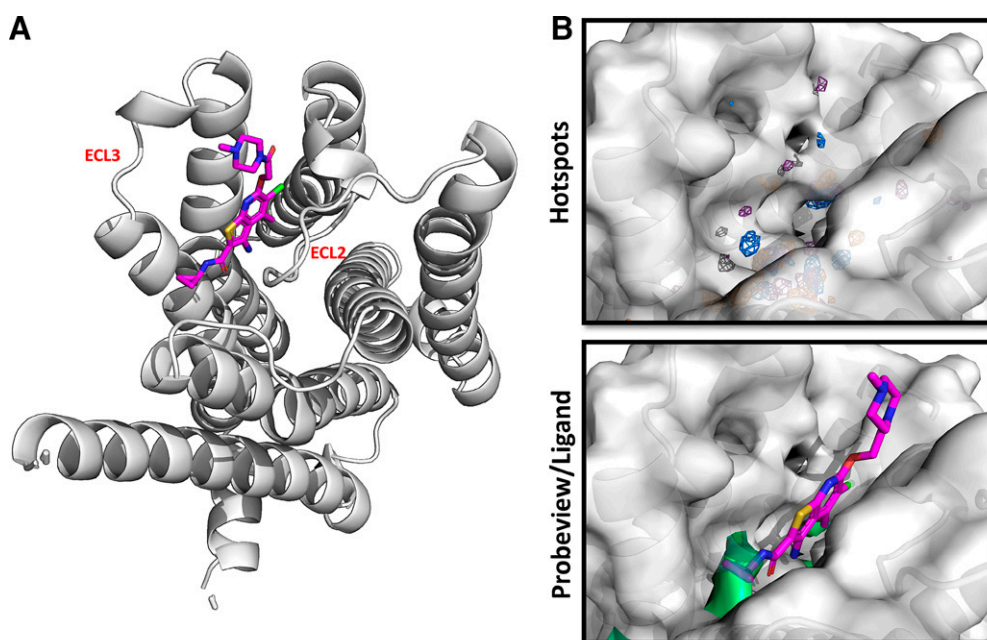
A

B
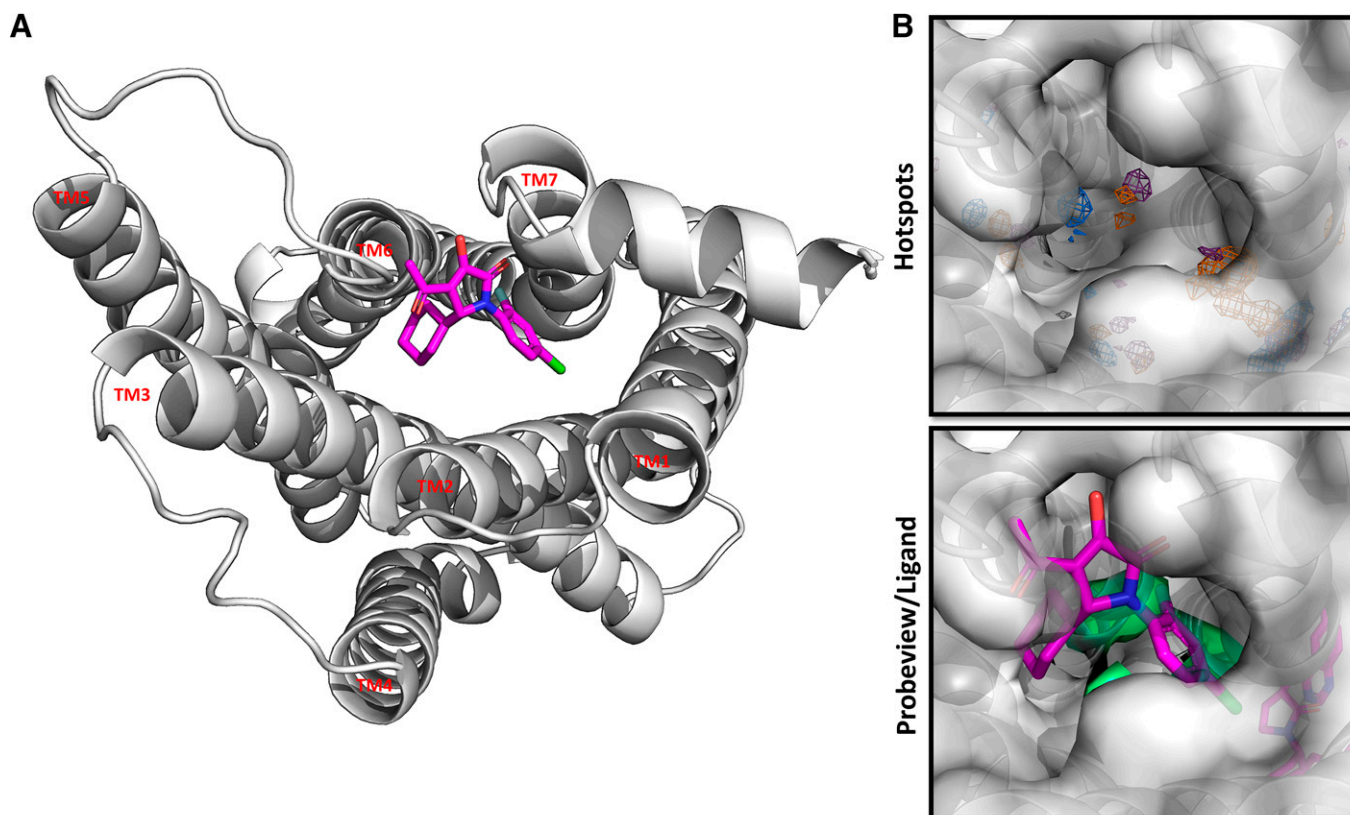
Fig. 3. Allosteric site prediction for C-C chemokine receptor type 2 (CCR2). (A) The 'holo' structure (PDB: 5T1A) was aligned onto the averaged structure (white) from the pyrimidine MixMD simulations. The crystal structure of the receptor was removed for clarity while only keeping CCR2-RA-[R] (magenta). (B) Pyrimidine (purple), acetonitrile (orange), and isopropyl alcohol (blue) hotspots were observed within the allosteric site at 50 σ. The site analysis from ProbeView is shown in green.

interacting with and stabilizing the probes overlapping with the 2-fluoro-4-chlorophenyl moiety of CCR2-RA-[R] (Supplemental Fig. 5B).

