Transition-based reweighting and neural rational inference analysis of effects of new psychoactive substances on cannabinoid receptors

Soumajit Dutta[†] and Diwakar Shukla[*,†,‡,¶,§]

†*Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801*

‡*Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL, 61801*

¶*Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801*

§*Cancer Center at Illinois, University of Illinois at Urbana-Champaign, Urbana, IL, 61801*

E-mail: diwakar@illinois.edu

Abstract

New psychoactive substances (NPS) targeting cannabinoid receptor 1 (CB_1) pose a significant threat to society as recreational abusive drugs that can avoid detection and have higher physiological side effects. These physiological side effects of NPS are shown to be linked to the higher β-arrestin signaling. We hypothesize that the difference in conformational dynamics of the NPxxY motif causes the distinct downstream signaling of NPS contrary to the classical cannabinoids. To compare the dynamic effects of the NPS and classical cannabinoid binding on the NPxxY conformational ensemble, we simulate (un)binding process of NPS MDMB-Fubinaca and classical cannabinoid HU-210 from CB_1 using unbiased and biased molecular dynamics simulations. The transition-based reweighing method (TRAM) is used to combine multi-ensemble simulations for the estimation of transition rates and underlying thermodynamics of (un)binding processes of ligands with nanomolar affinities, where it is more expensive to obtain local reversible sampling. Our analyses suggest that the ligands unbind from the receptors using the same pathway but by a different mechanism. Further analyses reveal higher conformational fluctuation in the NPxxY motif for NPS bound CB_1, supporting our hypothesis. The observation is further validated using a Variational autoencoder (VAE) based on Neural rational inference, which shows higher dynamic allostery-based interactions between the binding pocket residues and NPxxY for NPS bound CB_1. Hence, in this work, MD simulation, data-driven statistical methods, and deep learning point out the significant differences in (un)binding and downstream signaling of NPS and classical cannabinoids.

Keywords

G-protein-coupled receptor, cannabinoid receptor 1, molecular dynamics, well-tempered metadynamics, umbrella sampling, Markov state model, transition-based reweighting analysis method, variational autoencoder

Introduction

Cannabinoid receptor 1 (CB_1), which is majorly expressed in the central nervous system (CNS) belongs to the class A G-protein coupled receptor (GPCR) family proteins.[1–4] GPCRs are expressed in the cellular membrane and help transduce chemical signals from the extracellular to the intracellular direction with the help of the downstream signaling proteins (G-proteins and β-arrestin).[5–7] In addition, GPCRs are the largest family of drug targets due to their substantial involvement in human pathophysiology and druggability.[8,9] Significant research efforts have been invested in the discovery of drugs targeting CB_1. CB_1 is part of the endocannabinoid signaling pathways, which helps to maintain homeostasis in neuron signaling and physiological process.[10,11]

Initial drug discovery efforts, especially the design of synthetic agonists, were based on modifying the scaffolds of phytocannabinoids (e.g., Δ^9-Tetrahydrocannabinol, cannabinol) and endocannabinoids (e.g., Anandamide, 2-arachidonoylglycerol) (Figure 1).[12–14] The synthetic molecules, which maintain the aromatic, pyran, and cyclohexenyl ring of the most common psychoactive phytocannabinoid Δ^9-THC, are known as classical cannabinoids (Figure S1).[15–17] However, the pharmacological potential of these molecules was diminished due to their psychological and physiological side effects ("tetrad" side effect).[18–20] One such example of a synthetic cannabinoid is 1,1-Dimethylheptyl-11-hydroxy-tetrahydrocannabinol (commonly known as HU-210), which is a Schedule I controlled substance in the United States.[21,22]

Apart from the canonical structures of synthetic cannabinoids, molecules with diverse scaffolds were also synthesized through structure-activity studies[23–25] . However, these molecules also lacked any pharmacological importance due to psychological side effects.[26,27] Due to the diverse structures and psychological effects, these molecules became unregulated substitutes for traditional illicit substances.[28] These synthetic cannabinoids belong to a class of molecules known as new psychoactive substances(NPS) as these molecules are not scheduled under the Single Convention on Narcotic Drugs (1961) or the Convention on

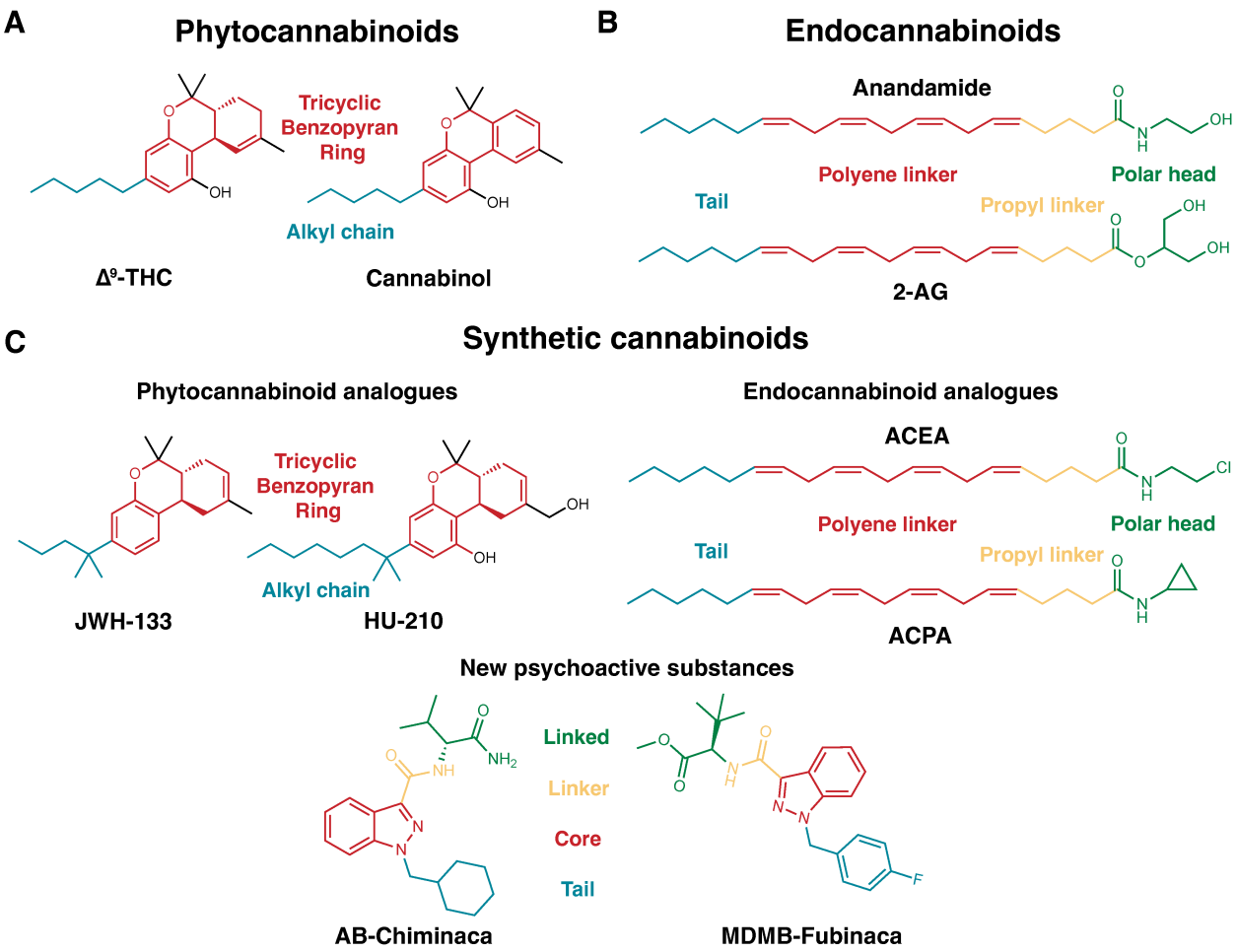

Figure 1: Classification of cannabinoid agonists: (A) Molecules derived from cannabis plants (phytocannabinoids) (B) endogenous agonists (Endocannabinoids) (C) synthetically designed molecules (Synthetic cannabinoids). Synthetic cannabinoids can be further classified based on scaffolds (phytocannabinoid analogues and endocannabinoid analogues or new psychoactive substances). Common pharmacophore groups of the ligands are shown in different colors. For phytocannabinoids and phytocannabinoid synthetic analogues, tricyclic benzopyran group and alkyl chains are colored in red and blue, respectively. Polar head group, propyl linker, polyene linker, and tail group of endocannabinoid and endocannabinoid analogues are colored with green, yellow, red, and orange, respectively. Linked, linker, core, and tail group of new psychoactive substances are colored with green, yellow, red, and orange, respectively.

Psychotropic Substances (1971).[28,29] Synthetic cannabinoids make up the largest category of NPS molecules.[30,31] NPS creates a significant challenge for drug enforcement agencies, as they appeal to drug users seeking "legal highs" to avoid the legal consequences of using traditional drugs and to be undetectable in drug screenings.[27]

The molecular structures of NPS synthetic cannabinoids consist of four pharmacophore

components: linked, linker, core, and tail groups.[27,32] The core usually consists of aromatic scaffolds (e.g., indole, indazole, Carbazole, Benzimidazole) (Figure S2).[24] The tail and linker groups are connected to the core. In the tail group, long alkyl chain-like scaffolds are ubiquitous in most NPSs; however, molecules with bulkier cyclic chains (e.g., AB-CHMINACA) are also present.[32] Frequently encountered scaffolds in linker groups are methanone, ethanone, carboxamide, and carboxylate ester groups.[33] The linker acts as a bridge between the core and the linked group. In the initial NPS synthetic cannabinoids, the linked group included polyaromatic rings; however, non-cyclic linked groups have also been identified in NPS recently.[24,32] Structural diversity in every component, while maintaining high binding affinity and potency for CB_1 make these molecules easier for drug manufacturers and harder to ban by drug enforcement agencies.[34,34–37]

The use of NPS synthetic cannabinoids has been found to cause more physiological side effects than traditional cannabinergic "tetrad" side effects.[38] These side effects include tachycardia, drowsiness, dizziness, hypertension, seizures, convulsions, nausea, high blood pressure, and chest pain.[38,39] For instance, Gatch and Forster have shown that the high concentrations of AMB-FUBINACA, the molecule which caused "zoombie outbreak" in New York, induced tremors.[40,41] A recent biochemical study has linked these discriminatory effects with the differential signaling of β-arrestin.[39] According to Finlay et al., NPS shows higher β-arrestin signaling compared to the classical cannabinoids, which has also been confirmed by other β-arrestin signaling studies.[39,42] However, mechanistic understanding of these differential downstream signaling effects between NPS and classical cannabinoids is still missing.

A recent mutagenesis study has shown that the conserved NPxxY motif of CB_1 have a larger role in downstream β-arrestin signaling than G-protein signaling.[43] However, structural comparison of the classical cannabinoid (AM841) and NPS (MDMB-FUBINACA) bound active CB_1-G_i complex shows a conformationally similar NPxxY motif (Figure 2).[44,45] Furthermore, the NPxxY motif of β-arrestin bound β_1-AR shows structural similarity with classical cannabinoid bound structure (Figure S3).[46] Therefore, we hypothesized that the

differential binding pocket interactions of NPS binding alters the dynamics of NPxxY motif allosterically compared to the classical cannabinoids. This phenomenon in turn leads to differential signaling for β-arrestin between NPS and Classical Cannabinoids.

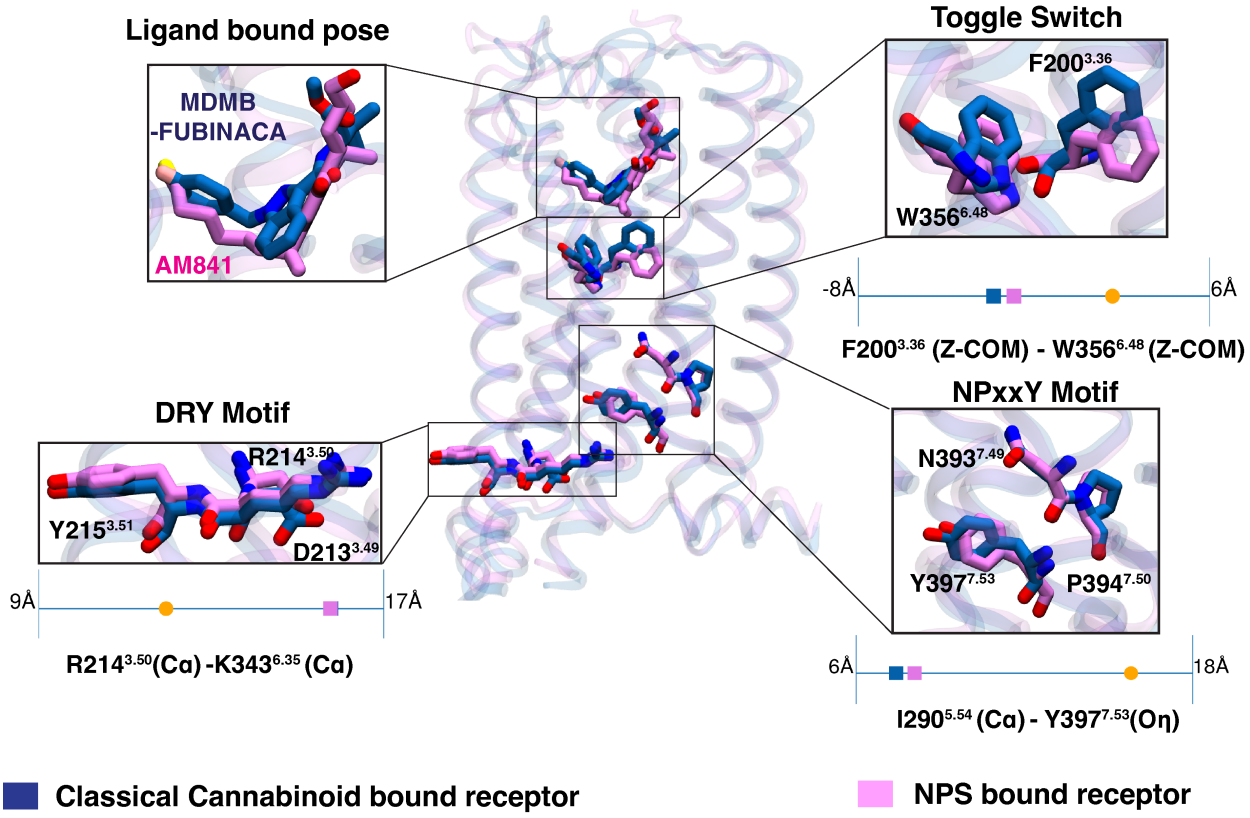

Figure 2: NPS bound CB_1 (PDB ID: 6N4B,[44] color: Blue) structure is superposed with the classical cannabinoid bound CB_1 (PDB ID: 6KPG,[45] color: Purple). Both structures are in G_i bound active state. Proteins are shown in transparent cartoon representation. Structural comparison of conversed activation metrices (Toggle switch, DRY motif, and NPxxY motif) and ligand poses are shown as separate boxes. Quantitative values of the activation metrics for both active structures are compared as scatter points on 1-D line with the CB_1 inactive structure (PDB ID: 5TGZ,[1] color: orange). These quantitative measurements were discussed in Dutta and Shukla[4]

.

