A millisecond coarse-grained simulation approach to decipher allosteric cannabinoid binding at the glycine receptor α_1

Alessio Bartocci,[†,‡] Paulo C.T. Souza,[¶] and Marco Cecchini[*,†,‡]

†*Institut de Chimie de Strasbourg, UMR 7177, CNRS, Universitè de Strasbourg, Strasbourg Cedex 67083, France*

‡*Laboratoire d'Ingenierie des Fonctions Molèculaires UMR7177, 4, rue Blaise Pascal, 67000 Strasbourg*

¶*Molecular Biology and Structural Biochemistry (MMSB, UMR 5086), CNRS & Université de Lyon, 7 Passage du Vercors, 69007, Lyon, France*

E-mail: mcecchini@unistra.fr

Abstract

Glycine receptors (GlyR) are regulated by small-molecule binding at several allosteric sites. Endocannabinoids like tetrahydrocannabinol (THC) and N-arachidonyl-ethanolamide (AEA) potentiate GlyR but their binding site is unknown. Using millisecond coarse-grained MD simulations powered by *Martini 3* we have characterized their binding site(s) at the Zebrafish GlyR-α1 with atomic resolution. Based on hundreds of thousand ligand-binding events, we find that cannabinoids bind at both intrasubunit and intersubunit sites in the transmembrane domain. For THC, intrasubunit binding is in excellent agreement with recent cryo-EM structures, while intersubunit binding recapitulates in full previous mutagenesis experiments. Surprisingly, AEA is found to bind at the same intersubunit site despite the strikingly different chemistry. Statistical analyses of the ligand-receptor interactions highlight potentially relevant residues for GlyR potentiation offering experimentally testable predictions. The results highlight an unanticipated complexity underlying allosteric regulation at synaptic receptors and establish an original simulation protocol for the identification and characterization of allosteric binding sites.

Introduction

Glycine receptors (GlyR) are pentameric ligand-gated ion channels that play a critical role in motor coordination and essential sensory functions such as vision and audition.[1] They are integral transmembrane protein assemblies featuring a large extracellular domain that hosts two or more glycine-binding sites and a compact transmembrane domain (TMD) that forms an axial chloride channel through the postsynaptic membrane. These receptors mediate synaptic inhibition in the spinal cord and the brain stem[2] and have since long been recognized as pharmacological targets for hyperhekplexia, temporal lobe epilepsy, and more recently chronic pain.[3] At the structural level, GlyR is by far the best characterized pentameric ligand-

gated ion channel with > 40 high-resolution structures solved in different conformations (i.e., resting, pre-active, active, and desensitized) and in complex with modulatory ligands such as agonists, partial agonists, antagonists and allosteric modulators.[4–8] A wide panel of small-molecule compounds including psychoactive drugs, general anesthetics, and neurotoxins are known to regulate the GlyR function. Recently, a library of 218 unique chemical entities with documented modulatory activity at homomeric GlyR-α1 and GlyR-α3 along with a structural annotation of their binding site on the receptor has been collected (GRALL).[9] Strikingly, about one third of this collection appears to target the TMD, highlighting the relevance of this region for the allosteric modulation of synaptic receptors.

Cannabis is the most commonly used psychoactive drug worldwide.[10] Endocannabinoids like tetrahydrocannabinol (THC) and cannabidiol (CBD), i.e., the primary psychoactive compounds in cannabis, are lipid-like signaling molecules that were shown to modulate the glycinergic response in addition to targeting the cannabinoid receptors CB1 and CB2.[11] N-arachidonyl-ethanol-amide (AEA), a.k.a. anandamide, was found to potentiate the glycinergic response in oocytes expressing recombinant GlyR-α1.[12] And THC and CBD along with other exogenous cannabinoids, i.e., ajulemic acid, HU-210, and WIN 55212-2, were found to potentiate GlyR non-competitively.[13] Despite its pharmacological relevance, the molecular mechanism of GlyR modulation by cannabinoids has remained elusive, mostly due to lack of structural information at high resolution.

Functional studies by patch-clamp electrophysiology in combination with site-directed mutagenesis have shown that the non-conservative mutation of serine 267 to iso-leucine on the transmembrane helix M2 abolishes GlyR-α1 co-activation by CBD, ajulemic acid and HU-210.[14] And serine substitution at position 296 on the transmembrane helix M3 of purified GlyR-α_3 was shown to abolish potentiation by CBD.[15] Most recently, Kumar *et al.* reported cryo-EM structures of the *Zebrafish* GlyR-α_1 reconstituted in lipid nanodisks in complex with THC.[16] These structures suggest that THC binds to a lipid-exposed intrasubunit pocket at

the interface of the transmembrane helices M3 and M4. However, the presence of THC density both in the resting and activated states of the receptor as well as the existence of unassigned lipid-like densities compatible with THC binding elsewhere indicate that further studies are needed to establish the mechanism of the THC-induced potentiation. Moreover, the allosteric modulation by chemically distinct endocannabinoids such as AEA remains to be explored.

Most drugs currently on the market were designed to target the primary active site of proteins, a.k.a. the orthosteric site. Allosteric ligands that bind to topographically distinct sites offer a competitive advantage over orthosteric compounds as they are more selective, they limit the risk of off-target effects, and can be used synergistically with known drugs to potentiate or attenuate the pharmacological action.[17] Moreover, allosteric sites have been recently exploited to develop therapeutics for proteins that were considered undruggable.[18] The identification of allosteric sites remains challenging and their discovery has so far been mostly serendipitous.[19] Systematic approaches for the identification and structural characterization of allosteric modulatory sites need to be developed, particularly for transmembrane proteins.