M2 Muscarinic Receptor. The 'holo' structure of M2R (PDB: 4MQT) is the only active state structure that was used

in the test set. It is bound to both an agonist, iperoxo, and a positive allosteric modulator (PAM), LY2119620 (Kruse et al., 2013). Despite MixMD simulations starting from the inactive state and the 'apo' structure, we were able to identify the allosteric site, albeit with a low ranking (Table 1). LY2119620

A

B

Fig. 4. Allosteric site prediction for M2 muscarinic receptor (M2R). (A) The 'holo' structure (PDB: 4MQT) was aligned onto the averaged structure (white) from the pyrimidine MixMD simulations. The crystal structure of the receptor was removed for clarity while only keeping LY2119620 (magenta). (B) Pyrimidine (purple) and isopropyl alcohol (blue) hotspots were observed within the allosteric site at 50 σ. The site analysis from ProbeView is shown in green.

A

B
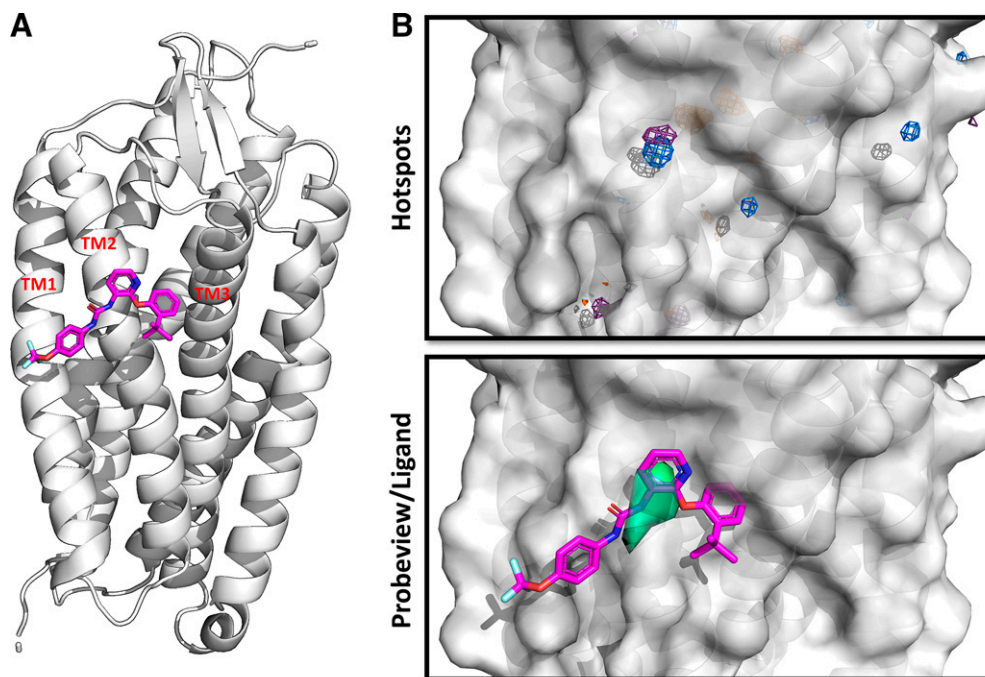
Fig. 5. Allosteric site prediction for P2Y purinoceptor 1 (P2RY1). (A) The 'holo' structure (PDB: 4XNV) was aligned onto the averaged structure (white) from the pyrimidine MixMD simulations. The crystal structure of the receptor was removed for clarity while only keeping BPTU (magenta). (B) Pyrimidine (purple) and isopropyl alcohol (blue) hotspots were observed within the allosteric site at 50 σ. The site analysis from ProbeView is shown in green.

binds in the extracellular vestibule above the orthosteric site, formed primarily by extracellular loops 2 and 3 (ECL2 and ECL3) (Fig. 4A). The predicted site overlapped with the cyclopropyl moiety of LY2119620 and isopropyl alcohol occupancy (Fig. 4B); this region remains in a similar conformation between the 'apo' and 'holo' states (Supplemental Fig. 1C). Important interactions with LY2119620 include a triple aromatic stacking interaction between Trp-422, the aromatic rings of LY2119620, and Tyr-177 and a charged interaction with Glu-172 (Supplemental Fig. 5C). In the former case, a pyrimidine hotspot was observed near Trp-422 (Supplemental Fig. 5C). Through the course of the simulations, ECL2 and ECL3 were unable to completely close to form the narrow pocket that holds LY2119620 (Supplemental Fig. 6). This could account for why stronger probe occupancy was not observed because Tyr-177 would be too far away to form an aromatic stacking interaction. As noted above, the charged interaction was not mapped because we did not employ charged probes in the study.

P2Y Purinoceptor 1. A NAM, BPTU (N-[2-[2-(1,1-Dimethylethyl)phenoxy]-3-pyridinyl]-N'-[4-(trifluoromethoxy)phenyl]urea), is bound to the 'holo' structure of P2Y1 (PDB: 4XNV) (Zhang et al., 2015). This site ranked 5 out of the 14 predicted sites (Table 1). BPTU lies in an open pocket formed between TM1, TM2, and TM3 (Fig. 5A). The pocket is largely similar between the 'apo' and 'holo' states (Supplemental Fig. 1D). The predicted site overlapped with the pyridyl and urea moieties of BPTU, in which the former corresponded with pyrimidine and isopropyl alcohol probe occupancies; an isopropyl alcohol hotspot was also observed near the site where the t-butylphenyl moiety resides, although it was not part of the predicted site (Fig. 5B). A multitude of hydrophobic and aromatic residues contributes to the pocket characteristics, of which Ala-106 and Phe-119 participate in a hydrophobic interaction with the pyridyl group of BPTU (Supplemental Fig. 5D). Mutagenesis of Ala-106 and Thr-103 was previously shown to abrogate the binding of BPTU to P2RY1 (Zhang

et al., 2015). Taken together, these data suggest that Ala-106, and perhaps Phe-119, are involved in forming the hotspot occupancy during MixMD simulations, which led to the correct prediction of the allosteric site.