To study these distinct dynamic effects, we compared the (un)binding of the classical cannabinoid (HU-210) and NPS (MDMB-FUBINACA) from the receptor binding site. These molecules have nanomolar affinities. Obtaining the initial pathway of ligand unbinding from unbiased sampling will be computationally expensive. Therefore, a well-tempered metadynamics approach was used to sample the unbinding event, where a time-dependent biased potential is deposited for the faster sampling of the metastable minima along the pathway.[47]

However, a detailed characterization of the unbinding processes is only possible through the thermodynamics and kinetics estimation of intermediate states. Thus, a transition operator-based approach is needed, which helps to estimate the transition timescale between the states and the stationary density of each state. Estimation from these approaches usually depends on the equilibrium between the local states, which can only be maintained by reversible sampling. For high-affinity ligands like MDMB-FUBINACA and HU-210, reversible sampling is expensive as ligands move from high energy unbound states to lower energy bound states irreversibly. Hence, we implemented a transition operator approach named the transition-based reweighting analysis (TRAM) method, which can tackle this lack of local equilibrium between states by combining unbiased and biased approaches.[48] TRAM has been used different simulation studies for estimating thermodynamics and kinetics of high energy processes. Characterization of the dimerization kinetics and sodium ion translocation of μ-opioid receptors have been studied using TRAM.[49,50] Furthermore, TRAM have been utilized for thermodynamics and kinetics estimations of ligand (un)binding for kinase and Serine protease.[48,51] To implement TRAM for our study, extensive sampling of the (un)binding process of both ligands was performed using a combination of umbrella sampling and unbiased simulations from the pathway obtained using metadynamics (see Methods section).[52] We showed that TRAM can produce consistent kinetic estimation with less unbiased simulation data compared to traditional methods like the Markov state model.[53]

Based on estimates of thermodynamics and kinetics, it was observed that both NPS and classical cannabinoids have similar unbinding pathways. However, their unbinding mechanisms differ due to the aromatic tail of the MDMB-FUBINACA compared to the alkyl side chain of HU-210. A comparison of downstream signaling for NPS and classical cannabinoid bound systems shows a significant difference in the dynamics of the NPxxY motif. Higher fluctuation of NPxxY motif in NPS bound systems is caused due to a lack of interaction between NPS ligands and TM7, especially with conserved polar interactions with S383$^{7.39}$. Allosteric strength between binding pocket residues and NPxxY motif was estimated with

the deep learning technique, Neural relational inference (NRI).[54] NRI network revealed that binding pocket residues of NPS bound ensemble have higher allosteric weights for the NPxxY motif compared to classical cannabinoids. These analyses validate our hypothesis that the differential dynamic allosteric control of the NPxxY motif might lead to the β-arrestin signaling for different ligands. This study provides a foundation for additional computational and experimental research to enhance our understanding of the connection between ligand scaffolds and downstream signaling. This knowledge will assist drug enforcement agencies in proactively banning these molecules and inform policies that can protect individuals from the effects of abuse.

Results and Discussion

Metadynamics simulations capture the unbinding paths of NPS and classical cannabinoids. The representative classical cannabinoid and NPS selected for this study are HU-210 and MDMB-FUBINACA.[22] Compared to Δ^9-THC, HU-210 has an extra hydroxyl group in the C-11 position and a 1',1'-Dimethylheptyl group instead of a pentyl side chain (Figure 1C and S1). MDMB-FUBINACA is a derivative of AB-FUBINACA, which was originally developed by Pfizer (Figure 1C).[44] These ligands binds to CB_1 receptor with nanomolar affinities (MDMB-FUBINACA K_i : 1.14 nM;[55] HU-210 K_i : 0.61 nM[14,56]).

Metadynamics simulation is a biased sampling method and has been widely used in protein-ligand binding and unbinding studies, as preexisting knowledge of the pathway is not necessary for performing these simulations.[57–60] In metadynamics, a time-dependent biased potential is deposited into the sampling process for the ligand to get out of stable minima at a faster pace.[61] In this work, two replicates of well-tempered metadynamics were performed to capture the unbinding pathway of HU-210 and MDMB-Fubinaca. The commonly used collective variables were selected for metadynamics simulations: (1) z component distance between the center of mass of ligand and residue in the ligand binding pocket ($W356^{6.48}$),

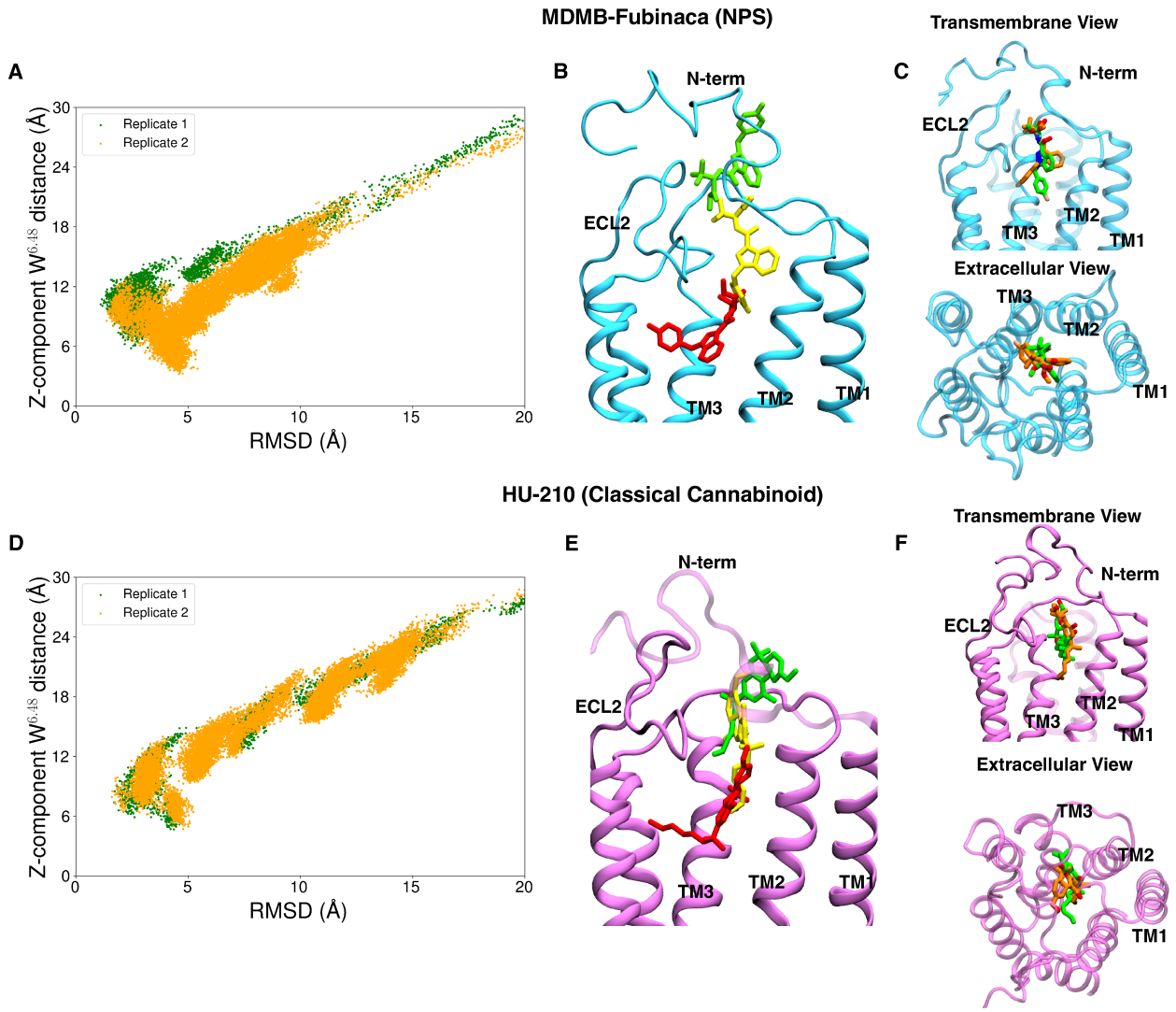

Figure 3: Unbinding ensemble of the MDMB-FUBINACA (A) and HU-210 (D) are projected as 2-D scatter plots where Z-component distance of ligand center of mass from W3566.48 is plotted against the ligand RMSD. These ensembles were obtained by running well-tempered metadynamics. Unbinding pathway of MDMB-FUBINACA (B) and HU-210 (E) obtained from the metadynamics. Ligand along the pathway is shown with different color in stick representation. The representative frames of both replicates from the unbinding ensembles are shown where the MDMB-FUBINACA (C) and HU-210 (F) are dissociating from the receptor. Both transmembrane (left panel) and extracellular (right panel) views are displayed here. Proteins and ligands are represented as cartoon and sticks, respectively.

and (2) Contact number with the ligand heavy atom and α carbon of all binding pocket residues (Equation 4).

The z-component distance was plotted against the RMSD of the ligands from the bound pose, which indicates that ligands follow a similar pathway for each replica (Figures 3A and

3C). It is also observed that the dissociation happens via the opening formed by TM2, TM3, ECL2, and N-terminus for both ligands (Figures 3B, 3C, 3E, and 3F). These observations indicate that the pathway may be the minimum free energy pathway for the ligand unbinding in CB_1. Previous metadynamics binding simulation of another cannabinoid ligand also points to a similar pathway.[60] Reweighted probability density obtained from metadynamics calculation shows one highly dense region in the pocket, depicting the stability of the bound pose of the ligands (Figures S4A and S4B). However, time dependent external force applied during the metadynamics makes the sampling in the orthogonal direction of the CVs less extensive. Thus, the biased simulation might not sample some protein-ligand interactions that helps to characterize intermediate states. To properly sample intermediate transition states during the underlying unbinding process, extensive unbiased simulations were performed from the initial pathway obtained from the limited sampling of the well-tempered metadynamics (discussed below).

Comparison of thermodynamics and kinetics estimates from Markov state Model and Transition-based reweighting analysis method. MSM and TRAM are both postprocessing techniques for estimating the kinetics and thermodynamics of underlying physical processes observed in MD simulation. MSM is applied to reversible equilibrium simulations, whereas TRAM estimations can be obtained from multi-ensemble simulations (combination of biased and unbiased simulations). The MSM depends on the local equilibrium between the Markovian states, which is also known as detailed balance (Equation 5). However, reversible local sampling becomes challenging with short parallel trajectories when the free energy difference between two local Markovian states is high. In those cases, reversibility is still assumed by forcing the detailed balance when estimating the transition probability matrix.[53] This leads to the incorrect estimation of the unbinding kinetics due to limited sampling from the stable bound state to the high energy unbound states.[48] Refining the state discretization (i.e., increasing the number of states) may resolve the issue. However, refined state discretization sometimes decreases the statistically significant tran-

sition count between all states, decreasing the model certainty. TRAM was shown to solve this problem by combining biased and unbiased simulations (see Methods section). Biased simulations (e.g., replica exchange, umbrella sampling) help to enhance the local sampling, either by increasing the temperature for faster sampling or by fixing collective variables with biased potential to have better sampling in orthogonal directions. It has been shown that compared to MSM, kinetics predicted using TRAM from the combination of biased and unbiased simulations are more aligned with the experiment results.[48]

As unbinding of ligands with high binding affinity (nanomolar) are being studied here, asymmetric transitions might be observed along the pathway. Therefore, we compared the use of MSM and TRAM in estimating the kinetics and thermodynamics of the (un)binding process. For TRAM, unbiased simulation and umbrella sampling were run from the clusters in conformational ensemble obtained from metadynamics (refer to the Methods section for more details). For MSM estimation, only unbiased simulations starting from the metadynamics pathway were considered.

For thermodynamics comparison, standard free energy was estimated for the ligands considering volume correction.[62] TRAM and MSM predictions of standard binding free energy are within 0.6 kcal/mol of each other for each ligand (Figure 4A). Although the absolute binding free energy differs from the experimentally predicted value by approximately 3 kcal/mol, the relative estimated free energy ($\Delta\Delta$G) values are also within 0.6 kcal/mol of the experimentally determined values. Therefore, it indicates that with sufficient sampling, both MSM and TRAM converge to the same predictions of relative free energy.

We also compared the kinetics obtained from the MSM and TRAM. Kinetic measurements were performed with transition path theory (TPT), which uses transition probability matrix from MSM or Tram to estimate mean free passage time between different states (see Methods section). Estimated binding times using TRAM and MSM match perfectly for both ligands. The estimated dissociation times are within one order of magnitude with each other. These observations agree with the previously reported computational research,

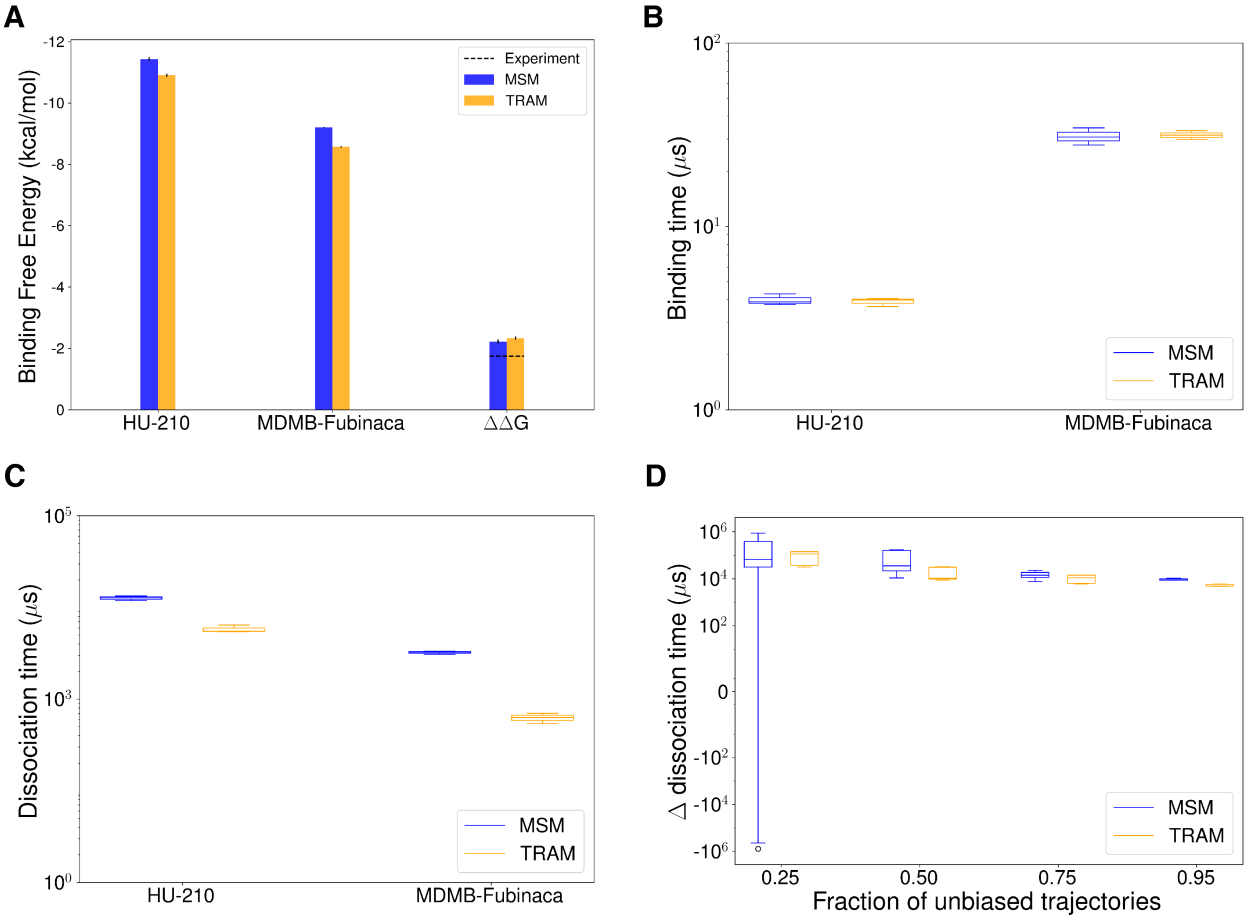

Figure 4: (A) The bar plot represents standard binding free energy for HU-210, MDMB-FUBINACA, and difference of standard binding free energy between the ligands. MSM and TRAM estimations are shown as blue and orange bars, respectively. Experimentally predicted values are shown as dotted line. (B, C) Binding (B) and dissociation (C) time for HU-210 and MDMB-FUBINACA are shown as box plots. (D) Difference in dissociation time of the two ligands is plotted as box plot against fraction of unbiased trajectories used for the estimation. This timescales were obtained from the mean free passage time calculation using TPT with transition probabilities estimated from MSM (color: blue) and TRAM (color: orange). Errors were calculated using bootstrapping method with 3 bootstrapped samples.

where experimentally comparable estimation of k_{off} rates were shown to be more challenging compared to k_{on}.[63]

Further analyses were performed to compare these methods in the low unbiased data regime. The difference between the dissociation time of ligands was measured with different amounts of unbiased data. It is observed that even with only 25% of original, unbiased data, TRAM can predict the kinetics within an order magnitude of the kinetics estimated

with full dataset (Figure 4D). On the other hand, error in MSM predicted kinetics more rapidly compared to TRAM with lesser amount of unbiased data. A similar trend can be observed for $\Delta\Delta$G prediction (Figure S5). Therefore, TRAM provides better predictions of thermodynamics and kinetics when less amount of unbiased data.