The use of coarse-grained (CG) models preserving chemical specificity such as the Martini force field[20] has emerged as a powerful and cost-effective strategy to probe the spatial and temporal evolution of biomolecules,[21,22] particularly for pharmacological applications.[23,24] This modeling strategy provides a smoother potential energy surface that allows for integration time steps up to 30 fs in molecular dynamics (MD) simulations[25]. The latter yields a speed-up of 2-3 orders of magnitude relative to all-atom MD, which opens to the exploration of time and size scales that were previously out of reach. Recently, the newly parameterized version of Martini, a.k.a. *Martini 3*,[26,27] in combination with unbiased MD was shown to reproduce the experimental binding affinity of several small molecules to a protein with high accuracy (< 0.5 kcal/mol) and no *a priori* knowledge of the ligand-binding site.[28] In addi-

tion, refinement of the ligand binding poses by back-mapping to all-atom representations[29] opens to multi-scale simulation approaches to explore protein-ligand binding.

Here, we combine efficient CG simulations powered by *Martini 3* with backmapping to decipher the allosteric site of two endocannabinoids, i.e., THC-Δ^9 and AEA, on the *Zebrafish* GlyR-α1. Our strategy allows mapping of the cannabinoid-receptor interaction in the membrane environment at equilibrium, it provides direct probing of the cannabinoid unbinding kinetics, and allows for the structural characterization of statistically relevant binding modes with atomic resolution. Based on hundred of thousands cannabinoid-binding events, we find that THC and AEA bind at both intrasubunit and intersubunit sites in the transmembrane domain of the receptor. The simulations reveals the existence of a previously uncharacterized intersubunit cannabinoid-binding site, which complements recent structural biology data and recapitulates in full the effect of mutations in functional experiments. The results presented here establish an effective simulation approach for the identification and characterization of allosteric binding sites at synaptic receptors.

Results

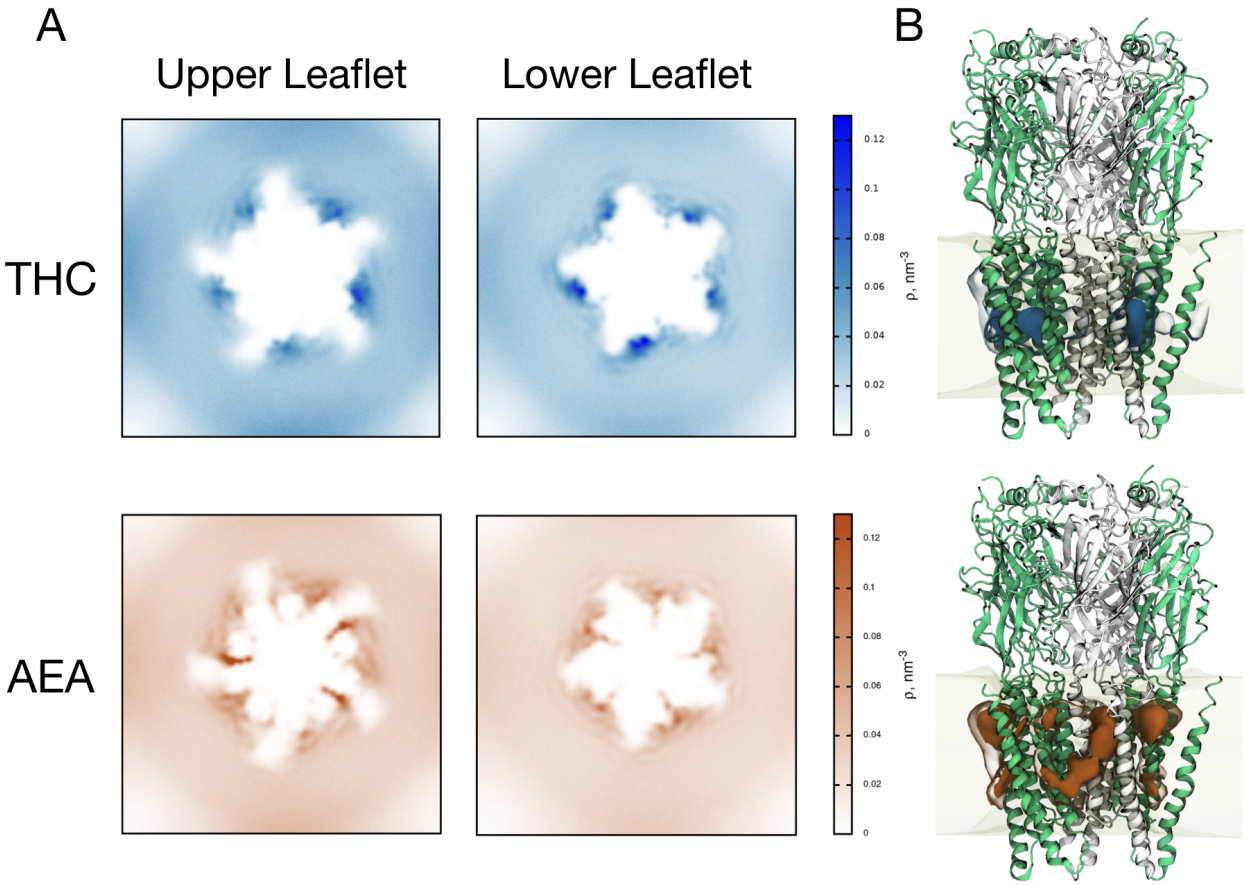

Figure 1: **Equilibrium distribution of cannabinoid ligands around GlyR.** (A) 2D-density maps (in nm^{-3}) for THC (blue) and AEA (orange) in the lipid membrane split in upper and lower leaflets as shown in Figure S1. (B) 3D-density maps. In all cases darker colors correspond to binding "hot-spots". The membrane environment is shown as a semi-transparent region.