Protease-Activated Receptor 2. AZ3451 is a NAM that was cocrystallized with the 'holo' structure of PAR2 (PDB: 5NDZ) (Cheng et al., 2017). From our ProbeView analysis, the allosteric site was ranked 2 out of 14 total predicted sites. The pocket is very clearly defined and formed between TM2, TM3, and TM4 (Fig. 6A). It is only partially open in the 'apo' state but fully open in the 'holo' state (Supplemental Fig. 1E). The 1,3-benzodioxole ring of AZ3451 was completely mapped, whereas acetonitrile and pyrimidine occupancies overlaid with the 1,3-dioxole moiety; overall, all three probes were involved in defining the predicted site (Fig. 6B). Numerous hydrophobic and aromatic residues form the allosteric pocket (Supplemental Fig. 5E), where Ala-157 (Gly-157 in wild type) and Tyr-210 were determined to be critical for AZ3451 binding through mutagenesis (Cheng et al., 2017). These two residues are positioned such that they form the predicted site, suggesting that they were involved in probe binding during the MixMD simulations.

Comparison of MixMD with FTMap Using the Test Set of GPCRs. FTMap is an algorithm that predicts druggable hot spots on biologic macromolecules by docking small organic probes to a multitude of positions, prioritizing regions that represent a cluster of several probe types (Ngan et al., 2012). FTMap is not analogous to MixMD but can be used for binding site prediction. To provide a means of comparing the predicted efficiency of our approach, we submitted the 'apo' structures from the test set to the FTMap server (Kozakov et al., 2015). Results for the five test GPCRs are shown in Supplemental Fig. 7. Whereas our MixMD method was able to identify hotspots corresponding with all allosteric sites (Figs. 2B, 3B, 4B, 5B, and 6B), FTMap was only able to correctly predict hotspots for CB1R, M2R, and PAR2 (Supplemental Fig. 5), although the two methods agreed in

A

B
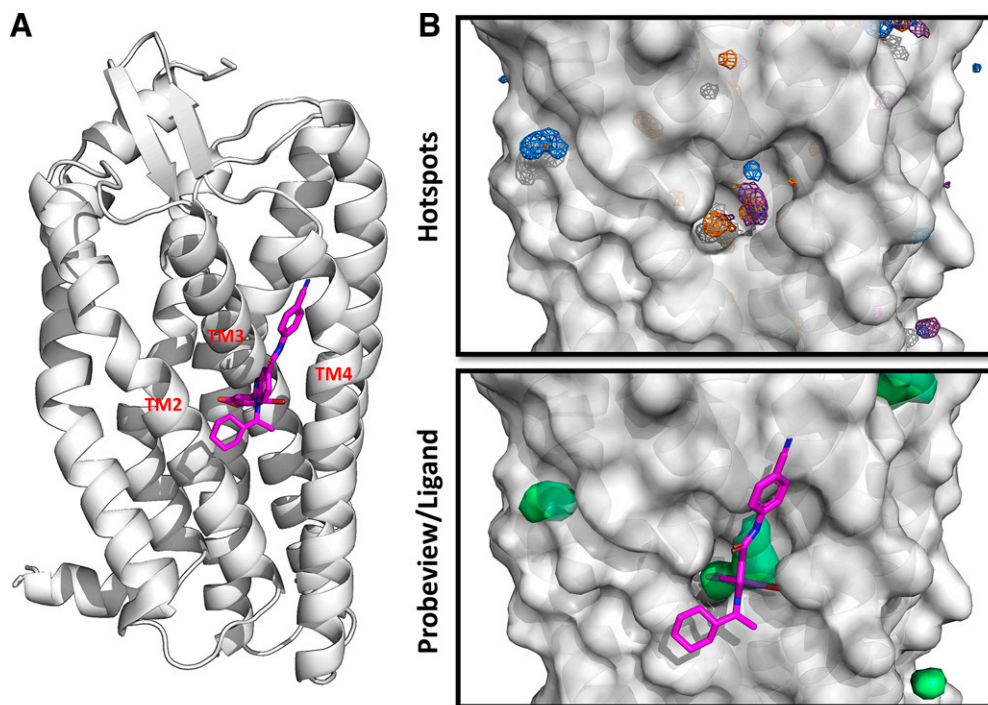
Fig. 6. Allosteric site prediction for protease-activated receptor 2 (PAR2). (A) The 'holo' structure (PDB: 5NDZ) was aligned onto the averaged structure (white) from the pyrimidine MixMD simulations. The crystal structure of the receptor was removed for clarity while only keeping AZ3451 (magenta). (B) Pyrimidine (purple), acetonitrile (orange), and isopropyl alcohol (blue) hotspots were observed within the allosteric site at 50 σ. The site analysis from ProbeView is shown in green.

the location for these. FTMap did not identify hotspots on CCR2 or P2RY1. In the former, it is likely that this is because the allosteric site is only partially formed in the 'apo' structure (Supplemental Fig. 1B); in the latter, it is possible that not enough probe types were clustered, preventing the site from entering the pool of candidate hotspots despite the 'apo' and 'holo' states being virtually identical in this region (Supplemental Fig. 1D).

Case Study To Identify Allosteric Sites on MOR. Active- (PDB: 5C1M) and inactive-state (PDB: 4DKL) structures of MOR, crystallized bound to the orthosteric agonist BU72 (Huang et al., 2015) or the irreversible orthosteric antagonist β-funaltrexamine (Manglik et al., 2012), were examined using MixMD. As there are currently no known structures of MOR in complex with an allosteric modulator, these structures are in the 'apo' state. Protein RMSDs during production simulations and corresponding electron density profiles are shown in Supplemental Figs. 2B and 3B, respectively. Sixteen potential sites were identified using the active structure of MOR and 20 sites using the inactive structure.

For the active structure of MOR, the phenyl ring of BU72 was mapped by all three probes, whereas its phenol ring overlaid with isopropyl alcohol occupancy (rank: 9/16; Fig. 7). Similarly, a site (rank: 1/16) corresponding to the position where nanobody 39 (Nb39) interacts with the receptor was mapped by all probes (Fig. 7); Nb39 is a camelid antibody fragment used to stabilize active-state MOR (Huang et al., 2015). With the inactive structure of MOR, β-funaltrexamine was in the vicinity of several probe occupancies (Fig. 7). Thus, using MixMD, we were able to successfully map out portions of ligands within the orthosteric pocket of MOR as well as the Nb39 binding site, validating the MixMD approach to identify binding sites on MOR.