Unbinding mechanism for new psychoactive substance. Although the binding position of the ligand and the overall binding pathway are similar for both the ligands, extensive biased and unbiased simulation analyzed by TRAM shows a significant difference in the unbinding mechanism of the ligands. To capture the unbinding pathway for MDMB-FUBINACA, we projected the TRAM weighted free energy landscape of the distance between the linked part of the ligand (Leucinate group) and TM5 with respect to the distance between the ligand tail group and TM7 (Figure 5A). The free energy landscape was divided into the non-overlapping intermediate macrostates to obtain better description of the unbinding process. In each macrostate, the contact frequency of ligand with binding pocket residues were calculated along with corresponding contact energies. A metastable minima is observed for macrostate representing the bound pose of the ligand depicting the stability of the ligand (Figure 5A). In the bound pose, the major interactions form between the aromatic (F170$^{2.57}$, W279$^{5.43}$, F268$^{\mathrm{ECL2}}$) and hydrophobic (L193$^{3.29}$) residues of the binding pocket (Figures 6A, 6B, and S6B).

The free energy landscape shows two probable mechanisms for MDMB-FUBINACA unbinding from the bound pose. The two pathways are differentiated by whether the linked or tail part of MDMB-FUBINACA dissociates first. One of the pathways, aromatic tail part of MDMB-FUBINACA moves away from TM5 and form interaction with aromatic residues in TM2 (F170$^{2.57}$ and F174$^{2.61}$) (Figures 5A, 6A, 6B, and S6C). This leads to the formation of intermediate metastable states, which we characterize as macrostate Intermediate state 1 (I1). This metastable minimum observed from I1 macrostate might be unique to the FUBINACA family of NPS synthetic cannabinoids as this family has the aromatic ring in tail group, unlike the long alkyl chain in other common synthetic cannabinoids. Along with the

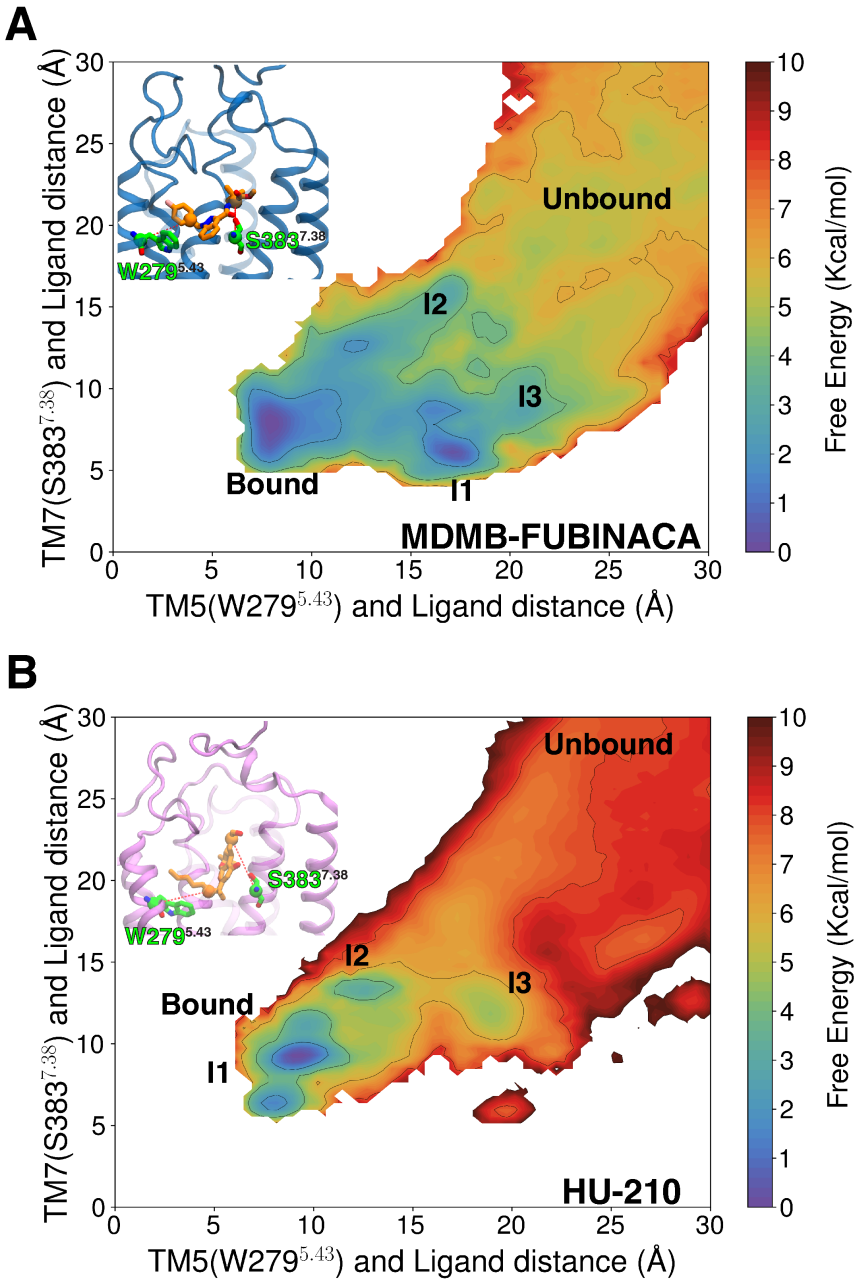

Figure 5: TRAM weighted Two dimensional projection of unbinding free energy landscape for MDMB-FUBINACA (A) and HU-210 (B). For MDMB-FUBINACA, distance between TM5 (W279$^{5.43}$-Cα) and tail part of the ligand is plotted against the distance between TM7 (S383$^{7.39}$-Cα) and ligand linked part. For HU-210, distance between the TM5 (W279$^{5.43}$-Cα) and tail is plotted against the TM7 (S383$^{7.39}$-Cα) and cyclohexenyl ring of the ligand. Measured distances are shown as red dotted lines in the inset figures. Macrostate positions are shown on the landscapes.

aromatic residues of TM2, major interaction with F268$^{\mathrm{ECL2}}$ is maintained in macrostate I1 (Figures 6A and 6B). K-L divergence analysis between bound and I1 macrostates show that

only minor changes in the binding pocket residues, especially in TM2 are needed to accommodate MDMB-FUBINACA in this conformational state (Figure 6C). The interconversion timescale (MFPT) between the macrostates were obtained from the transition path theory. MFPT calculations show that both the timescales are similar with slightly higher timescales for the bound pose compared to I1 transition ($20.6 \pm 2.3\mu s$) (Figure 6B). In this pathway, ligand moves from I1 metastable state to space between N-terminus, TM2, TM3, and ECL2 before dissociating from the receptor (Figure 6B). This region between the unbinding ensemble have been characterized as macrostate I3 (Figure 5A). Contact analysis show significant drop in ligand residue contacts with only aromatic residues in TM2, TM3 and ECL2 forming dominant interactions (Figures 6A and S6D). K-L divergence shows that ligand positioning in this particular regions causes relatively higher divergence on TM2 compared to I1 (Figure 6C). Kinetically, the transition from I1 to I3 ($33.7 \pm 3.1\mu s$) is much slower compared to reverse transition ($0.8 \pm 0.0\mu s$), validating the higher stability of the I1 compared to I3 macrostate (Figure 6B). According to the TPT analysis, breaking the aromatic interactions for complete dissociation of MDMB-FUBINACA requires $\sim 371.9 \pm 40.2\mu s$, making it the slowest step in this pathway (Figure 6B).

In the other possible unbinding pathway, orientation of MDMB-FUBINACA in the pocket doesnot change compared to the bound pose. The linked part of the ligand moves to space between N-terminus, TM2, TM3, and ECL2 (Figure 6B). We label this macrostate as I2. In this state, we observe stable polar interaction with K376$^{7.32}$ and hydrophobic interactions with aromatic and other hydrophobic residues (F177$^{2.64}$, F268$^{\mathrm{ECL2}}$, P269$^{\mathrm{ECL2}}$) (Figures 6A and S6A). However, free energy of this macrostate is higher than the bound pose, depicting higher entropic cost associated with this state. This can be shown by the higher intrastate RMSD of I3 compared to the bound pose (Figure S7). Transition timescale from the bound pose to the I2 ($136.7 \pm 13.9\mu s$) is one order of magnitude higher compared to the reverse transition ($9.8 \pm 13.9\mu s$) (Figure 6B). K-L divergence analysis also shows higher divergence in the extracellular region of TM2 and N-terminus compared to bound pose (Figure 6C).

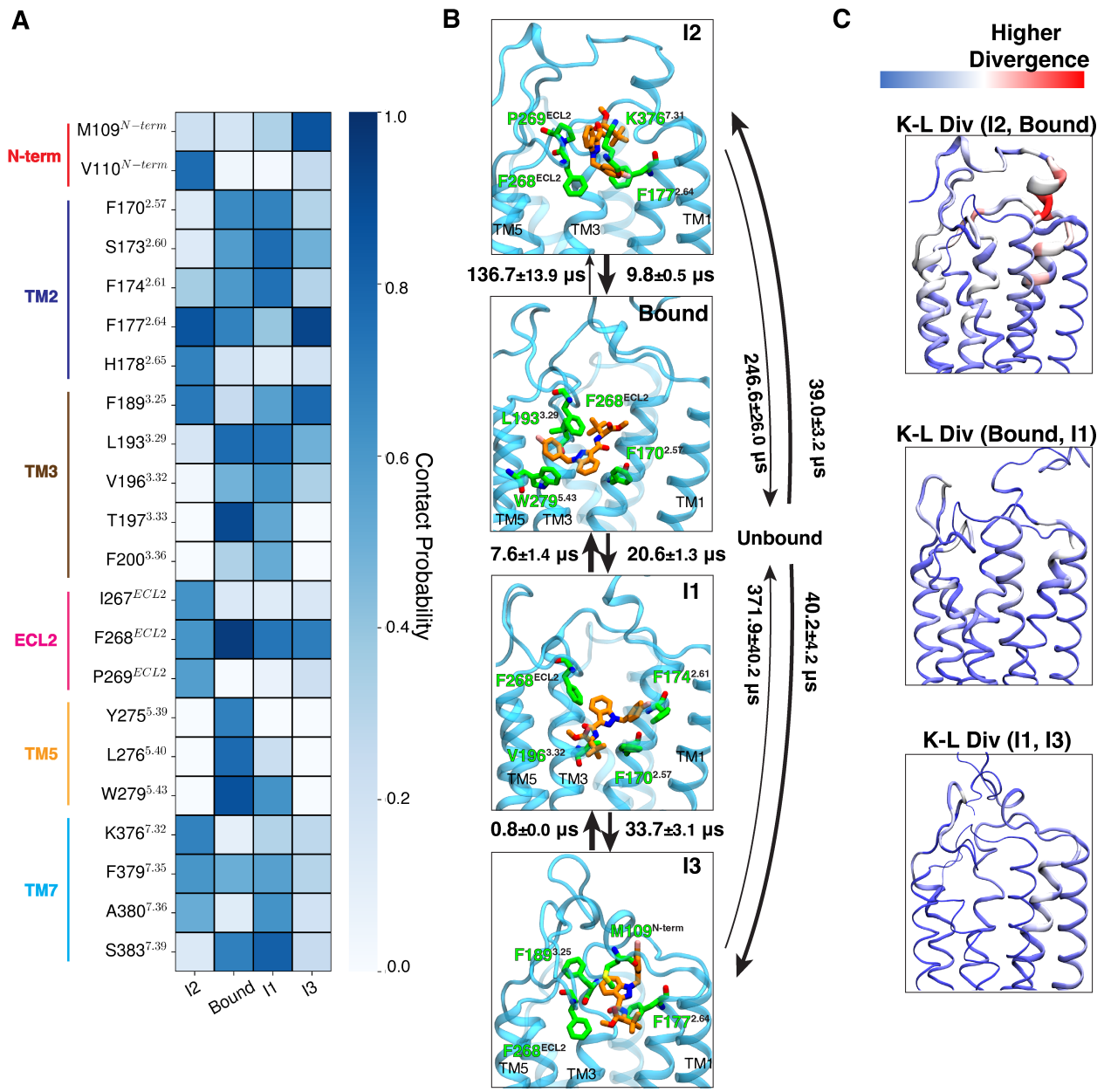

Figure 6: (A) The contact probabilities with binding pocket residues of MDMB-FUBINACA are shown as a heatmap for different macrostates, where ligand maintains contact with the receptor. Residues in different structural elements (loops and helices) are distinguished by distinct color bars. (B) Representative structures are shown where ligand (color: orange) and four residues (color: green) with highest interaction energies are shown as sticks. Proteins are shown as purple cartoon. Timescales between interstate transitions are shown as arrows. Arrow thickness is inversely proportional to the order of magnitude of the timescale. (C) K-L divergence between protein conformations of different states are shown with color (blue to red) and thickness (lower to higher) gradient. Thickness gradient are shown as rolling average to highlight a region of high K-L divergence. Errors in MFPT calculations were estimated based on 3 bootstrapped TRAM calculation with randomly selected 95% of unbiased trajectories.

Dissociation of MDMB-FUBINACA from I2 to the bulk is faster compared to dissociation from I3 ($246.6 \pm 26.0 \mu s$) (Figure 6B). However, the overall kinetic barrier for dissociation from the binding pose for the both unbinding mechanisms are relatively similar.