GlyR-THC Interaction

To explore the spatial distribution of THC around the transmembrane domain of the GlyR active state, 0.5 milliseconds of unbiased CG/MD simulations with 5% THC in the lipid membrane were collected; with respect to the number of lipids in the membrane. The results in Figure 1A show a highly symmetric distribution with little or no THC density

in the interior of the protein including the ion transmembrane pore (white region). The highly homogeneous color at a distance from the protein-lipid interface (light-blue region) and a few symmetric hot-spots (dark blue) speak for converged sampling. In addition, the bright color of the hot-spots indicates that THC binds at these sites specifically, suggesting the existence of cannabinoid-recognition sites. When the THC density is analysed in 3D, the data reveal a density peak in the mid of the membrane bilayer in proximity of the transmembrane helices M3 and M4 from the same subunit (i.e. intrasubunit); see Figure 1B. Moreover, additional density is found between subunits at the interface of the M3 ($-$) and M1 ($+$) helices. The results altogether indicate that the CG/MD simulation approach provides converged sampling of the GlyR-THC interaction in the lipid membrane and highlight the presence of one or more recognition sites for THC in the active state of the receptor.

In order to characterize the GlyR-THC interaction and isolate specific recognition events, the simulation trajectories were processed by monitoring the number of protein-ligand contacts over time (see *Methods*). Using a dual cut-off scheme to reduce statistical noise, 91266 THC-binding events were collected along with corresponding residence times. Statistics for the bound and unbound states of the ligand over 0.5 ms simulation at 300 K and 5% of THC reveals a free energy of binding in the membrane environment of 0.70 ± 0.01 kcal/mol (see Eq.S6). Remarkably, when the THC-binding affinity was evaluated over increasingly smaller fractions of simulation time, i.e. from 0.5 ms to 1 μs, the uncertainty over the predicted affinity increased by more than one order of magnitude but remained well below 1 kcal/mol (Figure S13). These results demonstrate that the CG/MD approach provides straightforward access to protein-ligand binding affinities (K_d) in the membrane with impressive statistical accuracy (see SI).

In addition, CG/MD simulations open to direct estimates of the ligand-binding kinetics (K_{off}). The distribution of the residence time of THC on GlyR (Figure S12) indicates that 98.6% of the binding events are short lived and correspond to non-specific interactions at

the protein-lipid interface. However, a small but significant fraction (i.e. 1290 events) are long-lived ($>$ 100ns) and represent receptor-ligand recognition events. By clustering binding events using an RMSD cutoff and correcting for the pentameric symmetry (see *Methods*), a collection of 255 unique THC-binding modes were extracted. Their probability distribution in Figure 2A highlights the existence of three dominant binding modes that account altogether for $\sim$40% of the GlyR-THC recognition sampled in simulation. Analysis of the residence times per binding mode reveals monoexponential distributions with a characteristic time of 160 ns for the most probable binding mode (orange), and 78 ns and 66 ns for the second (red) and third (cyan); (Figure 2C). Interestingly, these data indicate that the longest-lived binding mode is also the most populated.

Upon "backmapping" to all-atom representations, the structural models of the three dominant binding modes highlight that THC may bind specifically both intersubunit and intrasubunit. When THC binds intersubunit, the terpenic core of the ligand is sandwiched between two tryptophanes from the helix M1 (W255 and W259), and the short alkyl tail inserts in a hydrophobic pocket at the interface of the M3 ($-$) and M1 ($+$) helices (orange, Figure 2C). When THC binds intrasubunit (red or cyan), the ligand binds at the interface between M3 and M4 and forms a H-bonding interaction with a nearby serine (red or cyan, Figure 2C). Notably, the two intrasubunit binding modes target the same pocket but involve flipping of THC by 180°, which re-positions its H-bonding moiety relative to the residues that line the pocket. Reassuringly, the representations of the three dominant binding modes by THC are consistent with the hot-spots highlighted by the density plots in Figure 1.

In order to pinpoint residues that are potentially relevant for the GlyR potentiation by THC, a statistically relevant ligand-receptor contact analysis was carried out on the simulation frames corresponding to the three dominant THC-binding modes (see *Methods*). The results in Figure S17 predict that five residues in the TMD are critical for binding THC inter-subunit, with W255, F258 and W259 from M1 stabilizing the terpenic core of the ligand

via stacking interactions and L315 from M3(−) and I252 from M1(+) forming an intersubunit hydrophobic pocket where the short alkyl chain slips in. In addition, six residues appear to be relevant for THC-binding intrasubunit, with S312 or S406 forming a H-bond with the phenol group of THC and A409, F410, and F414 from M4 and L315 from M3 forming the hydrophobic pocket that accomodates the terpenic core of the ligand. Strikingly, the recent mutagenesis analysis by Kumar et al[16] revealed that residue substitution at W255, F258, W259, S312 and F410, which are all identified as critical in this analysis (Figure S17), significantly decreases, if not eliminates, GlyR potentiation by THC; for residue numbering consider a +8 shift in the sequence of Kumar (7M6O) relative to the sequence of 6PM6 used here. Moreover, the literature observation that serine substitution at position 296 (i.e. S312 in 6PM6) abolishes GlyR potentiation by CBD[15] is consistent with our predictions. These comparisons show that the statistical analysis of the receptor-THC interactions sampled by CG/MD is fully consistent with available mutagenesis data.