Study of the active-state structure of MOR identified additional potentially allosteric sites. Three of these sites formed a near contiguous pocket between TM2, TM3, and TM4

(Fig. 7). For two of the sites, all probes mapped strongly. One of these sites (rank: 3/16) was similarly located as the known allosteric site on PAR2 and had a similar shape as evidenced by how the 1,3-benzodioxole ring of AZ3451 overlays this site (Fig. 7A). The other major predicted site (rank: 5/16) was analogous to the ORG27569 site on CB1R, as can be seen by overlay of ORG27569 to the site on MOR (Fig. 7B). The third site (rank: 8/16) was smaller and had strong pyrimidine occupancy; it could be connected to one or both of the other sites. Two additional predicted sites lay near the G protein binding site (Fig. 7, view 2). The first of these sites (rank: 4/16) was mapped by all probes and was not accessible in the starting structure, suggesting it is a cryptic site (Vajda et al., 2018) (Fig. 7C, left). Additional sites were predicted along TM6, TM7, and TM1, but unlike the other sites, these did not have precedent 'holo' structures and thus were of lower confidence (Supplemental Fig. 8, left).

Potential allosteric sites were also identified using the inactive state structure of MOR. Similar to the active-state MOR structure, a putative site identified by all three probes (rank: 10/20) was observed to overlap with the chloro-indole group of ORG27569 from the CB1R structure (Fig. 8A). A second site (rank: 6/20) mapped by pyrimidine and acetonitrile occupied a region analogous to where CCR2-RA-[R] binds on CCR2, as shown by the overlay of CCR2-RA-[R] onto this site on MOR (Fig. 8C). A further site (rank: 7/20) mapped by all probes was predicted to be formed between TM3, TM4, and TM5 closer to the cytosolic region (Fig. 8B). This site is in a similar position to the binding site for a PAM on the β2-adrenergic receptor (PDB: 6N48) (Liu et al., 2019). Although this site is closed in the BU72-bound, active-state structure (PDB: 5C1M), it is open in both DAMGO-bound, active-state structures (PDB: 6DDE, 6DDF), neither of which were employed in the present study (Supplemental Fig. 9A). The residues that comprise this site are shown in Supplemental Fig. 9B. Like the active-state structure,
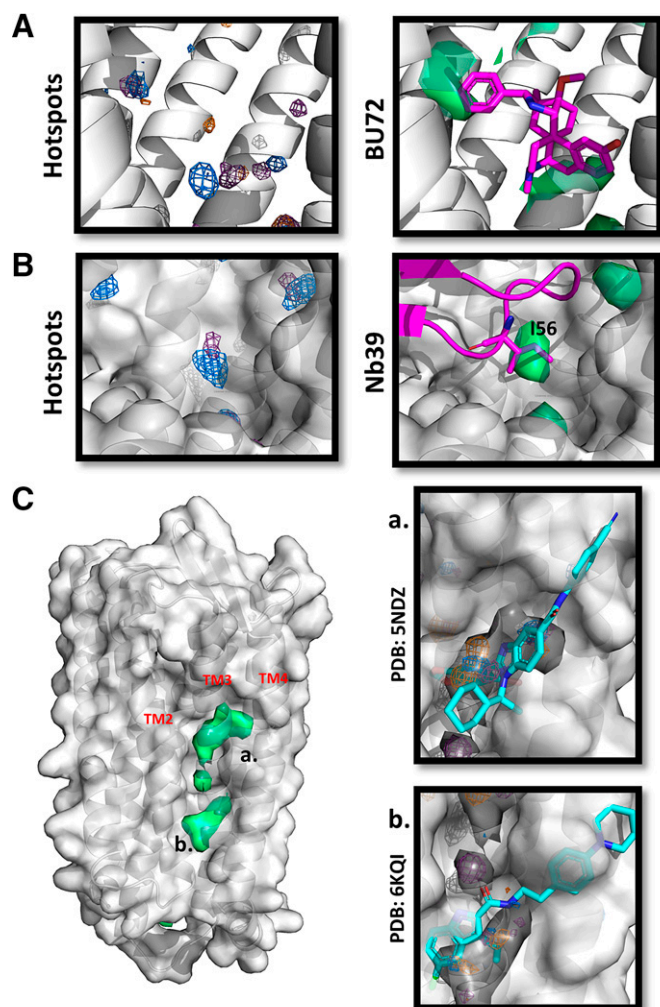
Fig. 7. Allosteric site prediction for active-state μ-opioid receptor (MOR). The binding sites for both (A) BU72 and (B) nanobody 39 (Nb39) were both successfully predicted and ranked at 9/16 and 1/16, respectively. Ligands are displayed in magenta in stick representation. Isoleucine 56 from Nb39 overlaid directly with pyrimidine and isopropyl alcohol probe occupancy. (C) The 'holo' structures for (a) PAR2 and (b) CB1R were aligned to the averaged MOR structure (white) from the pyrimidine MixMD simulations. The original crystal structures were removed for clarity, whereas their respective allosteric modulators (cyan) were kept. Hotspots corresponding to pyrimidine, acetonitrile, and isopropyl alcohol are purple, orange, and blue, respectively, and they are displayed at 50 σ. Site analyses from ProbeView are shown as green, whereas they are gray in the subfigures.

additional predicted sites were also identified along TM5, TM6, TM7, and TM1, which did not possess precedent 'holo' structures in the test set of GPCRs (Supplemental Fig. 8, right).

Discussion

We have developed a novel application of the MixMD simulation method to identify potential binding sites on GPCRs. By employing a test set of five GPCRs that have known structures with bound orthosteric or allosteric modulators in a variety of locations, we were able to validate our method through both "hotspot" and, consequently, binding site prediction. We also identified sites on the test set of GPCRs that are not yet experimentally proven to be allosteric sites. For

example, with the P2Y receptor, high-ranking sites included a region at the cytosolic end of TM domains 1, 2, and 7, a cleft formed between TM5 and TM6 and a pocket between TM6 and TM7. Conversely, the established alternative FTMap method (Ngan et al., 2012) was not able to identify the hotspots for two of the structures from our test set.

Only neutral probes were used in the present study because most of the compounds associated with the test set GPCRs were neutral. However, charged functional moieties on several ligands were not mapped well. For example, with MRS2500 (P2Y1, orthosteric site) and LY2119620 (M2R, allosteric site) salt bridges are crucial for their respective interactions. Thus, using charged probes would have potentially augmented the coverage of our method. We have previously used charged probes with soluble proteins (Graham et al., 2018), and it would be of interest to expand upon their use to membrane-bound proteins.