Unbinding mechanism for classical cannabinoid. For capturing the classical cannabinoid (un)binding mechanism, distances from the two terminal scaffolds (cyclohexenyl and alkyl chain) to TM5 and TM7 were measured similar to the NPS (Figure 5B). The free energy landscape of the unbinding of the HU-210 shows the differences in the mechanism from MDMB-FUBINACA. Similar to MDMB-FUBINACA, the HU-210 unbinding landscapes were also divided into non-overlapping macrostates. Macrostate representing HU-210 bound pose shows a metastable energy minimum. Comparing the bound macrostate interactions of MDMB-FUBINACA, classical cannabinoid HU-210 shows higher interactions with TM7 residues ($S383^{7.39}$, $F379^{7.35}$) (Figures 6A, 7A, S6B and S8B). Previous experimentally determined structures of classical cannabinoid bound CB_1 have pointed out these conserved polar interactions of the hydroxyl group at C-1 position with $S383^{7.39}$.[45,64] Although MDMB-FUBINACA also maintain this polar interaction with carboxylic oxygen, the interaction energy for the HU-210 is much higher, depicting the importance of this conserved residue in stabilizing classical cannabinoids (Figure 6B). Mutagenesis studies also support this difference in interaction with $S383^{7.39}$ between the classical cannabinoids with hydroxyl group (HU-210) and CB_1 ligands, which have carboxylic oxygen in the equivalent position (WIN-55,212-2).[65,66] Alanine mutation of $S383^{7.39}$ have shown to decrease the ligand affinity and downstream efficacy of classical cannabinoids by orders of magnitude, while having minimal effect on WIN-55,212-2, which have carboxylic oxygen in the linked part as MDMB-FUBINACA.[65,66] Other major interactions ($F170^{2.57}$ and $F268^{ECL2}$) in the bound pose are common between the two ligands (Figures 6B and 7B).

A relatively weaker metastable state is observed when the ligand moves relatively deeper (closer to TM5) inside the binding pocket. The flexible alkyl chain of HU-210 allows the ligand to have this deeper position (Figures 7A and 7B). Protein-ligand interaction analysis

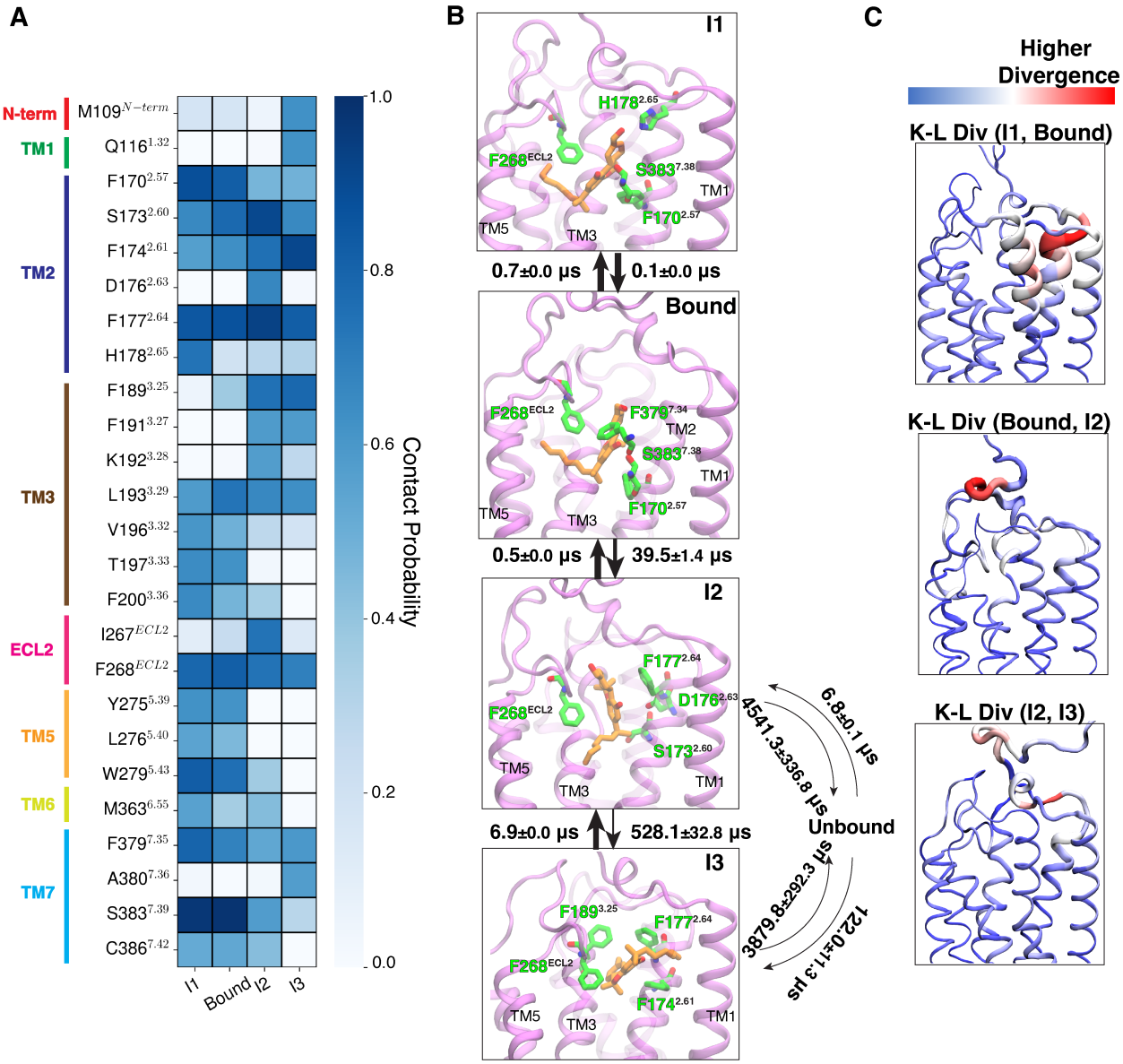

Figure 7: (A) The contact probabilities with binding pocket residues of HU-210 are shown as a heatmap for different macrostates, where ligand maintains contact with the receptor. Residues in different structural elements (loops and helices) are distinguished by distinct color bars. (B) Representative structures are shown where ligand (color: orange) and four residues (color: green) with highest interaction energies are shown as sticks. Proteins are shown as purple cartoon. Timescales between interstate transitions are shown as arrows. Arrow thickness is inversely proportional to the order of magnitude of the timescale. (C) K-L divergence between protein conformations of different states are shown with color (blue to red) and thickness (lower to higher) gradient. Thickness gradient are shown as rolling average to highlight a region of high K-L divergence. Errors in MFPT calculations were estimated based on 3 bootstrapped TRAM calculation with randomly selected 95% of unbiased trajectories.

in the macrostate representing this region (I1) show that hydroxyl group at C-11 forms a major polar interaction with H178$^{2.65}$ (Figure S8A). The bound and I1 macrostates are kinetically close, as indicated by the rapid interconversion between these states (Figure 7B). K-L divergence between the two states show the highest divergence in extracellular TM2 and TM7, where major interaction switch have happened (Figure 7C).

Contrasting to MDMB-FUBINACA, only one pathway was discovered with classical cannabinoid cyclic scaffold departing from the receptor first. Major interactions that break when the ligand moves out of the binding pose to macrostate I2 is the polar interaction with S383$^{7.39}$ and hydrophobic interaction with aromatic F170$^{2.57}$ (Figures 7A, 7B, and S8C). Breaking of these bonds leads to larger kinetic barrier of approximately $39.5 \pm 1.4\mu s$ (Figure 7B). In this macrostate, the HU-210 forms major interactions with aromatic residues F268$^{\mathrm{ECL2}}$ and F177$^{2.64}$ and polar interactions with S173$^{2.60}$ and D176$^{2.63}$ (Figure 7B). From this pose, HU-210 either dissociates from the receptor or obtain another relatively weak stabilized state (I3) in the receptor. In I3, the alkyl chain of the ligand is flipped in the pocket and stabilized by aromatic residues in TM2, TM3 and ECL2 (Figure 7B). This transition from I2 to I3 ($528.1 \pm 32.8\mu s$) kinetically much slower compared to the reverse transition ($6.9 \pm 0.0\mu s$) (Figure 7B). From both I2 and I3 macrostates, the ligand can dissociate from the pocket and mean free passage time for these transitions appear to be in the milisecond timescale, which is one order magnitude higher compared to the MDMB-FUBINACA unbinding (Figure 6B). This phenomena supports the relatively high affinity of the classic cannabinoid HU-210 compared to the NPS MDMB-FUBINACA.

Neural rational inference analysis shows distinct downstream signaling between new psychoactive substances and classical cannabinoids. It was hypothesized that distinct physiological side effects of NPS, compared to cannonical "tetrad" side effects of classical cannabinoids, was caused by higher β-arrestin signaling of NPS. Mechanistic understanding of distinct downstream signaling of NPS and classical cannabinoids is still missing. Additional mutagenesis experiments have shown that β-arrestin signaling for CB_1 are linked

to NPxxY motif (Figure 2). Therefore, we hypothesized that NPS binding might lead to differential allosteric control of NPxxY motif movement to assist β-arrestin recruitment.

To observe the difference in downstream signaling due to binding of NPS and classical cannabinoids, we compared the conformational ensemble of TM6 movement and NPxxY, two cannonical matrices of class A GPCR activation. TRAM weighted probability density plot shows that TM6 outward movement have identical distribution for HU-210 and MDMB-FUBINACA (Figures 8A and 8B). In contrast, NPxxY motif have distinct distributions for two ligands. NPS binding leads to higher NPxxY motif RMSD compared to classical cannabinoid ligands, validating the dynamic effect of the NPS ligand on the NPxxY motif.

we adopted a data-driven deep learning network known as Neural relational inference (NRI) to further validate our hypothesis of allosteric control. NRI network has an architecture of variational autoencoder. The encoder part of the network predicts the interactions between the residues from the trajectory dynamics, and the decoder predicts the trajectories from the interaction. With this network, we try to produce alpha carbon coordinates at $t+\tau$ from the coordinates at time t. In the process of regenerating the future coordinates, the latent space of the network learns the dynamic interactions between different residues in the protein. These interactions are calculated from the estimated posterior probability $q(z_{ij}|x)$. In this work, we trained the network from 10 unbiased trajectories (100 ns per trajectory), the NPS, and classical cannabinoid-bound unbiased trajectories (Method Section). Here, we compared the allosteric interaction weights between the binding pocket and the NPxxY motif. Results show that each binding pocket residue in NPS bound ensemble shows higher allosteric weights with the NPxxY motif, indicating larger dynamic interactions between the NPxxY motif and binding pocket residues.

To explain larger dynamics in the NPxxY motif for NPS bound receptor, we propose that it may be caused by relatively weaker interactions with TM7 binding pocket residues, as discussed previously. To validate this hypothesis, we projected NPxxY motif RMSD at the bound macrostate with respect to the S383$^{7.39}$ (Oγ) distance from the closest oxygen

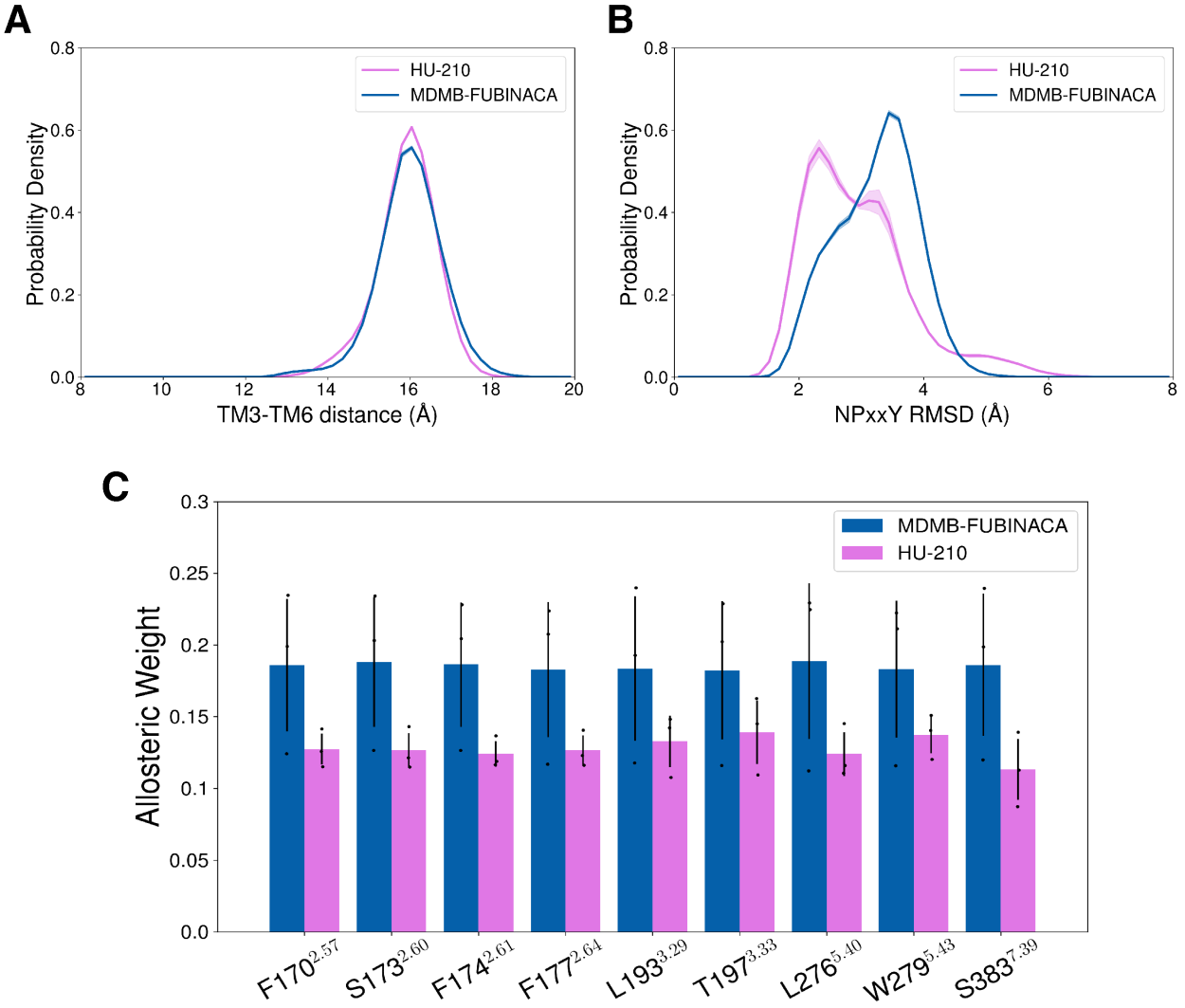

Figure 8: (A) TRAM weighted probability density of TM3 ($R214^{3.50}$) and TM6 ($K343^{6.35}$) distance distribution are plotted for HU-210 (color: purple) and MDMB-FUBINACA (color: blue) unbinding ensemble. (B) TRAM weighted probability density of NPxxY RMSD distribution are plotted for HU-210 (color: purple) and MDMB-FUBINACA (color: blue) unbinding ensemble. Error in the probability densities is estimated using bootstrapping approach, where TRAM was built for 3 bootstrapped samples with 95% of total data. (C) Allosteric weights between the binding pocket residues and the NPxxY motif are plotted as bar plots for HU-210 (color: purple) and MDMB-FUBINACA (color: blue) bound trajectories. Allosteric weights are estimated from the posterior probability of the NRI network. Error in the allosteric weights is calculated by training the network 3 different training data.

atom of both ligands (Figure S9). These oxygen atoms for HU-210 and MDMB-FUBINACA are the hydroxyl oxygen at C1 carbon and acyl oxygen at the linker group, respectively. The projection shows that this distance, which represents the conversed polar interaction, is

much higher for MDMB-FUBINACA bound ensemble, causing the weaker interaction for this ligand. It is also observed the NPxxY motif RMSD increases with the larger distance of this polar interaction (Pearson correlation coefficient: 0.32), suggesting that weaker interactions with TM7 in the binding pocket lead to higher fluctuations in the downstream NPxxY motif of TM7.