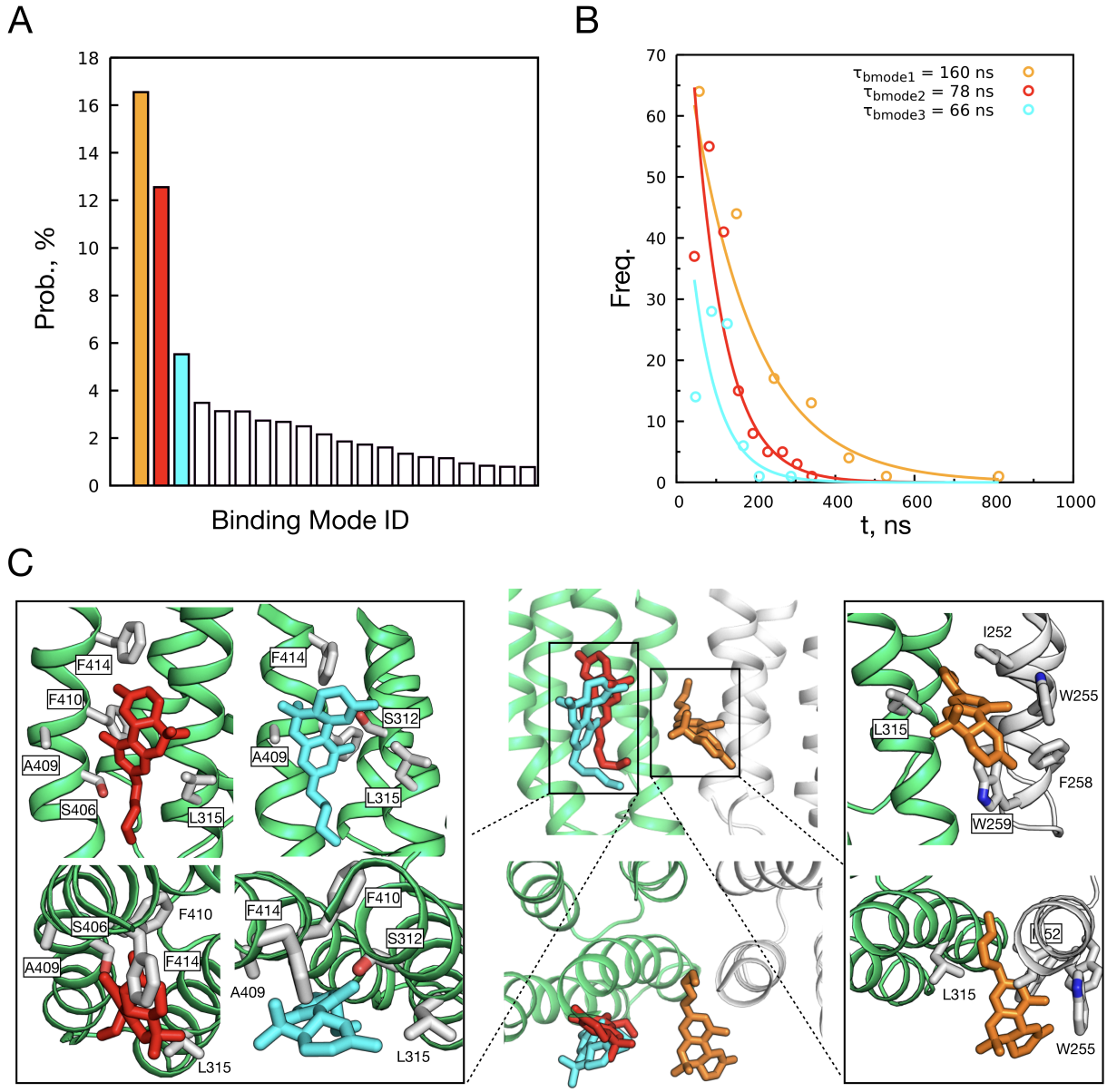

Figure 2: **THC allosteric binding at GlyR**. (A) Specific binding modes are represented with associated probabilities. The dominant binding modes are color-coded. (B) Residence time (τ_r) distributions of the three most important binding events. The characteristic residence time per binding mode is given. The fit was done as described in the *Methods*. (C) Atomistic representation of the dominant binding modes. Protein chains are color-coded in green and white to highlight adjacent subunits. The THC molecule is shown in sticks and colored following the scheme in panel A. For each binding mode, the protein residues involved in the ligand-receptor recognition are highlighted in the insets.

GlyR-AEA Interaction

A similar investigation based on millisecond CG/MD simulations was carried out to characterize the interaction between GlyR and anandamide (AEA). The results show a different interaction pattern relative to THC with AEA penetrating much deeper in the TMD and being mostly localized intrasubunit in the upper leaflet and intersubunit in the lower leaflet (Figure 1). Similar to THC, the probability of finding AEA at the receptor-lipid interface is higher than the bulk at more than one interaction site.