Although we were successful in identifying known allosteric sites for all test set cases, higher probe occupancies would have been desirable for ranking purposes. A previous study using MixMD with ProbeView analysis correctly identified allosteric sites with the most favorably ranked sites, in soluble proteins (Graham et al., 2018). However, the situation is more complex with membrane-bound GPCRs. Indeed, despite having the same global topology, GPCR structures solved in complex with allosteric modulators show sites in diverse locations (Wakefield et al., 2019). Furthermore, the notion that multiple allosteric sites exist within a single GPCR is becoming a distinct possibility (Congreve et al., 2017), which would convolute site identification if evaluated by ranking only. One example that supports this is the β2-adrenergic receptor. The structure of this receptor has been solved in complex with either a PAM (PDB: 6N48, active state) (Liu et al., 2019) or NAM (PDB: 5X7D, inactive state) (Liu et al., 2017) that bind to separate sites, although these may be specific to the activation state of the receptor.

One solution to the ranking problem is to use normal mode analysis. Previous studies have used protein normal modes to assess whether low-frequency modes are altered between the 'apo' and 'holo' states (Panjkovich and Daura, 2012; Chan et al., 2021). Another recent algorithm "Ohm" maps allosteric networks within a static protein structure and can predict an allosteric site given the orthosteric or active site (Wang et al., 2020). Additional analyses of the predicted hotspots in identifying allosteric sites in GPCRs will be explored in the future to further refine our conclusions from ProbeView analysis.

A series of studies published by the McCammon group utilizing FTMap (Ngan et al., 2012) on static protein structures to identify druggable hotspots and to analyze distinct conformations of GPCRs used long-timescale molecular dynamics simulations to predict allosteric sites. Overall, their method identified seven distinct allosteric sites on M2R across a series of active, intermediate, and inactive state receptor conformations (Miao et al., 2014). Using the inactive state of the receptor, the group identified two sites: one that corresponded with the extracellular vestibule and a lipid-exposed pocket formed by TM5 and TM6 (Miao et al., 2014). The former case matched with our MixMD predictions, signifying both methods were able to identify the known allosteric site to which LY2119620 binds. However, the latter was not identified by our method. In contrast, we identified additional sites on M2R not found in the previous study in the vicinity of the G protein binding site
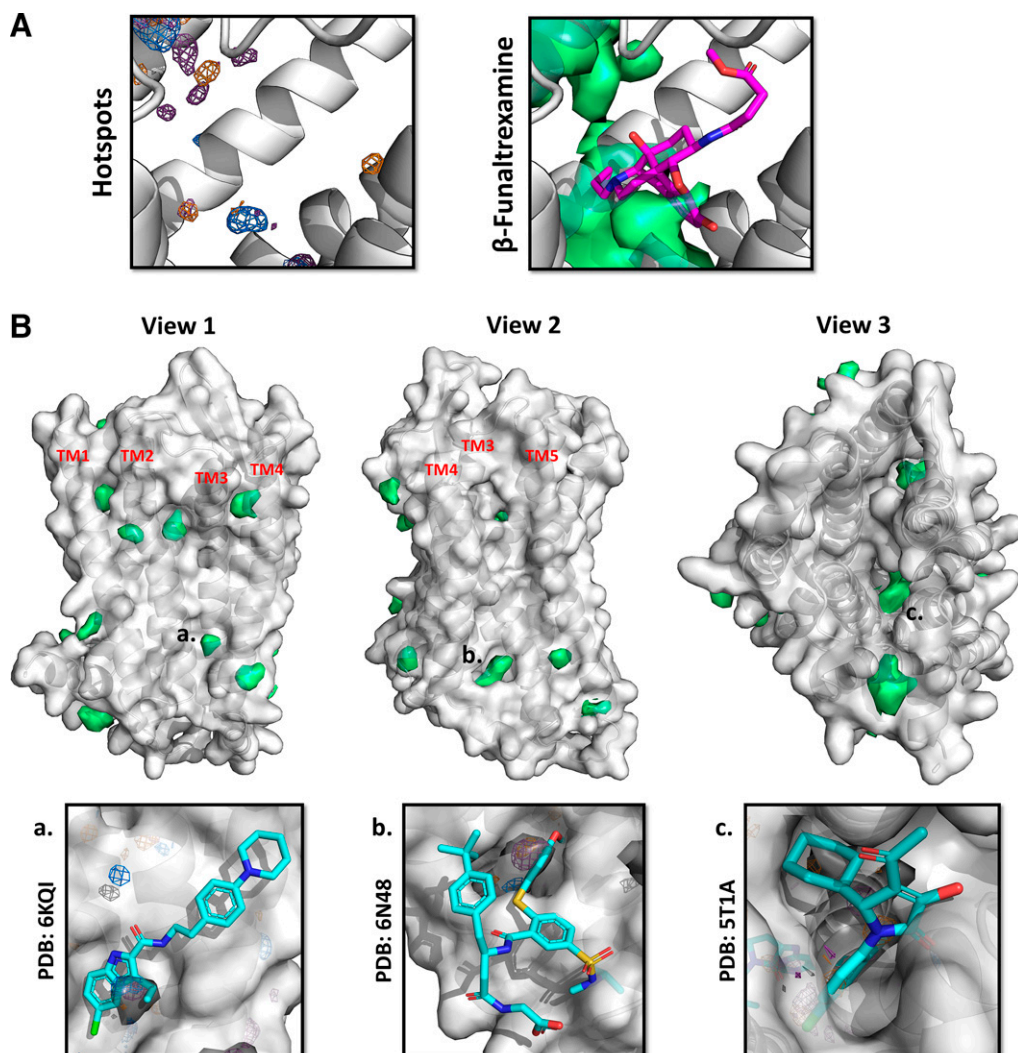
Fig. 8. Allosteric site prediction for inactive-state μ-opioid receptor (MOR). (A) The binding site for β-funaltrexamine was both successfully predicted and ranked at 1/20; ligands are displayed in magenta in stick representation. (B) The 'holo' structures for (a) CB1, (b) β2-adrenergic receptor, and (c) CCR2 were aligned to the averaged MOR structure (white) from the pyrimidine MixMD simulations. The original crystal structures were removed for clarity, whereas their respective allosteric modulators (cyan) were kept. Hotspots corresponding to pyrimidine, acetonitrile, and isopropyl alcohol are purple, orange, and blue, respectively, and they are displayed at 50 σ. Site analyses from ProbeView are shown as green, whereas they are gray in the subfigures.