In order to show the universality of higher allosteric effects on NPxxY motif of NPS bound systems, unbiased microsecond long simulations were performed for three other classical cannabinoids (AMG-41, JWH-133, and O-1317) and NPS (AMB-FUBINACA, CUMYL-FUBINACA, 5F-AMP) (Figure S10A). Simulation revealed that NPS linker oxygen distance from with S383$^{7.39}$ (Oγ) are higher compared to the classical cannabinoid hydroxyl oxygen and S383$^{7.39}$ (Oγ) distance (Figure S10B). Although classical cannabinoid JWH-133 doesnot contain hydroxyl oxygen, equivalent hydrogen atom still maintains shorter distance compared to NPS linker oxygen. These validates the lesser polar interactions of NPS ligands with TM7. These ligand bound unbiased simulations were used to train NRI network to estimate the allosteric weights between the binding pocket residues and NPxxY motif (Figure S10B). Results show that NPS ligands show an relatively higher allosteric weights over the classical cannabinoids. These observations bolster the claim of higher the β-arrestin signaling for NPS ligands.

Conclusions

Synthetic cannabinoids were designed as a potential therapeutics to target cannabinoid receptors. However, major side effects of these ligands diminish their therapeutic potential. Although both classical cannabinoids and NPS synthetic cannabinoids have been abused as recreational drugs, later poses larger threats for the society due the chemical diversity of the NPS structures makes them harder to control from being abused. Furthermore, physiological studies have shown NPS targeting cannabinoid receptors lead to the dangerous physiologi-

cal effects compared to "tetrad" side effects associated with classical cannabinoids. Previous studies have relate these side effects with the higher downstream β-arrestin signaling of NPS. Mutagenesis studies have shown NPxxY motif have larger role to play in β-arrestin signaling. In this work, we proposed that NPS and classical cannabinoid have distinct allosteric control on NPxxY motif when bound to orthostetic pocket of CB_1. In this hypothesis driven study, we compared ligand-protein interactions of NPS MDMB-FUBINACA and classical cannabinoid HU-210 for CB_1 by studying their unbinding mechanism and downstream signaling.

As both ligands are stable binders with nanomolar affinity, well-tempered metadynamics simulations were performed to obtain the initial unbinding pathway. These simulations were able to find similar pathway via the opening formed by TM2, TM3, N-terminus, and ECL2 which matches with previous metadynamics binding simulations for cannabinoid receptors. For the proper characterization of intermediate states, the unbinding processes were further extensively sampled using unbiased simulation and umbrella sampling.

Effectiveness of the post processing techniques TRAM and MSM were compared in predicting the ligands with binding affinity and kinetics. MSM predicts the kinetics and thermodynamics from the eigendecomposition of the tranisition probability matrix. MSM assumes that the local equilibrium is maintained between the Markovian states. However, with limited sampling, this creterion may not valid between local high and low energy states. TRAM tries to solve this issue by combining biased and unbiased simulation, where biased simulations enhances the local sampling to maintain the equilibrium. We observed that with sufficient data both methods performed in similar way in estimating the standard binding free energy. The relative free energy estimated by both methods match with the experimental result within 0.6 kcal/mol. With lesser amount of unbiased data, TRAM predictions of kinetics and thermodynamics remain more consistent than the MSM as the biased simulations help to maintain local equilibrium.

TRAM estimated thermodynamics helped to decipher the differences between the unbinding of NPS MDMB-FUBINACA and classical cannabinoid HU-210. First, for MDMB-

FUBINACA, a larger conformational change is observed within the pocket. A metastable intermediate state is observed when the aromatic tail of FUBINACA flip inside the pocket and reorient itself close to the aromatic residues of TM2. It was observed that both linked part and tail part of the ligands can lead the dissociation of the ligand from the receptor. Second, for HU-210, conserved cyclic group leads the dissociation from the receptor. It supports previous simulation where the alkyl side chain of the ligand binds to the receptor first.[17] Third, aromatic residues in the pocket ($F268^{ECL2}$, $F170^{2.57}$) form major interactions with both classical cannabinoid and NPS molecules. Major differences in protein-ligand interactions were observed in TM7. Stronger interactions are observed for the classical cannabinoid with TM7, especially polar interaction with $S383^{7.39}$ and hydrophobic interaction with $F379^{7.35}$ compared to NPS molecules.

Finally, downstream signaling comparison shows the difference in NPxxY motif for NPS and classical cannabinoids. NPxxY motif of NPS bound system undergoes larger dynamics which can be explained by weak interaction with TM7. Further validation of these dynamic allosteric communications were performed with variational autoencoder (NRI). NRI network learns the dynamic interactions between residues in the latent space by learning to reconstruct the dynamics. Dynamic allostery measured by the posterior probability of VAE shows higher allosteric weights from the binding pocket residues to the NPxxY motif region for NPS molecules. These observations support the higher β-arrestin signaling of these molecules. Therefore, these data driven computational study helps us to distinguish between the receptor-protein interaction, unbinding mechanism and downstream signaling NPS compared to other classical cannabinoids.

Methods

System Preparation

For NPS unbinding simulation, G-protein bound active structure (PDB: 6N4B[44]) was selected as the initial structure. G-protein subunits and Non-Protein residues other than orthosteric ligand MDMB-Fubinaca were removed from the PDB structure file. Missing residues in ICL3 (21 residues, 314-334) and ECL2 (6 residues, 258-263) were modeled sequentially using Remodel protocol of Rosetta loop modeling.[67,68] In each step, the remodeled structure with least energy was further refined using kinematic closure protocol.[69]

Starting 108 residues from CB_1 N-terminus were also missing from the cryo-EM structure. However, it is not feasible to model the entire N-terminus because of two following reasons.[67] First, a proper template is not available for modeling N-terminus regions as most of the class A GPCRs do not contain large N-terminus.[70] Second, it is challenging to model these large numbers of residues accurately with template-free *ab initio* modeling because of the combinatorial expansions of conformational space. Therefore, the closest 20 residues (89-108) were modeled as membrane proximal regions of the N-terminus were shown to be important for CB_1 signaling by allosterically modulating ligand affinity.[71] Furthermore, $\Delta 89CB_1$ (CB_1 with first 88 residues truncated in N-terminus) have similar ligand binding affinity compared to CB_1 with full sequence.[72] Modeling of this membrane proximal region was also performed Remodel protocol of Rosetta loop modeling. A distance constraint is added during this modeling step between $C98^{N-term}$ and $C107^{N-term}$ to create the disulfide bond between the residues.[71,73]

The "Ligand Reader & Modeler" module of CHARMM-GUI was used for ligand (e.g., MDMB-Fubinaca) parameterization using CHARMM General Force Field (CGenFF).[74] The ligand bound receptor was embedded in the bilayer membrane and salt solution (extracellular and intracellular region) using CHARMM-GUI.[75] As CB_1 is majorly expressed in central nervous system, an average brain membrane composition of asymmetric complex membrane

was selected. The membrane composition was obtained from Ingólfsson et al. and proportionally downsized according to our system (Table S1).[76] 150 mM NaCl salt solution with TIP3P water model was used in the extracellular and intracellular regions.[77] CHARMM36m forcefield was used to parameterize the protein, lipid, water, and ions.[78]

For building the classical cannabinoid system, the modeled PDB structure was used. In this case, the orthosteric NPS ligand was removed, and a classical cannabinoid HU-210 was docked into the orthosteric pocket using Autodock Vina.[79] The docked bound pose was selected based on best overalled structure of HU-210 to the experimentally determined crystal structure of another bound classical cannabinoid (Ligand: AM841, PDB:6KPG[45]) (Figure S11). The classical cannabinoid-bound system was built with identical complex membrane composition, salt concentration, and forcefield with NPS bound system.

For allosteric calculation, other classical cannabinoids (AMG-41, JWH-133, and O-1317) and NPS (AMB-FUBINACA, CUMYL-FUBINACA, 5F-AMP) bound systems were also set up. These ligands are docked into the orthosteric pocket. Best docking poses were selected based on optimizing the distance between the hydroxyl group of classical cannabinoid (linker oxygen for NPS) to S383$^{7.39}$ and the furtherest tail atom distance to W279$^{5.43}$. These systems also have identical complex membrane composition, salt concentration, and forcefield with previously described systems.

System minimization and Equilibration

All ligand-bound systems are minimized and equilibrated before the production run. Ten thousand minimization steps were performed with the conjugate gradient method. Six sequential equilibration steps were carried out to stabilize the systems at 300 K temperature and 1 atm pressure for the production simulations. The systems were heated to 300 K in the NVT ensemble in the initial two stages. Each of these steps was performed for 250 ps. Langevin dynamics was used to control the temperature with additional damping and random force. The damping coefficient for the damping force was set as 1/picosecond. The

Langevin dynamics was turned off for the hydrogen atoms. All the bonded hydrogen atoms were constrained with the SHAKE algorithm with default parameters of NAMD.[80] The integration time step for these two NVT ensemble equilibration was one femtosecond (fs). Harmonic constraints were used for fixing the temperature coupling of the protein residues with constraint scaling term set to 10 in the first NVT ensemble equilibration, followed by 5. Temperature coupling of lipid molecules was also restrained with harmonic force with a force constant equal to 5. CHARMM-GUI Selected Dihedral and Improper bonds were also restrained with an extra bonded term with a force constant of 500 in the first NVT ensemble equilibration, followed by 200. The non-bonded cutoff distance for the van der Waals interactions was set to be 12 Å with a switch distance of 10 Å, at which a switching function is turned on to truncate van der Waals interactions at the cutoff distance. Non-bonded interactions for three consecutively bonded atoms were excluded. The particle mesh Ewald method was implemented for electrostatics calculation with grid size 1 Å.[81]

The next four equilibrations were performed in the NPT ensemble, where pressure was fixed to 1 atm with Langevin piston pressure control. The barostat oscillation period was set to 100 fs with a damping time 50 fs. These four NPT ensemble equilibrations were performed for 250, 500, 500, and 500 ps, respectively. The integration timestep for these equilibration steps was increased to two fs. The constraint for temperature coupling on the protein residues was decreased gradually with the constraint scaling term for the four NPT ensemble simulations set to 2.5, 1.0, 0.5, and 0.1, respectively. Similarly, the restraints on temperature coupling on the lipid molecules were also decreased gradually with force constant for the four NPT ensemble simulations set to 2, 1, 0.2, and 0.0, respectively. Furthermore, the restraints on the CHARMM-GUI Selected dihedral and improper bonds were decreased gradually with force constant for the four steps set to 100, 100, 50, and 0.0, respectively.

Well-tempered metadynamics

Well-tempered metadynamics was implemented for finding unbinding pathways.[47,82] Simulations were performed with the Collective variables module (Colvars) of NAMDv2.14.[83,84] In metadynamics, a history-dependent biasing potential ($V_{meta}(\boldsymbol{S}, t)$) is added to the Hamiltonian of the MD simulation, which discourages the system from revisiting configurations that have already been sampled.[61,85] The $V_{meta}(\boldsymbol{S}, t)$ is a sum of Gaussians deposited along the system trajectory in the CVs space ($\boldsymbol{S} = (S_1(r), S_2(r)..S_d(r))$ as shown in Equation 1, where W, σ, τ are Gaussian height, width, and deposition time step, respectively. With a sufficiently long simulation, the bias potential estimates the underlying free energy along the CVs.[61,82] The Well-tempered metadynamics was introduced to increase the convergence of bias potential by decreasing the Gaussian height with time (Equation 2).[61] In the Equation 2, ω is the bias deposition rate and $\omega\tau$ is equivalent to constant gaussian height for well-tempered metadynamics. The free energy is estimated from the bias potential using Equation 3, where the ΔT is an user defined parameter.

$$V_{meta}(\boldsymbol{S}, t) = \sum_{t'=\tau}^{t} W \prod_{i=1}^{N_{cv}} \exp(\frac{(S(r(t)) - S(r(t')))^2}{2\sigma_i^2}) \tag{1}$$

$$W = \omega\tau \exp(\frac{-V_{meta}(S, t)}{k_B \Delta T}) \tag{2}$$

$$\Delta G(\boldsymbol{S}) = -\frac{T + \Delta T}{\Delta T} V_{meta}(\boldsymbol{S}, t) \tag{3}$$

In this work, we selected two collective variables (CVs) for NPS (MDMB-FUBINACA) and classical cannabinoid (HU-210) unbinding according to Mahinthichaichan et al..[59] The first collective variable is the z-direction distance from the converged toggle switch to the center of mass of all heavy carbon atoms of ligands. The second collective variable is the coordination number (CN) as defined in Equation 4 where d_{ij} represents the distance from

ith atom of the ligand and the alpha carbon of jth residue in the binding pocket. Selected binding pocket residues for CN calculation are shown in the Supplementary table. The ΔT, gaussian height ($\omega\tau$), deposition time step (τ) are selected as 4200K, 0.4 kcal/mol and 100 ps, respectively. The gaussian width for the two CVs are set to be 0.5 and 0.1 ,respectively.

$$CN = \sum_i \sum_j \frac{1 - (d_{ij}/4.5)^8}{1 - (d_{ij}/4.5)^{16}} \tag{4}$$

Umbrella Sampling and Unbiased Sampling

Umbrella sampling was performed along ligand distance from TM5 to capture the unbinding process (Figures S12A and S12B).[52] The unbinding pathway obtained from the metadynamics was clustered into 300 bins by dividing the selected distances from 5 to 35 Å. The center of each bin was used as the center of each window for umbrella sampling. Five independent structures were selected from each cluster to simulate five independent umbrella runs in each umbrella window. If a cluster does not contain any structure, starting structures for that window were selected from the closest clusters. A constant harmonic biased potential of 10 kcal/mol is used for each window. openMM v7.8 MD engine was used to run the umbrella sampling runs.[86] The temperature and pressure of the systems are controlled at 300K and 1 atm by the Langevin thermostats and Monte Carlo barostats. The integration timestep was chosen to be two fs. Movements of the containing Hydrogen atom were constainted using HBonds commands with SHAKE (or SETTLE for water) algorithm. The cutoff distance for Non-bonded interaction other than electrostatic interaction was set to 12 Å, with a switching potential at 10 Å to make the potential to zero smoothly at the cutoff. The particle weld method was used to calculate the long-range electrostatic. Each simulation was run for 20 ns.

Identical starting structures and simulation conditions (Thermostat, barostat, cutoff, electrostatic calculation method, integration timestep, and constraints on Hydrogen bond) were selected for unbiased simulations. openMMv7.8 simulation software was used to run

simulations. Each trajectory was run for 100 ns. All the simulations were performed on the distributive computing facility folding@home.[87]

Markov State Model

Markov state model (MSM) is used to estimate the thermodynamics and kinetics from the unbiased simulation.[53,88] The prevalence of MSM as a post-processing technique for MD simulations was due to its reliance on only local equilibration of MD trajectories to predict the global equilibrium properties.[89,90] Hence, MSM can combine information from distinct short trajectories, which can only attain the local equilibrium.[91–93] This attribute of MSM helps to utilize the power of parallel supercomputers to run parallel short simulations to sample a process with a high energy barrier.