A detailed analysis of the receptor-ligand contacts reveals that 129346 binding events were sampled by 0.5 ms CG/MD at 300 K and 5% (with respect to the number of lipids in the membrane) of AEA, 1.5 % of which have a residence time > 100 ns and represent ligand-receptor specific recognition events (Figure S12). The distribution of the residence time reveals that AEA binding to the GlyR active state is longer-lived than THC, with binding events lasting up to 6 μs, which is six-fold longer than the longest lived THC-binding event (Figure 3). Similar to THC, AEA binding to the GlyR active state is slightly unfavorable with a $\Delta G = 0.51 \pm 0.03$ kcal/mol, albeit being more favorable than THC by 0.2 kcal/mol at same experimental conditions. Statistics over the long-lived binding events highlight the existence of four dominant binding modes over 192 distinct binding modes, which account for more than 60 % of the specific AEA-receptor recognition (Figure 3A). Two of them (green and cyan) have a residence time $\tau_r > 800$ ns, whereas the other two (red and yellow) are shorter-lived with residence times $\tau_r \simeq$ 200-300 ns (Figure 3B). Upon "backmapping", atomistic representations of the dominant binding modes reveal that the two longest binding modes correspond to AEA intrasubunit binding at an upper leaflet site (green and cyan, Figure 3C), whereas the two shorter-lived binding modes correspond to intersubunit binding either at an upper leaflet site (orange) or at a lower leaflet site (red).

To identify residues potentially relevant for GlyR potentiation by AEA, a statistical analysis of the per-residue ligand-receptor contacts was carried out over 0.5 ms of CG/MD

simulation (Figure 4). In short, the analysis reveals that: i) binding modes 1 and 3 are essentially equivalent (Figure S15); ii) when AEA binds intrasubunit, the ligand inserts in a hydrophobic cavity at the interface of the transmembrane helices M1, M3, and M4 and anchors its head via polar contacts with the side chains of S247 from M1 and C306 from M3, while the rest of the molecule makes non-polar contacts with the side chains of M243, Y244, Y251 and W255 from M1, W302, M303 and F309 from M3, and L415, N418 and W422 from M4; iii) when AEA binds at the upper intersubunit site (orange), the ligand inserts the polar head in the cleft formed by M3(+) and M1(−) helices and establishes a H-bonding interaction with the side chain of S283 from M2 similar to ivermectin binding[4,5], while forming non-polar contacts with two isoleucines (I241 and I245) and one tyrosine (Y244) from M1 and one alanine (A304) from M3; and iv) when AEA binds to the lower inter-subunit site, molecular recognition involves the surprising insertion of the polar head in the same hydrophobic pocket occupied by the alkylic tail of THC in its intersubunit binding mode along with extensive interaction with W255 and W259 from M1 that sandwich most of the alkylic tail of the ligand. We note that in the lower intersubunit binding mode (red) the polar head of AEA inserts deeper than THC and establishes polar interactions with water molecules percolated in the interior of the protein from the ion pore via several water channels (see Figure S16). Intriguingly, the CG/MD simulations predict that the same residues implicated in THC binding (i.e. I252, W255, W259, and L315) are critical for AEA binding, despite the strikingly different chemistry of these two cannabinoids. In addition, they highlight the implication of S283 (S267 in human) in AEA recognition, whose role in GlyR potentiation by cannabinoids in functional assays[14] as well as ivermectin binding by cryo-EM studies[4,5] has been previously recognized.