(Supplemental Fig. 10A). This site appears to be analogous to that in the CCR2R in terms of shape and could accommodate a molecule like CCR2-RA-[R] (Supplemental Fig. 10B). Furthermore, the site is partially closed in the initial 'apo' structure but opens up during MixMD simulations (Supplemental Fig. 10B), suggesting that it is a potential cryptic site. We note that FTMap has the same issue of poorly ranking sites (Supplemental Fig. 7) that was seen in our ProbeView analysis. However, we believe the present comparison shows that MixMD provides certain advantages over FTMap. First, MixMD utilizes small organic probes during the simulation that interact with the receptor in key hotspots to potentially open up cryptic sites (Smith and Carlson, 2021). Second, multiple independent simulations for each probe are run, which might better sample the free energy landscape and confer greater confidence as opposed to analysis of a single long trajectory. Nonetheless, the previous method does allow for larger global conformational changes due to its longer simulation lengths, suggesting that both methods are complementary.

For a test case study, we chose to analyze MOR. This GPCR is the target for all clinically used opioid analgesics (Jutkiewicz and Traynor, 2023). Despite their effectiveness as pain medications, consumption of opioid drugs commonly results in deleterious side effects such as addiction,

constipation, and respiratory depression. However, opioids remain the gold standard for pain relief. Therefore, there is a clear and unmet need for novel analgesics for the management of pain. In this regard, PAMs of MOR have become a viable strategy (Burford et al., 2015b). A number of PAMs have been discovered by means of high-throughput screening, including BMS-986122 (Burford et al., 2013; Livingston and Traynor, 2014), which has been shown to afford pain modulation in vivo (Kandasamy et al., 2021). A prior in silico study proposed a site in the extracellular vestibule for BMS986122 through molecular dynamics simulations (Bartuzi et al., 2016), although no experimental validation was performed.

Recently, an NMR study of human MOR identified T162 (T160 in mouse MOR) as crucial for the binding of BMS-986122 (Kaneko et al., 2022). Our study identified a potential allosteric site around this residue formed between TM3, TM4, and TM5 and comprising numerous hydrophobic residues (Fig. 9; Supplemental Fig. 9B). This site is open in the β-funaltrexamine-bound inactive structure and DAMGO ([D-Ala[2], N-MePhe[4], Gly-ol[5]]-enkephalin)-bound active structure, whereas it is closed in the BU72-bound active structure (Supplemental Fig. 9A). Although the allosteric site was predicted by ProbeView for only the inactive structure (Fig. 8B), hotspots were observed for both active and inactive structures
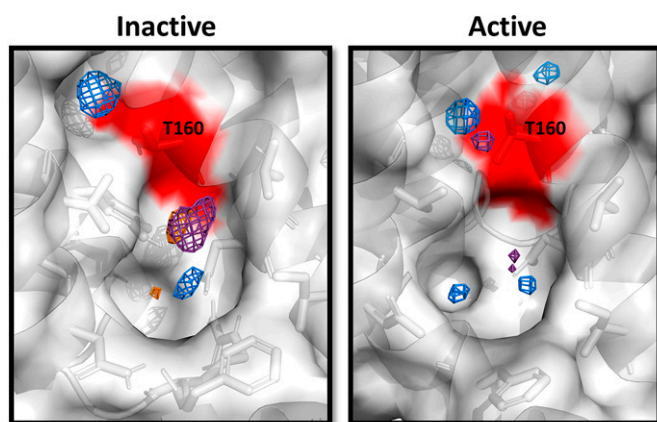
Fig. 9. Predicted allosteric site for BMS986122. The inactive and active state structures of MOR are shown on the left and right, respectively. The DAMGO-bound active structure (PDB: 6DDE) is shown instead of the original BU72-bound active structure used in simulations (PDB: 5C1M) to better illustrate that the pocket is open. Hotspots corresponding to pyrimidine, acetonitrile, and isopropyl alcohol are purple, orange, and blue, respectively, and they are displayed at 50 σ. T162 in human MOR (T160 in mouse MOR) was previously reported in an NMR study to be critical for the binding of BMS986122 and is shown in red with stick representation (Kaneko et al., 2022).

(Fig. 9), suggesting that the allosteric site opens in the active structure. This is consistent with a model of MOR activation in which binding of BMS986122 to MOR together with the agonist DAMGO increases the population of active state conformations (Kaneko et al., 2022). Extending the timescale of the simulations could have provided a more equilibrated population of the open conformation of the pocket and thus a ranked site through ProbeView. We propose to pursue docking and molecular dynamic simulations to determine if reported allosteric modulators of MOR, including the PAMs BMS-986122, BMS-986121 (Burford et al., 2013), BMS-986187 (Burford et al., 2015a; Livingston et al., 2018), and MS1 (Bisignano et al., 2015), as well as putative NAMs such as the cannabinoids (Kathmann et al., 2006; Livingston and Traynor, 2018), interact with putative sites identified on MOR.

It is worth noting that several of the putative allosteric sites on MOR identified by MixMD are analogous to experimentally proven sites on other GPCRs, suggesting that these may be common regions for initiating and/or transmitting allosteric effects. These include positions where NAMs bind to PAR2 or CB1R and a site near the G protein binding region where the negative modulator CCR2-RA-[R] binds to CCR2; as discussed above, we also found this site in M2R. This may suggest that identifying ligands for these sites on MOR could generate NAM molecules. In contrast, a site identified on MOR is in the same location as a known PAM for the β-adrenergic receptor. However, until we identify and examine molecules for binding to these newly predicted sites on MOR, we are unable to determine whether they would act as negative, positive, or neutral allosteric modulators.

During preparation of this manuscript, the development of a cosolvent molecular dynamics–based approach to identify possible nonorthosteric binding sites on GPCRs employing fragments derived from known allosteric modulators as probes was published (Ciancetta et al., 2021). Two of the test cases employed in that study (P2RY1 and M2R) overlapped with ours, and both correctly mapped the allosteric sites. However, key

differences exist between the methods. Our process places the organic probes near the extracellular- and cytosolic-facing regions of the GPCR and allows for free diffusion. The approach of Ciancetta and colleagues (2021) uses a cylindrical "barrier" for the organic probes to allow for confined diffusion in a limited area. Although this was reported to work well in both retrospective and prospective validations, in our study we were able to successfully map allosteric sites in the absence of a barrier. Moreover, the technique of Ciancetta and colleagues (2021) generated fragments derived from known allosteric modulators, whereas we used standard small organic probes, thus providing a wider applicability to GPCRs with no known allosteric modulators. Nonetheless, both methods appear complementary, and it would be interesting to see both incorporated into drug discovery pipelines to aid in the identification of allosteric sites on GPCRs.