In MSM, the conformational ensemble is discretized into states, assuming that the interstate transition will be Markovian. According to the Markovian property, interstate transitions will only depend on the current state, not on the past state. In order to uphold the principle of Markovian transition, the structures of the discretized states must be conformationally and kinetically similar, allowing for rapid interconversion compared to interstate transitions.[88] These faster intrastate conversions ensure that Markovian states lose their memory of the previous interstate transition after a lag time (τ).

Another necessary criterion for MSM is the reversibility in the interstate transition, which is also known as the detailed balance.[94] The detailed balance assures the local equilibrium between the state population, as shown in Equation 5, where μ_i is the stationary density of ith Markovian states and P_{ij} represents the probability of transition between ith to jth at lag time. These transition probabilities form a transition matrix with $N \times N$ dimension where N is the total number of states. From the eigendecomposition of this transition matrix, biomolecular processes are discretized into orthogonal processes (orthogonal eigenvectors). The timescale of these processes is calculated based on eigenvalues. The highest eigenvalue is one; the corresponding left eigenvector of the transition matrix determines the stationary

density. Eigenvectors with eigenvalues close to one denote the slow processes.

$$\mu_i P_{ij} = \mu_j P_{ji} \tag{5}$$

We can estimate the thermodynamics and kinetics of processes from the transition matrix, whose elements are transition probabilities. These probabilities are usually calculated from the transition counts (C_{ij}) using the maximum likelihood approach, where the likelihood function (Equation 6) is maximized with two constraints: detailed balance and normalization constraints ($\sum_j p_{ij} = 1$). The constrained optimization problem is usually solved iteratively, as shown in Prinz et al..[53]

$$L_{MSM} = \prod_{i,j} (p_{ij})^{c_{ij}} \tag{6}$$

The following steps are taken for the practical implementation of the MSM from the MD data.[4,17,95–97]

1. Each frame obtained from the MD simulation was featurized using features important for capturing the conformational ensemble. In this case, the unbinding process for each ligand was featurized using distances that characterize the ligand distances to the binding pocket and binding pocket conformational change. Specifically, all heavy atom distances from each the Cα carbon atom of all binding pocket residues were calculated (Figure S13). Additionally, all possible combinations of Cα carbon atom distances between all the binding pocket residues were included to capture the binding pocket motion. Feature calculations were performed with the python library MDTraj v1.9.8.[98] The total number features selected for MSM building of MDMB-FUBINACA and HU-210 are 297 and 288, respectively.

2. Dimensionality reduction was performed using time-independent component analysis (TICA).[99,100] We found the orthogonal projections (time-independent components) with TICA, which are linear combinations of the slowest features. In tIC space, two

spatially close points are kinetically close. The lagtime selected for tiC building was 5ns.

3. Clustering was performed on the tICs using k-means clustering algorithms to discretize the space into Markovian states.

4. Lag time for the MSM was calculated by estimating the shortest time at which the timescale of the slowest processes has converged to a particular value (Figures S14A and S14B).

5. To optimize MSM based on the cluster numbers and tIC components on which clustering is performed, we calculated the VAMP-2 score from the MSM, where VAMP stands for Variational Approach for Markov Processes (Figure S15).[101] For a reversible MSM, this score represents the summation of the square of the k slowest eigenvalues, where k is a hyperparameter. Closer the eigenvalue to 1, the corresponding eigenvector captures a slower process. Therefore, we optimize the MSM by maximizing the VAMP-2 score.

6. To validate the Markovian property of our optimized models, ChapmanKolmogorov test (c-k) test was performed (Figure S16). C-K test states that for a Markov model kth power of $P(\tau)$ needs to be equal transition probability matrix determined at $k\tau$ time ($P(\tau)^k \approx P(k\tau)$). We showed that differences between the elements of transition probability matrix at higher lag times remain relatively small.

Dimensionality reduction, clustering Markov state model building, and VAMP-2 calculations are performed with the pyEMMA v2.5.6 library.[102] The optimized MSM for MDMB-FUBINACA unbinding simulations were built with 700 clusters, 7 tiCs and 35 ns of lag time. For HU-210, optimized MSM were build with 800 clusters, 6 tiCs and 35 ns of lag time.

Transition-based reweighting analysis method

Markov State Models have been extensively used to investigate the protein-ligand binding process.[17,62,103–107] However, these studies were mainly performed for ligands with high off-rates which could be sampled using the unbiased trajectories. For ligand with low off rates, the use of reversible transition matrix would yield incorrect estimates of unbinding kinetics. Therefore, we use The Transition-based reweighting analysis (TRAM)[48,108] method to accurately estimate the unbinding kinetics of new psychoactive substances. TRAM is a thermodynamics and kinetics estimator method, which, unlike MSM, can combine unbiased and biased simulation data to estimate thermodynamics and kinetics. TRAM utilizes the advantages of the local equilibrium approximation of MSM and the benefits of biased simulations to enforce local equilibrium in interstate transitions where it is difficult to attain.

As the simulations are obtained from multiple ensembles (biased and unbiased), it is paramount to classify the MD frames (or the conformations) based on which ensemble it belongs to. Each ensemble represents simulations that are performed with identical Hamiltonian energy functions. Therefore, unbiased simulations are considered as one ensemble, whereas, in umbrella sampling, each biasing window is considered a single ensemble.

Like MSM, in TRAM, the conformational space is also discretized into non-overlapping states. The interstate transitions should follow the following relationship shown in the Equation 7, where f_i^k is the local free energy of the ith state and kth ensemble. The term $e^{-f_i^k}$ is proportional to the stationary density ($\mu(x)$) of state S_i in ensemble k . The $\mu(x)$ of each conformation (x) of S_i is weighted with negative exponential of bias energy ($b^k(x)$) on x in ensemble k ($e^{-b^k(x)}$) (Equation 8).

$$e^{-f_i^k} p_{ij}^k = e^{-f_j^k} p_{ji}^k \tag{7}$$

$$e^{-f_i^k} = \int_{S_i} e^{-b^k(x)} \mu(x) \tag{8}$$

To obtain kinetics and thermodynamics information from TRAM, we have to derive interstate transitions (p_{ij}^k) and the stationary density of the entire ensemble ($\mu(x)$), where both the terms follow normalization constraint (Equation 9 and 10). Therefore, there are $m^2K + X$ unknown variables. Therefore, to solve these unknown variables, the maximum likelihood approach has been considered, where the likelihood function is defined as Equation 11, which is the combination of the likelihood function of MSM and local equilibrium. This maximum likelihood problem was subjected to the constraints of Equation 7, 9 and 10.

$$\sum_j p_{ij}^k = 1 \tag{9}$$

$$\sum_{x \in X} \mu(x) = 1 \tag{10}$$

$$L_{TRAM} = \Pi_{k=1}^{K} (\underbrace{\Pi_{i,j} (p_{ij}^k)^{c_{ij}^k}}_{L_{MSM}}) (\underbrace{\Pi_{i=1}^{m} \Pi_{x \in X_i^k} e^{f_i^k - b^k(x)} \mu(x))}_{L_{LEQ}} \tag{11}$$

Wu et al. showed that the solution of these maximum-likelihood problem can be turned into system of non-linear algebraic equations (Equation 12, 13 and 14), where c_{ij}^k is count of the interstate transitions between state S_i and S_j in ensemble k. This system of equations are solved iteratively to estimate v_i^k and f_i^k, which provides the prediction of p_{ji}^k and $\mu(x)$ (Equation 15 and 16).

$$\sum_j \frac{c_{ij}^k + c_{ji}^k}{exp[f_j^k - f_i^k] v_j^k + v_i^k} = 1 \tag{12}$$

$$\sum_{x \in X_i} \frac{exp(f_i^k - b^k(x))}{\sum_l R_i^l exp[f_i^l - b^l(x)]} = 1 \tag{13}$$

$$R_i^k = \sum_j \frac{(c_{ij}^k + c_{ji}^k) v_j^k}{v_j^k + exp[f_i^k - f_j^k] v_i^k} + N_i^k - \sum_j c_{ji}^k \tag{14}$$

$$p_{ij}^k = \frac{c_{ij}^k + c_{ji}^k}{exp[[f_j^k - f_i^k]v_j^k + v_i^k} \tag{15}$$

$$\mu(x) = \frac{1}{\sum_k R_{i(x)}^k exp[f_{i(x)}^k - b^k(x)]} \tag{16}$$

We used the python package pyEMMA v2.5.6 for the practical implementation of TRAM.[102] For calculating transition counts in ensemble k (c_{ij}^k), the lagtime of 15 ns was chosen. In the implementation, we need to preprocess each trajectory into three arrays.

1. One of the arrays represents the spatial discretization of each trajectory frame, where each frame belongs to a particular state. Therefore, each element can take values from 0 to m-1 (m is the cluster number). Before the discretization of the space, time-independent component analysis was performed on the biased and unbiased data separately. The number of tIC components for each system was selected based on the number of the tIC components of optimized MSM. Each frame from the unbiased simulation is represented by the unbiased tICs, concatenated with its feature projections on the biased tICs. Similarly, each frame from the biased simulation is represented by its feature projections on the unbiased tic, concatenated with biased tic projection. Therefore, NPS unbinding simulations have 14 tICs, whereas classical cannabinoid unbinding simulations have 12 tICs. The number of clusters is also obtained from the optimized MSMs.

2. Another array represents the corresponding ensemble to which each trajectory frame belongs. There are 300 windows for the umbrella sampling. Therefore, there are 301 ensembles, as the unbiased simulations represent a separate ensemble.

3. Third array represents the corresponding bias potential ($b_k(x)$) a particular frame feels if it were to be in a particular ensemble. For umbrella sampling, the biased potential is represented as Equation 17, where c_k is selected to be 10 kcal/mol and y_k is the center

of each umbrella window.

$$b_k(x) = \frac{c_k}{2kT}(x - y_k)^2 \tag{17}$$

Transition Path Theory

Transition path theory (TPT) analysis is applied to calculate the transition pathway and timescale between different macrostates, representing different configurational spaces in the unbinding process.[109,110] In this work, we define macrostates as a collection of Markovian states present in the area of interest in the unbinding free energy landscape. An essential concept of transition path theory is the committer probability (q_i^+), which is defined as the probability of any Markovian state reaching the final metastable state before it returns to the initial state. Therefore, the Markovian states present in metastable state B has a committer probability of 1. It has been shown that committer probability follows the following system of linear equation as shown in Equation 18, where P_{ik} is the transition probability between state S_i and S_j as discussed in the previous section.

$$-q_i^+ + \sum_{k \notin B} p_{ik} q_i^+ = \sum_{k \in B} \sum_{k \in B} p_{ik} q_k^+ \tag{18}$$

In this work, the quantity of interest from TPT is the timescale (or rate) between the metastable state transitions as shown in Equation 19, where π_i is the stationary probability of state S_i. TPT calculations were performed by PyEMMA v2.5.6.[102]

$$k_{AB} = \frac{\sum_{k \in A} \sum_{k \notin A} \pi_i p_{ik} q_k^+}{\tau \sum_{i=1}^{m} (\pi_i (1 - q_i^+)} \tag{19}$$

K-L divergence analysis

KullbackLeibler divergence (K-L divergence) analysis was performed to show the structural differences in protein conformations in different macrostates[4,111] . In this study, this tech-

nique was used in this work to calculate the difference in the pairwise inverse distance distribution in macrostates. Although K-L divergence is an asymmetric measurement, for this study, we used a symmetric version of the K-L divergence by taking the average of the K-L divergence between two residues. Per residue contribution of K-L divergence is calculated by taking the sum of all the pairwise distances containing that residue. This analysis was performed by in-house Python code.

Trajectory Analysis

Python package GetContacts is used to perform the contact calculation.[112] Linear interaction energy analysis was performed to calculate the interaction energy between ligand and receptor using AMBERTools CPPTraj v18.01.[113,114] Trajectory visualization and figure generation are performed with VMD v1.9.3.[115]

Deep Learning Network for Allosteric Prediction

Neural relational inference (NRI) network was implemented to predict allosteric dependence between the residues in the different parts of the receptors.[54,116] This network is a Variational autoencoder (VAE) comprising encoding and decoding parts.[117] The encoder ($q_\phi(\boldsymbol{z}|\boldsymbol{x})$) takes the input C$\alpha$ coordinates of protein conformations at time t ($\boldsymbol{x_t}$) and tries to learn the interactions between two residues (z_{ij}) in the protein as a latent space. The decoder ($p_\theta(\boldsymbol{x}|\boldsymbol{z})$) network try to regenerate the protein conformation at time $t + \tau$ ($\boldsymbol{x_{t+\tau}}$). Similar to other VAE, the learning process maximizes the evidence lower bound (ELBO) as shown in Equation 20, where $p_\theta(\boldsymbol{z})$ represents the prior distribution for $\boldsymbol{z}$. Here, the prior distribution is selected as default presented in the original paper, where it is represented as a categorical distribution with $K = 4$ ($P_1 = 0.91$, $P_2 = 0.03$, $P_3 = 0.03$, $P_4 = 0.03$).

As shown in Equation, the ELBO consists of two terms. In the first term, further mathematical derivations can show that the first term can be represented as the the reconstruction error (Equation 21), where σ^2 is variance of the distribution, a user defined parameter.

$$L(\phi, \theta) = \mathbb{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\log p_\theta(\boldsymbol{x}|\boldsymbol{z})] - KL[q_\phi(\boldsymbol{x}|\boldsymbol{z})||p_\theta(\boldsymbol{z})] \tag{20}$$

$$\mathbb{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\log p_\theta(\boldsymbol{x}|\boldsymbol{z})] = \sum_j \sum_{t=2}^{T} \frac{\parallel x_j^t - \mu_j^t \parallel^2}{2\sigma^2} + const \tag{21}$$

The second term is also called regularization term which is the the K-L divergence between estimated posterior ($q_\phi(\boldsymbol{z}|\boldsymbol{x})$) and prior distribution ($p_\theta(\boldsymbol{z})$) (Equation 22). As the prior distribution is a categorical distribution, the K-L divergence becomes entropy of the posterior distribution. We obtained the code for the NRI network from the GitHub implementation and kept most of hyperparameters as default for our training, except for decreasing the hidden layer size to 64.[118] From each unbinding simulations, 10 unbiased trajectories were selected where the ligand remain in the bound pose. Each trajectory has a length of 100 ns. Both cases, the τ was selected to be 5 ns. The allosteric weights (posterior probability) were obtained from the validation data (2 trajectories), where training was performed with remaining 8 trajectories (Figure S18). This procedure was repeated three times, where training and validation data were selected randomly.

$$KL[q_\phi(\boldsymbol{z}|\boldsymbol{x})||p_\theta(\boldsymbol{z})] = \sum_{i \neq j} H(q_\phi(z_{ij}|x)) + const \tag{22}$$

To train the NRI network for other NPS and classical cannabinoid bound systems, ten 100 ns unbiased independent trajectories was run per ligand . Training and validation for all these systems were performed identically as described for the previous systems.