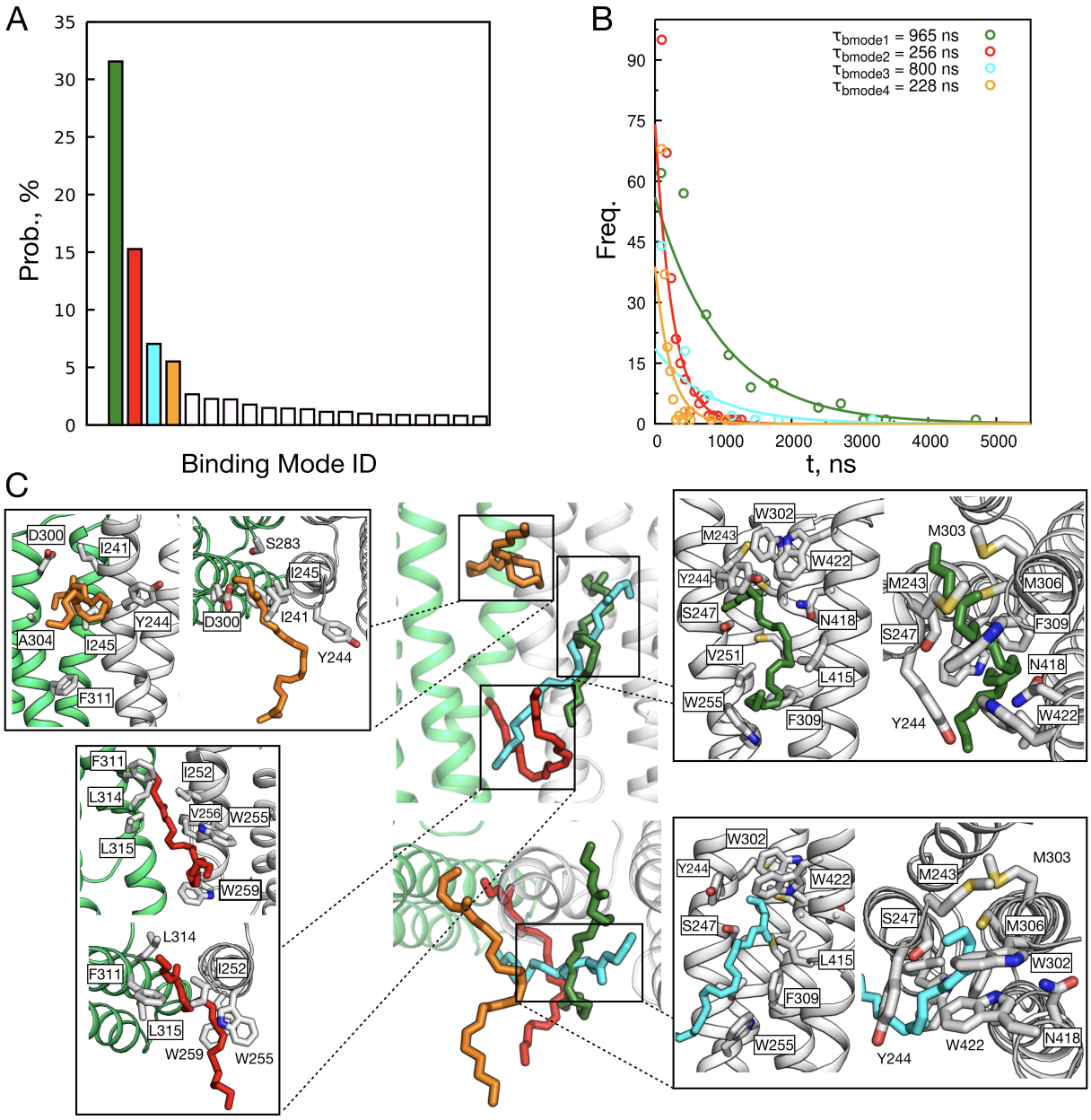

Figure 3: **AEA allosteric binding at GlyR**. (A) Specific binding modes are represented with associated probabilities. The dominant binding modes are color-coded. (B) Residence time (τ_r) distributions of the four most important binding events. The characteristic residence time per binding mode is given. The fit was done as described in the *Methods*. (C) Atomistic representation of GlyR-AEA interaction in the dominant binding modes. Protein chains are color-coded in green and white to highlighting adjacent subunits, while AEA molecule is shown in sticks and color-coded following the scheme in panel A. For each binding mode, the protein residues involved in the ligand-receptor recognition are highlighted in the insets.