References

Bartuzi D, Kaczor AA, and Matosiuk D (2016) Interplay between two allosteric sites and their influence on agonist binding in human μ-opioid receptor. *J Chem Inf Model* **56**:563–570.

Bisignano P, Burford NT, Shang Y, Marlow B, Livingston KE, Fenton AM, Rockwell K, Budenholzer L, Traynor JR, Gerritz SW, et al. (2015) Ligand-based discovery of a new scaffold for allosteric modulation of the μ-opioid receptor. *J Chem Inf Model* **55**:1836–1843.

Burford NT, Clark MJ, Wehrman TS, Gerritz SW, Banks M, O'Connell J, Traynor JR, and Alt A (2013) Discovery of positive allosteric modulators and silent allosteric modulators of the μ-opioid receptor. *Proc Natl Acad Sci U S A* **110**:10830–10835.

Burford NT, Livingston KE, Canals M, Ryan MR, Budenholzer LM, Han Y, Shang Y, Herbst JJ, O'Connell J, Banks M, et al. (2015a) Discovery, synthesis, and molecular pharmacology of selective positive allosteric modulators of the δ-opioid receptor. *J Med Chem* **58**:4220–4229.

Burford NT, Traynor JR, and Alt A (2015b) Positive allosteric modulators of the μ-opioid receptor: a novel approach for future pain medications. *Br J Pharmacol* **172**:277–286.

Case DA, Aktulga HM, Belfon K, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham III TE, Cisneros GA, Cruzeiro VWD, Darden TA, et al. (2021) *AMBER 2021*, University of California, San Francisco, San Francisco, CA.

Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham III TE, Cruzeiro VWD, Darden TA, Duke RE, Ghoreishi D, Gilson MK, et al. (2018) *AMBER 2018*, University of California, San Francisco, San Francisco, CA.

Chan WKB, DasGupta D, Carlson HA, and Traynor JR (2021) Mixed-solvent molecular dynamics simulation-based discovery of a putative allosteric site on regulator of G protein signaling 4. *J Comput Chem* **42**:2170–2180.

Chan WKB and Zhang Y (2020) Virtual screening of human class-A GPCRs using ligand profiles built on multiple ligand-receptor interactions. *J Mol Biol* **432**:4872–4890.

Cheng RKY, Fiez-Vandal C, Schlenker O, Edman K, Aggeler B, Brown DG, Brown GA, Cooke RM, Dumelin CE, Doré AS, et al. (2017) Structural insight into allosteric modulation of protease-activated receptor 2. *Nature* **545**:112–115.

Ciancetta A, Gill AK, Ding T, Karlov DS, Chalhoub G, McCormick PJ, and Tikhonova IG (2021) Probe confined dynamic mapping for G protein-coupled receptor allosteric site prediction. *ACS Cent Sci* **7**:1847–1862.

Congreve M, Oswald C, and Marshall FH (2017) Applying structure-based drug design approaches to allosteric modulators of GPCRs. *Trends Pharmacol Sci* **38**:837–847.

Ghanakota P and Carlson HA (2016) Moving beyond active-site detection: MixMD applied to allosteric systems. *J Phys Chem B* **120**:8685–8695.

Graham SE, Leja N, and Carlson HA (2018) MixMD ProbeView: robust binding site prediction from cosolvent simulations. *J Chem Inf Model* **58**:1426–1433.

Gregory KJ, Noetzel MJ, and Niswender CM (2013) Pharmacology of metabotropic glutamate receptor allosteric modulators: structural basis and therapeutic potential for CNS disorders. *Prog Mol Biol Transl Sci* **115**:61–121.

Hollingsworth SA, Kelly B, Valant C, Michaelis JA, Mastromihalis O, Thompson G, Venkatakrishnan AJ, Hertig S, Scammells PJ, Sexton PM, et al. (2019) Cryptic pocket formation underlies allosteric modulator selectivity at muscarinic GPCRs. *Nat Commun* **10**:3289.

Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn AL, Kato HE, Livingston KE, Thorsen TS, Kling RC, et al. (2015) Structural insights into μ-opioid receptor activation. *Nature* **524**:315–321.

Ichiyama S, Oka Y, Haga K, Kojima S, Tateishi Y, Shirakawa M, and Haga T (2006) The structure of the third intracellular loop of the muscarinic acetylcholine receptor M2 subtype. *FEBS Lett* **580**:23–26.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, and Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**:926–935 DOI: 10.1063/1.445869.

Jutkiewicz EM and Traynor JR (2023) Opioid analgesics, in *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 14th ed (Bruton LL and Knollmann B, eds) pp 443–470, McGraw Hill.

Kandasamy R, Hillhouse TM, Livingston KE, Kochan KE, Meurice C, Eans SO, Li M-H, White AD, Roques BP, McLaughlin JP, et al. (2021) Positive allosteric modulation of the μ-opioid receptor produces analgesia with reduced side effects. *Proc Natl Acad Sci U S A* **118**:e2000017118.

Kaneko S, Imai S, Asao N, Kofuku Y, Ueda T, and Shimada I (2022) Activation mechanism of the μ-opioid receptor by an allosteric modulator. *Proc Natl Acad Sci U S A* **119**:e2121918119.

Kathmann M, Flau K, Redmer A, Tränkle C, and Schlicker E (2006) Cannabidiol is an allosteric modulator at μ- and δ-opioid receptors. *Naunyn Schmiedebergs Arch Pharmacol* **372**:354–361.

Kenakin T (2017) Allosteric drug effects, *Pharmacology in Drug Discovery and Development*, 2nd ed, pp 101–129, Academic Press, New York NY.

Kozakov D, Grove LE, Hall DR, Bohnuud T, Mottarella SE, Luo L, Xia B, Beglov D, and Vajda S (2015) The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat Protoc* **10**:733–755.

Kruse AC, Ring AM, Manglik A, Hu J, Hu K, Eitel K, Hübner H, Pardon E, Valant C, Sexton PM, et al. (2013) Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**:101–106.

Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, Wei S, Buckner J, Jeong JC, Qi Y, et al. (2016) CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J Chem Theory Comput* **12**:405–413.

Lexa KW, Goh GB, and Carlson HA (2014) Parameter choice matters: validating probe parameters for use in mixed-solvent simulations. *J Chem Inf Model* **54**:2190–2199.