Standard Binding Free Energy Calculations

To calculate the standard binding free energy from simulation, we adopted a procedure described in Buch et al..[62] In this procedure, a volumetric correction term is added to the pmf to calculate the final binding free energy (Equation 23). The volumetric correction term

is used to predict the free energy at the standard condition (1M) as shown in Equation 24, where V_o corresponds to the volume of a molecule should occupy at the standard condition (1661 $\mathring{A}^3$) and V_u is the volume of the unbounded state in the simulation box. The expression for the pmf contribution of the free energy is shown is in Equation 25, where the denominator of the equation can be represented as $\int_u exp(-\beta W(r))dr = exp(-\beta \Delta W)V_u$. Therefore, the final derivation of ΔG is shown in Equation 26

$$\Delta G = \Delta G_{pmf} + \Delta G_V \tag{23}$$

$$\Delta G_V = -k_B T \log(\frac{V_u}{V_o}) \tag{24}$$

$$\Delta G_{pmf} = -k_B T \log \frac{\int_b exp(-\beta W(r))dr}{\int_u exp(-\beta W(r))dr} \tag{25}$$

$$\Delta G = -k_B T \log \frac{\int_b exp(-\beta W(r))dr}{V_o} - \Delta W \tag{26}$$

In this work, to estimate the free energy (ΔG) x, y, z component of the ligand center of mass is calculated compared to the center of mass of the alpha carbons of binding pocket residues. The three dimensional space was descretized into 25 × 25 × 50 bins and each bin is weighted using TRAM calculated probability density. Depth of the pmf (ΔW) was calculated by averaging the pmf of the 100 bins with highest pmf in the bulk. To evaluate the weighted binding volume ($\int_b exp(-\beta W(r))dr$), we selected the bins with pmf less than 1 kcal/mol.

Acknowledgements

D.S. acknowledges support from NIGMS MIRA award R35GM-142745 and NSF Early CAREER Award (MCB-1845606). S.D. and D.S. thank folding@home donors for providing computational resources for the study.

Conflict of interest

The authors declare no conflict of interest.

Code and data availability

Unbinding simulation trajectories and topology files that has been used for the analysis can be obtained from `https://uofi.box.com/s/gcpafgj5l184c3woqk9m92dk468r58mg`. Python scripts and necessary files to generate the figures is provided in the github repository. `https://github.com/ShuklaGroup/Dutta_Shukla_Cannabinoid_2023a.git`

References

(1) Hua, T.; Vemuri, K.; Pu, M.; Qu, L.; Han, G. W.; Wu, Y.; Zhao, S.; Shui, W.; Li, S.; Korde, A.; Laprairie, R. B.; Stahl, E. L.; Ho, J.-H.; Zvonok, N.; Zhou, H.; Kufareva, I.; Wu, B.; Zhao, Q.; Hanson, M. A.; Bohn, L. M.; Makriyannis, A.; Stevens, R. C.; Liu, Z.-J. Crystal Structure of the Human Cannabinoid Receptor CB1. *Cell* **2016**, *167*, 750–762.e14.

(2) Mackie, K. Cannabinoid Receptors: Where They are and What They do. *Journal of Neuroendocrinology* **2008**, *20*, 10–14.

(3) Zou, S.; Kumar, U. Cannabinoid Receptors and the Endocannabinoid System: Signal-

ing and Function in the Central Nervous System. *International Journal of Molecular Sciences* **2018**, *19*, 833.

(4) Dutta, S.; Shukla, D. Distinct activation mechanisms regulate subtype selectivity of Cannabinoid receptors. *Communications Biology* **2023**, *6*.

(5) Rosenbaum, D. M.; Rasmussen, S. G. F.; Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **2009**, *459*, 356–363.

(6) Latorraca, N. R.; Venkatakrishnan, A. J.; Dror, R. O. GPCR Dynamics: Structures in Motion. *Chemical Reviews* **2016**, *117*, 139–155.

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

(8) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nature Reviews Drug Discovery* **2017**, *16*, 829–842.

(9) Yang, D.; Zhou, Q.; Labroska, V.; Qin, S.; Darbalaei, S.; Wu, Y.; Yuliantie, E.; Xie, L.; Tao, H.; Cheng, J.; Liu, Q.; Zhao, S.; Shui, W.; Jiang, Y.; Wang, M.-W. G protein-coupled receptors: structure- and function-based drug discovery. *Signal Transduction and Targeted Therapy* **2021**, *6*.

(10) Smith, T. H.; Sim-Selley, L. J.; Selley, D. E. Cannabinoid CB1 receptor-interacting proteins: novel targets for central nervous system drug discovery? *British Journal of Pharmacology* **2010**, *160*, 454–466.

(11) An, D.; Peigneur, S.; Hendrickx, L. A.; Tytgat, J. Targeting Cannabinoid Receptors: Current Status and Prospects of Natural Products. *International Journal of Molecular Sciences* **2020**, *21*, 5064.

(12) Pertwee, R. G. Cannabinoid pharmacology: the first 66 years. *British journal of pharmacology* **2006**, *147*, S163–S171.

(13) Pertwee, R. G.; Ross, R. Cannabinoid receptors and their ligands. *Prostaglandins, Leukotrienes and Essential Fatty Acids (PLEFA)* **2002**, *66*, 101–121.

(14) Pertwee, R. G.; Howlett, A. C.; Abood, M. E.; Alexander, S. P. H.; Marzo, V. D.; Elphick, M. R.; Greasley, P. J.; Hansen, H. S.; Kunos, G.; Mackie, K.; Mechoulam, R.; Ross, R. A. International Union of Basic and Clinical Pharmacology. LXXIX. Cannabinoid Receptors and Their Ligands: Beyond CB1 and CB2. *Pharmacological Reviews* **2010**, *62*, 588–631.

(15) Razdan, R. K. *The Cannabinoid Receptors*; Humana Press, 2009; pp 3–19.

(16) Maurer, H. H., Brandt, S. D., Eds. *New Psychoactive Substances*; Springer International Publishing, 2018.

(17) Dutta, S.; Selvam, B.; Das, A.; Shukla, D. Mechanistic origin of partial agonism of tetrahydrocannabinol for cannabinoid receptors. *Journal of Biological Chemistry* **2022**, *298*, 101764.

(18) Moore, C. F.; Weerts, E. M. Cannabinoid tetrad effects of oral Δ9-tetrahydrocannabinol (THC) and cannabidiol (CBD) in male and female rats: sex, dose-effects and time course evaluations. *Psychopharmacology* **2021**, *239*, 1397–1408.

(19) Wang, X.-F.; Galaj, E.; Bi, G.-H.; Zhang, C.; He, Y.; Zhan, J.; Bauman, M. H.; Gardner, E. L.; Xi, Z.-X. Different receptor mechanisms underlying phytocannabinoid- versus synthetic cannabinoid-induced tetrad effects: Opposite roles of CB1/CB2 versus GPR55 receptors. *British Journal of Pharmacology* **2020**, *177*, 1865–1880.

(20) Tummino, T. A.; Iliopoulos-Tsoutsouvas, C.; Braz, J. M.; O'Brien, E. S.; Stein, R. M.; Craik, V.; Tran, N. K.; Ganapathy, S.; Shiimura, Y.; Tong, F.; Ho, T. C.; Rad-

chenko, D. S.; Moroz, Y. S.; Liu, F.; Rosado, S. R.; Bhardwaj, K.; Benitez, J.; Liu, Y.; Kandasamy, H.; Normand, C.; Semache, M.; Sabbagh, L.; Glenn, I.; Irwin, J. J.; Kumar, K. K.; Makriyannis, A.; Basbaum, A. I.; Shoichet, B. K. Structure-based discovery of cannabinoid-1 receptor agonists with reduced side effects. **2023**,

(21) Diversion Control Division, Drug Enforcement Administration, U.S. Department of Justice, October, 2019. `https://www.deadiversion.usdoj.gov/drug_chem_info/spice/` (Accessed 2023-19-08).

(22) Farinha-Ferreira, M.; Rei, N.; Fonseca-Gomes, J.; Miranda-Lourenço, C.; Serrão, P.; Vaz, S. H.; Gomes, J. I.; Martins, V.; de Alves Pereira, B.; Sebastião, A. M. Unexpected short- and long-term effects of chronic adolescent HU-210 exposure on emotional behavior. *Neuropharmacology* **2022**, *214*, 109155.

(23) Wiley, J. L.; Marusich, J. A.; Thomas, B. F. *Neuropharmacology of New Psychoactive Substances (NPS)*; Springer International Publishing, 2016; pp 231–248.

(24) Schoeder, C. T.; Hess, C.; Madea, B.; Meiler, J.; Müller, C. E. Pharmacological evaluation of new constituents of "Spice": synthetic cannabinoids based on indole, indazole, benzimidazole and carbazole scaffolds. *Forensic Toxicology* **2018**, *36*, 385–403.

(25) Walsh, K. B.; Andersen, H. K. Molecular Pharmacology of Synthetic Cannabinoids: Delineating CB1 Receptor-Mediated Cell Signaling. *International Journal of Molecular Sciences* **2020**, *21*, 6115.

(26) Akram, H.; Mokrysz, C.; Curran, H. V. What are the psychological effects of using synthetic cannabinoids? A systematic review. *Journal of Psychopharmacology* **2019**, *33*, 271–283.

(27) Worob, A.; Wenthur, C. DARK Classics in Chemical Neuroscience: Synthetic Cannabinoids (Spice/K2). *ACS Chemical Neuroscience* **2019**, *11*, 3881–3892.

(28) Peacock, A.; Bruno, R.; Gisev, N.; Degenhardt, L.; Hall, W.; Sedefov, R.; White, J.; Thomas, K. V.; Farrell, M.; Griffiths, P. New psychoactive substances: challenges for drug surveillance, control, and public health responses. *The Lancet* **2019**, *394*, 1668–1684.

(29) Madras, B. K. *Neuropharmacology of New Psychoactive Substances (NPS)*; Springer International Publishing, 2016; pp 1–18.

(30) Shafi, A.; Berry, A. J.; Sumnall, H.; Wood, D. M.; Tracy, D. K. New psychoactive substances: a review and updates. *Therapeutic Advances in Psychopharmacology* **2020**, *10*, 204512532096719.

(31) Alam, R. M.; Keating, J. J. Adding more "spice" to the pot: A review of the chemistry and pharmacology of newly emerging heterocyclic synthetic cannabinoid receptor agonists. *Drug Testing and Analysis* **2020**, *12*, 297–315.

(32) Potts, A. J.; Cano, C.; Thomas, S. H. L.; Hill, S. L. Synthetic cannabinoid receptor agonists: classification and nomenclature. *Clinical Toxicology* **2019**, *58*, 82–98.

(33) Hill, S. L.; Dunn, M.; Cano, C.; Harnor, S. J.; Hardcastle, I. R.; Grundlingh, J.; Dargan, P. I.; Wood, D. M.; Tucker, S.; Bartram, T.; Thomas, S. H. L. Human Toxicity Caused by Indole and Indazole Carboxylate Synthetic Cannabinoid Receptor Agonists: From Horizon Scanning to Notification. *Clinical Chemistry* **2018**, *64*, 346–354.

(34) Banister, S. D.; Moir, M.; Stuart, J.; Kevin, R. C.; Wood, K. E.; Longworth, M.; Wilkinson, S. M.; Beinat, C.; Buchanan, A. S.; Glass, M.; Connor, M.; McGregor, I. S.; Kassiou, M. Pharmacology of Indole and Indazole Synthetic Cannabinoid Designer Drugs AB-FUBINACA, ADB-FUBINACA, AB-PINACA, ADB-PINACA, 5F-AB-PINACA, 5F-ADB-PINACA, ADBICA, and 5F-ADBICA. *ACS Chemical Neuroscience* **2015**, *6*, 1546–1559.

(35) Ametovski, A.; Macdonald, C.; Manning, J. J.; Haneef, S. A. S.; Santiago, M.; Martin, L.; Sparkes, E.; Reckers, A.; Gerona, R. R.; Connor, M.; Glass, M.; Banister, S. D. Exploring Stereochemical and Conformational Requirements at Cannabinoid Receptors for Synthetic Cannabinoids Related to SDB-006, 5F-SDB-006, CUMYL-PICA, and 5F-CUMYL-PICA. *ACS Chemical Neuroscience* **2020**, *11*, 3672–3682.

(36) Cannaert, A.; Sparkes, E.; Pike, E.; Luo, J. L.; Fang, A.; Kevin, R. C.; Ellison, R.; Gerona, R.; Banister, S. D.; Stove, C. P. Synthesis and in Vitro Cannabinoid Receptor 1 Activity of Recently Detected Synthetic Cannabinoids 4F-MDMB-BICA, 5F-MPP-PICA, MMB-4en-PICA, CUMYL-CBMICA, ADB-BINACA, APP-BINACA, 4F-MDMB-BINACA, MDMB-4en-PINACA, A-CHMINACA, 5F-AB-P7AICA, 5F-MDMB-P7AICA, and 5F-AP7AICA. *ACS Chemical Neuroscience* **2020**, *11*, 4434–4446.

(37) Banister, S. D.; Stuart, J.; Kevin, R. C.; Edington, A.; Longworth, M.; Wilkinson, S. M.; Beinat, C.; Buchanan, A. S.; Hibbs, D. E.; Glass, M.; Connor, M.; McGregor, I. S.; Kassiou, M. Effects of Bioisosteric Fluorine in Synthetic Cannabinoid Designer Drugs JWH-018, AM-2201, UR-144, XLR-11, PB-22, 5F-PB-22, APICA, and STS-135. *ACS Chemical Neuroscience* **2015**, *6*, 1445–1458.

(38) Tai, S.; Fantegrossi, W. E. Synthetic Cannabinoids: Pharmacology, Behavioral Effects, and Abuse Potential. *Current Addiction Reports* **2014**, *1*, 129–136.

(39) Finlay, D. B.; Manning, J. J.; Ibsen, M. S.; Macdonald, C. E.; Patel, M.; Javitch, J. A.; Banister, S. D.; Glass, M. Do Toxic Synthetic Cannabinoid Receptor Agonists Have Signature in Vitro Activity Profiles? A Case Study of AMB-FUBINACA. *ACS Chemical Neuroscience* **2019**, *10*, 4350–4360.

(40) Gatch, M. B.; Forster, M. J. Cannabinoid-like effects of five novel carboxamide synthetic cannabinoids. *NeuroToxicology* **2019**, *70*, 72–79.

(41) Adams, A. J.; Banister, S. D.; Irizarry, L.; Trecki, J.; Schwartz, M.; Gerona, R. "Zombie" Outbreak Caused by the Synthetic Cannabinoid AMB-FUBINACA in New York. *New England Journal of Medicine* **2017**, *376*, 235–242.

(42) Grafinger, K. E.; Cannaert, A.; Ametovski, A.; Sparkes, E.; Cairns, E.; Banister, S. D.; Auwärter, V.; Stove, C. P. Systematic evaluation of a panel of 30 synthetic cannabinoid receptor agonists structurally related to MMB-4en-PICA, MDMB-4en-PINACA, ADB-4en-PINACA, and MMB-4CN-BUTINACA using a combination of binding and different CB1 receptor activation assays—Part II: Structure activity relationship assessment via a β-arrestin recruitment assay. *Drug Testing and Analysis* **2021**, *13*, 1402–1411.

(43) Leo, L. M.; Al-Zoubi, R.; Hurst, D. P.; Stephan, A. P.; Zhao, P.; Tilley, D. G.; Miess, E.; Schulz, S.; Abood, M. E.; Reggio, P. H. The NPXXY Motif Regulates β-Arrestin Recruitment by the CB1 Cannabinoid Receptor. *Cannabis and Cannabinoid Research* **2022**,

(44) Kumar, K. K.; Shalev-Benami, M.; Robertson, M. J.; Hu, H.; Banister, S. D.; Hollingsworth, S. A.; Latorraca, N. R.; Kato, H. E.; Hilger, D.; Maeda, S.; Weis, W. I.; Farrens, D. L.; Dror, R. O.; Malhotra, S. V.; Kobilka, B. K.; Skiniotis, G. Structure of a Signaling Cannabinoid Receptor 1-G Protein Complex. *Cell* **2019**, *176*, 448–458.e12.