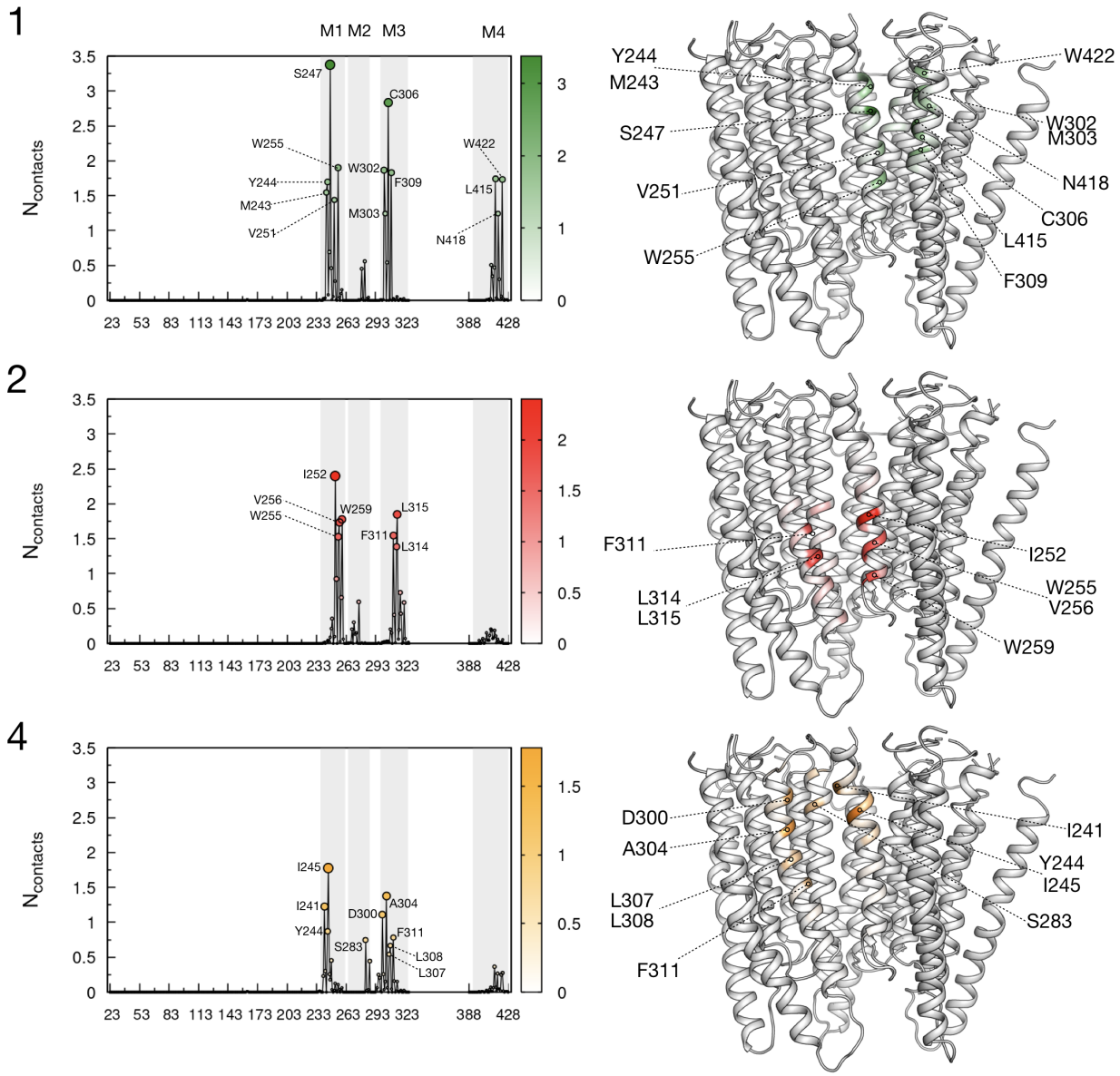

Figure 4: **GlyR-AEA contact analysis**. The analysis was carried out for the three dominant GlyR-AEA binding modes (green, red and orange in Figure 3) as described in the *Methods*. On the left-hand side, the average number of contacts per residue is plotted along the sequence of the protein. Gray boxes indicate protein stretches corresponding to the transmembrane helices M1-M4. On the right-hand side, the residues identified by the contact analysis are indicated on the protein structure. The color code is the same as that in Figure 3.

Discussion

Endogenous cannabinoids such as tetrahydrocannabinol (THC) and anandamide (AEA) regulate the function of brain receptors including the ionotropic glycine receptor (GlyR)[1]. Due to the lack of structural information at high resolution, the molecular mechanism of GlyR potentiation by THC and AEA has remained elusive. Very recently, cryoEM structures of the *Zebrafish* GlyR-α1 reconstituted in lipid nanodiscs illuminated the receptor in complex with THC.[16] These structures highlighted the existence of an intrasubunit cannabinoid-binding site located in the transmembrane protein domain at the interface of the M3 and M4 helices. The intrasubunit nature of this site is surprising and inconsistent with previous structural biology of GlyR and pentameric homologues, which shows that allosteric modulatory sites are typically located between subunits.[9] Moreover, the different chemistry of THC and AEA prevents from a straightforward extrapolation of the structural results for THC, so that the nature of the allosteric modulatory site for fatty-acid cannabinoids remains to be elucidated.

We set out to provide an atomistic representation of the GlyR-cannabinoid interaction using a multi-scale simulation approach powered by the synergistic combination of coarse-grained MD simulations with *Martini 3* and "backmapping" to explore the details of molecular recognition with atomic resolution. Based on hundreds of thousand ligand-binding events, the simulations reveal the existence of multiple cannabinoid-binding sites in the transmembrane domain of GlyR and provide atomistic representations of the dominant binding modes. As a bonus, the simulations provide converged results for the ligand binding-affinity (K_d) and residence time (K_{off}), which are useful to explore the thermodynamics and the kinetics of cannabinoid-receptor binding in the limit of the approximations of the energy model. The results for THC reveal the existence of two allosteric sites for di-terpene cannabinoids that account for $> 30\%$ of the specific ligand-receptor recognition sampled by millisecond CG/MD. One binding site is intrasubunit and lies at the interface between the transmembrane helices

M3 and M4. The other is intersubunit and located between the M3(+) and M1(−) helices. Remarkably, the intrasubunit allosteric site predicted in simulation perfectly matches the THC-binding mode illuminated by the cryoEM structures of Kumar et al[16] (Figure 5A), which validates the modeling approach. In addition, the intersubunit binding mode overlaps with the neurosteroid-binding mode illuminated by the X-ray structure of GlyR-GABA$_A$R chimera in complex with tetrahydrodeoxycorticosterone (PDB ID: 5OSB)[30] (Figure 5B) and is consistent with additional unassigned density present in the cryo-EM maps of Kumar et al[16]; see Figure S18. Intriguingly, the discovery of an intersubunit binding site for THC offers a plausible re-interpretation of recent site-directed mutagenesis studies showing that residue substitution (to alanine or phenyl-alanine) at W259, P266 and F410 almost completely removes GlyR potentiation by THC without changing the sensitivity to glycine.[16] To explain the data, the existence of an allosteric coupling between the intrasubunit THC-binding site and the desensitization gate of GlyR (P266) via a network of aromatic residues involving the M1-M2 linker has been postulated.[16] The existence of an intersubunit THC-binding site at the M1-M3 interface, offers a more natural explanation of the mutagenesis results and provides further support to our simulations.

The results on AEA predict that despite the strikingly different chemistry, fatty-acid cannabinoids and di-terpene cannabinoids bind to the same "lower" intersubunit site. Also and unlike THC, they show that AEA may bind to an "upper" intersubunit pocket that strongly overlaps with the binding site of the positive allosteric modulator ivermectin (PDB ID: 5VDH)[5] (Figure 5C). Given the similar mode of action of THC and AEA on the glycinergic response,[13] the existence of a common lower intersubunit binding site illuminated by CG/MD (i.e. same binding pocket and comparable binding kinetics) offers a plausible mechanism for the GlyR allosteric modulation by endocannabinoids. In addition, our analysis offers precise experimentally testable predictions. In fact, if AEA binding to the lower intersubunit site is involved in GlyR potentiation, the same mutagenesis analysis of Kumar et al[16]

with AEA is predicted to show a significant reduction, if not abolition, of GlyR potentiation in some of those mutants. Vice versa, if AEA binding to the upper intersubunit site is mostly involved in GlyR potentiation, the serine to iso-leucine non-conservative mutation of S283 (S267 in human), which was found to abolish GlyR potentiation by cannabidiol, ajulemic acid and HU-210[14], is predicted to have similar effects with AEA. In case site-directed mutagenesis at both upper and lower intersubunit sites were shown to affect GlyR potentiation, the illumination of two distinct intersubunit sites by our CG/MD analysis would provide a direct explanation for the finding.