Liu X, Ahn S, Kahsai AW, Meng K-C, Latorraca NR, Pani B, Venkatakrishnan AJ, Masoudi A, Weis WI, Dror RO, et al. (2017) Mechanism of intracellular allosteric β2 AR antagonist revealed by X-ray crystal structure. *Nature* **548**:480–484.

Liu X, Masoudi A, Kahsai AW, Huang L-Y, Pani B, Staus DP, Shim PJ, Hirata K, Simhal RK, Schwalb AM, et al. (2019) Mechanism of β2AR regulation by an intracellular positive allosteric modulator. *Science* **364**:1283–1287.

Livingston KE, Stanczyk MA, Burford NT, Alt A, Canals M, and Traynor JR (2018) Pharmacologic evidence for a putative conserved allosteric site on opioid receptors. *Mol Pharmacol* **93**:157–167.

Livingston KE and Traynor JR (2014) Disruption of the Na+ ion binding site as a mechanism for positive allosteric modulation of the μ-opioid receptor. *Proc Natl Acad Sci U S A* **111**:18369–18374.

Livingston KE and Traynor JR (2018) Allostery at opioid receptors: modulation with small molecule ligands. *Br J Pharmacol* **175**:2846–2856.

Lomize MA, Pogozheva ID, Joo H, Mosberg HI, and Lomize AL (2012) OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res* **40**:D370–D376.

Lu S and Zhang J (2019) Small molecule allosteric modulators of G-protein-coupled receptors: drug–target interactions. *J Med Chem* **62**:24–45.

Maia EHB, Assis LC, de Oliveira TA, da Silva AM, and Taranto AG (2020) Structure-based virtual screening: from classical to artificial intelligence. *Front Chem* **8**:343.

Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, and Simmerling C (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* **11**:3696–3713.

Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, and Granier S (2012) Crystal structure of the μ-opioid receptor bound to a morphinan antagonist. *Nature* **485**:321–326.

Miao Y, Nichols SE, and McCammon JA (2014) Mapping of allosteric druggable sites in activation-associated conformers of the M2 muscarinic receptor. *Chem Biol Drug Des* **83**:237–246.

Mousaei M, Kudaibergenova M, MacKerell Jr AD, and Noskov S (2020) Assessing hERG1 blockade from Bayesian machine-learning-optimized site identification by ligand competitive saturation simulations. *J Chem Inf Model* **60**:6489–6501.

Ngan CH, Bohnuud T, Mottarella SE, Beglov D, Villar EA, Hall DR, Kozakov D, and Vajda S (2012) FTMAP: extended protein mapping with user-selected probe molecules. *Nucleic Acids Res* **40**:W271–W275.

O'Hayre M, Vázquez-Prado J, Kufareva I, Stawiski EW, Handel TM, Seshagiri S, and Gutkind JS (2013) The emerging mutational landscape of G proteins and G-protein-coupled receptors in cancer. *Nat Rev Cancer* **13**:412–424.

Panjkovich A and Daura X (2012) Exploiting protein flexibility to predict the location of allosteric sites. *BMC Bioinformatics* **13**:273.

Prakash P, Sayyed-Ahmad A, and Gorfe AA (2015) pMD-membrane: a method for ligand binding site identification in membrane-bound proteins. *PLOS Comput Biol* **11**:e1004469.

Römpler H, Stäubert C, Thor D, Schulz A, Hofreiter M, and Schöneberg T (2007) G protein-coupled time travel: evolutionary aspects of GPCR research. *Mol Interv* **7**:17–25.

Shao Z, Yan W, Chapman K, Ramesh K, Ferrell AJ, Yin J, Wang X, Xu Q, and Rosenbaum DM (2019) Structure of an allosteric modulator bound to the CB1 cannabinoid receptor. *Nat Chem Biol* **15**:1199–1205.

Smith RD and Carlson HA (2021) Identification of cryptic binding sites using MixMD with standard and accelerated molecular dynamics. *J Chem Inf Model* **61**:1287–1299.

Stornaiuolo M, Bruno A, Botta L, La Regina G, Cosconati S, Silvestri R, Marinelli L, and Novellino E (2015) Endogenous vs exogenous allosteric modulators in GPCRs: a dispute for shuttling CB1 among different membrane microenvironments. *Sci Rep* **5**:15453.

Vajda S, Beglov D, Wakefield AE, Egbert M, and Whitty A (2018) Cryptic binding sites on proteins: definition, detection, and druggability. *Curr Opin Chem Biol* **44**:1–8.

Wagner JR, Lee CT, Durrant JD, Malmstrom RD, Feher VA, and Amaro RE (2016) Emerging computational methods for the rational discovery of allosteric drugs. *Chem Rev* **116**:6370–6390.

Wakefield AE, Mason JS, Vajda S, and Keserű GM (2019) Analysis of tractable allosteric sites in G protein-coupled receptors. *Sci Rep* **9**:6180.

Wang J, Jain A, McDonald LR, Gambogi C, Lee AL, and Dokholyan NV (2020) Mapping allosteric communications within individual proteins. *Nat Commun* **11**:3862.

Wenthur CJ, Gentry PR, Mathews TP, and Lindsley CW (2014) Drugs for allosteric sites on receptors. *Annu Rev Pharmacol Toxicol* **54**:165–184.

Zhang D, Gao Z-G, Zhang K, Kiselev E, Crane S, Wang J, Paoletta S, Yi C, Ma L, Zhang W, et al. (2015) Two disparate ligand-binding sites in the human P2Y1 receptor. *Nature* **520**:317–321.

Zheng Y, Qin L, Zacarías NVO, de Vries H, Han GW, Gustavsson M, Dabros M, Zhao C, Cherney RJ, Carter P, et al. (2016) Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. *Nature* **540**:458–461.

Zweemer AJ, Bunnik J, Veenhuizen M, Miraglia F, Lenselink EB, Vilums M, de Vries H, Gibert A, Thiele S, Rosenkilde MM, et al. (2014) Discovery and mapping of an intracellular antagonist binding site at the chemokine receptor CCR2. *Mol Pharmacol* **86**:358–368.

Address correspondence to: Dr. John R. Traynor, Department of Pharmacology, University of Michigan, 1150 West Medical Center Drive, Ann Arbor, MI 48109. E-mail: jtraynor@umich.edu; or Heather A. Carlson, Department of Medicinal Chemistry, University of Michigan, 1150 West Medical Center Drive, Ann Arbor, MI 48109. E-mail: carlsonh@med.umich.edu