(45) Hua, T.; Li, X.; Wu, L.; Iliopoulos-Tsoutsouvas, C.; Wang, Y.; Wu, M.; Shen, L.; Brust, C. A.; Nikas, S. P.; Song, F.; Song, X.; Yuan, S.; Sun, Q.; Wu, Y.; Jiang, S.; Grim, T. W.; Benchama, O.; Stahl, E. L.; Zvonok, N.; Zhao, S.; Bohn, L. M.; Makriyannis, A.; Liu, Z.-J. Activation and Signaling Mechanism Revealed by Cannabinoid Receptor-Gi Complex Structures. *Cell* **2020**, *180*, 655–665.e18.

(46) Lee, Y.; Warne, T.; Nehmé, R.; Pandey, S.; Dwivedi-Agnihotri, H.; Chaturvedi, M.; Edwards, P. C.; García-Nafría, J.; Leslie, A. G. W.; Shukla, A. K.; Tate, C. G. Molec-

ular basis of β-arrestin coupling to formoterol-bound β1-adrenoceptor. *Nature* **2020**, *583*, 862–866.

(47) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters* **2008**, *100*.

(48) Wu, H.; Paul, F.; Wehmeyer, C.; Noé, F. Multiensemble Markov models of molecular thermodynamics and kinetics. *Proceedings of the National Academy of Sciences* **2016**, *113*.

(49) Meral, D.; Provasi, D.; Prada-Gracia, D.; Möller, J.; Marino, K.; Lohse, M. J.; Filizola, M. Molecular details of dimerization kinetics reveal negligible populations of transient μ-opioid receptor homodimers at physiological concentrations. *Scientific Reports* **2018**, *8*.

(50) Hu, X.; Wang, Y.; Hunkele, A.; Provasi, D.; Pasternak, G. W.; Filizola, M. Kinetic and thermodynamic insights into sodium ion translocation through the μ-opioid receptor from molecular dynamics and machine learning analysis. *PLOS Computational Biology* **2019**, *15*, e1006689.

(51) Spiriti, J.; Noé, F.; Wong, C. F. Simulation of ligand dissociation kinetics from the protein kinase PYK2. *Journal of Computational Chemistry* **2022**, *43*, 1911–1922.

(52) Kästner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 932–942.

(53) Prinz, J.-H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schütte, C.; Noé, F. Markov models of molecular kinetics: Generation and validation. *The Journal of Chemical Physics* **2011**, *134*, 174105.

(54) Zhu, J.; Wang, J.; Han, W.; Xu, D. Neural relational inference to learn long-range

allosteric interactions in proteins from molecular dynamics simulations. *Nature Communications* **2022**, *13*.

(55) Gamage, T. F.; Farquhar, C. E.; Lefever, T. W.; Marusich, J. A.; Kevin, R. C.; McGregor, I. S.; Wiley, J. L.; Thomas, B. F. Molecular and Behavioral Pharmacological Characterization of Abused Synthetic Cannabinoids MMB- and MDMB-FUBINACA, MN-18, NNEI, CUMYL-PICA, and 5-Fluoro-CUMYL-PICA. *Journal of Pharmacology and Experimental Therapeutics* **2018**, *365*, 437–446.

(56) Stern, E.; Lambert, D. Medicinal Chemistry Endeavors around the Phytocannabinoids. *Chemistry & Biodiversity* **2007**, *4*, 1707–1728.

(57) Ibrahim, P.; Clark, T. Metadynamics simulations of ligand binding to GPCRs. *Current Opinion in Structural Biology* **2019**, *55*, 129–137.

(58) Saleh, N.; Ibrahim, P.; Saladino, G.; Gervasio, F. L.; Clark, T. An Efficient Metadynamics-Based Protocol To Model the Binding Affinity and the Transition State Ensemble of G-Protein-Coupled Receptor Ligands. *Journal of Chemical Information and Modeling* **2017**, *57*, 1210–1217.

(59) Mahinthichaichan, P.; Vo, Q. N.; Ellis, C. R.; Shen, J. Kinetics and Mechanism of Fentanyl Dissociation from the μ-Opioid Receptor. *JACS Au* **2021**, *1*, 2208–2215.

(60) Saleh, N.; Hucke, O.; Kramer, G.; Schmidt, E.; Montel, F.; Lipinski, R.; Ferger, B.; Clark, T.; Hildebrand, P. W.; Tautermann, C. S. Multiple Binding Sites Contribute to the Mechanism of Mixed Agonistic and Positive Allosteric Modulators of the Cannabinoid CB1 Receptor. *Angewandte Chemie* **2018**, *130*, 2610–2615.

(61) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *WIREs Computational Molecular Science* **2011**, *1*, 826–843.

(62) Buch, I.; Giorgino, T.; Fabritiis, G. D. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proceedings of the National Academy of Sciences* **2011**, *108*, 10184–10189.

(63) Wang, J.; Do, H. N.; Koirala, K.; Miao, Y. Predicting Biomolecular Binding Kinetics: A Review. *Journal of Chemical Theory and Computation* **2023**, *19*, 2135–2148.

(64) Hua, T.; Vemuri, K.; Nikas, S. P.; Laprairie, R. B.; Wu, Y.; Qu, L.; Pu, M.; Korde, A.; Jiang, S.; Ho, J.-H.; Han, G. W.; Ding, K.; Li, X.; Liu, H.; Hanson, M. A.; Zhao, S.; Bohn, L. M.; Makriyannis, A.; Stevens, R. C.; Liu, Z.-J. Crystal structures of agonist-bound human cannabinoid receptor CB1. *Nature* **2017**, *547*, 468–471.

(65) Kapur, A.; Hurst, D. P.; Fleischer, D.; Whitnell, R.; Thakur, G. A.; Makriyannis, A.; Reggio, P. H.; Abood, M. E. Mutation Studies of Ser7.39 and Ser2.60 in the Human CB1 Cannabinoid Receptor: Evidence for a Serine-Induced Bend in CB1 Transmembrane Helix 7. *Molecular Pharmacology* **2007**, *71*, 1512–1524.

(66) Sitkoff, D. F.; Lee, N.; Ellsworth, B. A.; Huang, Q.; Kang, L.; Baska, R.; Huang, Y.; Sun, C.; Pendri, A.; Malley, M. F.; Scaringe, R. P.; Gougoutas, J. Z.; Reggio, P. H.; Ewing, W. R.; Pelleymounter, M. A.; Carlson, K. E. Cannabinoid CB1 receptor ligand binding and function examined through mutagenesis studies of F200 and S383. *European Journal of Pharmacology* **2011**, *651*, 9–17.

(67) Stein, A.; Kortemme, T. Improvements to Robotics-Inspired Conformational Sampling in Rosetta. *PLoS ONE* **2013**, *8*, e63090.

(68) Alford, R. F.; Leaver-Fay, A.; Jeliazkov, J. R.; O'Meara, M. J.; DiMaio, F. P.; Park, H.; Shapovalov, M. V.; Renfrew, P. D.; Mulligan, V. K.; Kappel, K.; Labonte, J. W.; Pacella, M. S.; Bonneau, R.; Bradley, P.; Dunbrack, R. L.; Das, R.; Baker, D.; Kuhlman, B.; Kortemme, T.; Gray, J. J. The Rosetta All-Atom Energy Function for

Macromolecular Modeling and Design. *Journal of Chemical Theory and Computation* **2017**, *13*, 3031–3048.

(69) Mandell, D. J.; Coutsias, E. A.; Kortemme, T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nature Methods* **2009**, *6*, 551–552.

(70) Pándy-Szekeres, G.; Munk, C.; Tsonkov, T. M.; Mordalski, S.; Harpsøe, K.; Hauser, A. S.; Bojarski, A. J.; Gloriam, D. E. GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic Acids Research* **2017**, *46*, D440–D446.

(71) Fay, J. F.; Farrens, D. L. The Membrane Proximal Region of the Cannabinoid Receptor CB1 N-Terminus Can Allosterically Modulate Ligand Affinity. *Biochemistry* **2013**, *52*, 8286–8294.

(72) Andersson, H.; DAntona, A. M.; Kendall, D. A.; Heijne, G. V.; Chin, C.-N. Membrane Assembly of the Cannabinoid Receptor 1: Impact of a Long N-Terminal Tail. *Molecular Pharmacology* **2003**, *64*, 570–577.

(73) Richter, F.; Leaver-Fay, A.; Khare, S. D.; Bjelic, S.; Baker, D. De Novo Enzyme Design Using Rosetta3. *PLoS ONE* **2011**, *6*, e19230.

(74) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry* **2009**, NA–NA.

(75) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **2008**, *29*, 1859–1865.

(76) Ingólfsson, H. I.; Carpenter, T. S.; Bhatia, H.; Bremer, P.-T.; Marrink, S. J.; Light-

stone, F. C. Computational Lipidomics of the Neuronal Plasma Membrane. *Biophysical Journal* **2017**, *113*, 2271–2280.

(77) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A* **2001**, *105*, 9954–9960.

(78) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods* **2016**, *14*, 71–73.

(79) Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **2009**, NA–NA.

(80) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23*, 327–341.

(81) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.

(82) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences* **2002**, *99*, 12562–12566.

(83) Fiorin, G.; Klein, M. L.; Hénin, J. Using collective variables to drive molecular dynamics simulations. *Molecular Physics* **2013**, *111*, 3345–3362.

(84) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26*, 1781–1802.

(85) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annual Review of Physical Chemistry* **2016**, *67*, 159–184.

(86) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Computational Biology* **2017**, *13*, e1005659.

(87) Beberg, A. L.; Ensign, D. L.; Jayachandran, G.; Khaliq, S.; Pande, V. S. Folding@home: Lessons from eight years of volunteer distributed computing. 2009 IEEE International Symposium on Parallel & Distributed Processing. 2009.

(88) Noé, F.; Fischer, S. Transition networks for modeling the kinetics of conformational change in macromolecules. *Current Opinion in Structural Biology* **2008**, *18*, 154–162.

(89) Husic, B. E.; Pande, V. S. Markov State Models: From an Art to a Science. *Journal of the American Chemical Society* **2018**, *140*, 2386–2396.

(90) Noé, F.; Rosta, E. Markov Models of Molecular Kinetics. *The Journal of Chemical Physics* **2019**, *151*.

(91) Bowman, G. R., Pande, V. S., Noé, F., Eds. *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*; Springer Netherlands, 2014.

(92) Wang, W.; Cao, S.; Zhu, L.; Huang, X. Constructing Markov State Models to elucidate the functional conformational changes of complex biomolecules. *WIREs Computational Molecular Science* **2017**, *8*.

(93) Shukla, D.; Hernández, C. X.; Weber, J. K.; Pande, V. S. Markov State Models Provide Insights into Dynamic Modulation of Protein Function. *Accounts of Chemical Research* **2015**, *48*, 414–422.

(94) Noé, F. Probability distributions of molecular observables computed from Markov models. *The Journal of Chemical Physics* **2008**, *128*, 244103.

(95) Dutta, S.; Selvam, B.; Shukla, D. Distinct Binding Mechanisms for Allosteric Sodium Ion in Cannabinoid Receptors. *ACS Chemical Neuroscience* **2022**, *13*, 379–389.

(96) Bansal, P. D.; Dutta, S.; Shukla, D. Activation mechanism of the human Smoothened receptor. *Biophysical Journal* **2023**, *122*, 1400–1413.

(97) Mi, X.; Desormeaux, E. K.; Le, T. T.; van der Donk, W. A.; Shukla, D. Sequence controlled secondary structure is important for the site-selectivity of lanthipeptide cyclization. *Chemical Science* **2023**, *14*, 6904–6914.

(98) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal* **2015**, *109*, 1528 – 1532.

(99) Pérez-Hernández, G.; Paul, F.; Giorgino, T.; Fabritiis, G. D.; Noé, F. Identification of slow molecular order parameters for Markov model construction. *The Journal of Chemical Physics* **2013**, *139*.

(100) Schwantes, C. R.; Pande, V. S. Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. *Journal of Chemical Theory and Computation* **2013**, *9*, 2000–2009.

(101) Wu, H.; Noé, F. Variational Approach for Learning Markov Processes from Time Series Data. *Journal of Nonlinear Science* **2019**, *30*, 23–66.

(102) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J.-H.; Noé, F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *Journal of Chemical Theory and Computation* **2015**, *11*, 5525–5542.

(103) Lawrenz, M.; Shukla, D.; Pande, V. S. Cloud computing approaches for prediction of ligand binding poses and pathways. *Scientific Reports* **2015**, *5*.

(104) Aldukhi, F.; Deb, A.; Zhao, C.; Moffett, A. S.; Shukla, D. Molecular Mechanism of Brassinosteroid Perception by the Plant Growth Receptor BRI1. *The Journal of Physical Chemistry B* **2019**, *124*, 355–365.

(105) Shukla, S.; Zhao, C.; Shukla, D. Dewetting Controls Plant Hormone Perception and Initiation of Drought Resistance Signaling. *Structure* **2019**, *27*, 692–702.e3.

(106) Zhao, C.; Shukla, D. Molecular basis of the activation and dissociation of dimeric PYL2 receptor in abscisic acid signaling. *Physical Chemistry Chemical Physics* **2022**, *24*, 724–734.

(107) Chen, J.; White, A.; Nelson, D. C.; Shukla, D. Role of substrate recognition in modulating strigolactone receptor selectivity in witchweed. *Journal of Biological Chemistry* **2021**, *297*, 101092.

(108) Galama, M. M.; Wu, H.; Krämer, A.; Sadeghi, M.; Noé, F. Stochastic Approximation to MBAR and TRAM: Batchwise Free Energy Estimation. *Journal of Chemical Theory and Computation* **2023**, *19*, 758–766.

(109) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proceedings of the National Academy of Sciences* **2009**, *106*, 19011–19016.

(110) Metzner, P.; Schütte, C.; Vanden-Eijnden, E. Transition Path Theory for Markov Jump Processes. *Multiscale Modeling & Simulation* **2009**, *7*, 1192–1219.

(111) Fleetwood, O.; Carlsson, J.; Delemotte, L. Identification of ligand-specific G protein-coupled receptor states and prediction of downstream efficacy via data-driven modeling. *eLife* **2021**, *10*.

(112) *GetContacts: Interactive analysis for atomic interaction in protein structures*, Stanford University. `https://getcontacts.github.io/` (Accessed 2023-19-08).

(113) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013**, *9*, 3084–3095.

(114) de Amorim, H. L.; Caceres, R.; Netz, P. Linear Interaction Energy (LIE) Method in Lead Discovery and Optimization. *Current Drug Targets* **2008**, *9*, 1100–1105.

(115) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38.

(116) Kipf, T.; Fetaya, E.; Wang, K.-C.; Welling, M.; Zemel, R. Neural Relational Inference for Interacting Systems. 2018; `https://arxiv.org/abs/1802.04687`.

(117) Kingma, D. P.; Welling, M. Auto-Encoding Variational Bayes. 2013; `https://arxiv.org/abs/1312.6114`.

(118) Zhu, J.; Wang, J.; Han, W.; Xu, D. *NRI-MD*, `https://github.com/ethanfetaya/NRI` (Accessed 2023-19-08).