More generally, the results in this paper establish an original simulation protocol for the identification and structural characterization of allosteric sites in transmembrane protein domains. Starting with high-resolution structures of the protein, CG/MD simulations powered by Martini have become broadly accessible and are straightforward to run.[31,32] The efficiency of these simulations provides converged statistical sampling, which allows for a thorough exploration of the protein-ligand interactions both structurally and energetically. As shown by the two cannabinoids explored here, this aspect is critical for protein-ligand recognition in lipid membranes, which results from the competition among almost equivalent binding modes at topographically distinct sites (intrasubunit/intersubunit). In light of this, the development of automated workflows combining extended CG/MD trajectories with robust backmapping procedures appears a promising avenue to harness such a molecular complexity and yield atomistic representations of the statistically relevant binding modes. The same result is currently out of reach using standard modeling approaches (e.g. all-atom MD or docking) and possibly structural biology too. In fact, while all-atom MD and docking are expected to fail because of incomplete sampling for the former and inappropriate ranking of binding modes for the latter, structural studies might fail because they typically provide coordinates for a single ligand-binding mode, which might be only marginally relevant for allosteric modulation. On the other hand, CG parameterization of small molecules remains

challenging and is the bottleneck of the CG/MD simulation protocol implemented here. Despite some developments have started to emerge,[33] the establishment of robust and efficient strategies for the automatic parameterization of ligands in *Martini 3* is a challenge for the future.

In conclusion, the results presented in this paper provide fundamental insights on the cannabinoid-GlyR interaction and offer precise experimentally testable predictions. This work highlights an unexpected complexity underlying ligand binding and allosteric regulation at synaptic receptors and presents an original simulation approach to harness it. The striking reproduction of the cryo-EM binding mode of THC in GlyR demonstrates that coarse-grained MD simulations may be predictive and are useful to identify and characterize allosteric transmembrane sites with no *a priori* knowledge of the ligand-binding site or mode.

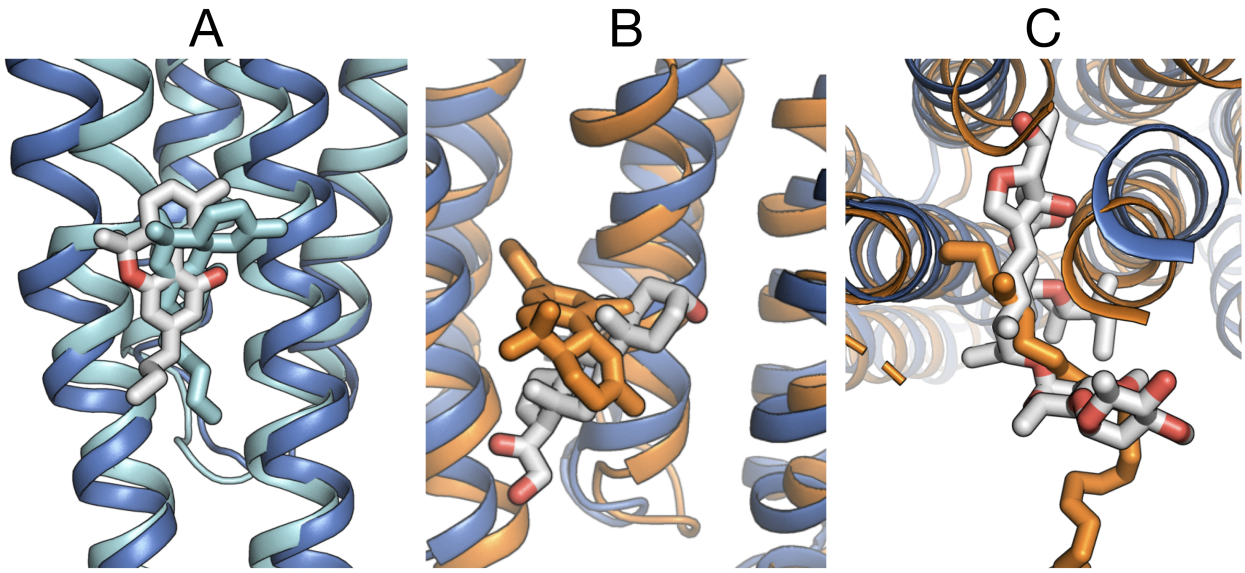

Figure 5: **Comparison of cannabinoid binding modes predicted by CG/MD with experiments**. (A) Superimposition of the cryo-EM structure of GlyR solved in complex with THC (PDB ID: 7M6O) with the THC intrasubunit binding mode predicted in simulation (cyan in Figure 2). The all-atom RMSD between experimental and predicted ligand-binding modes is 2.7 Å. (B) Superimposition of X-ray structure of the GlyR-GABA$_A$R chimera solved in complex with tetrahydrodeoxycorticosterone (PDB ID: 5OSB) with the THC intersubunit binding mode predicted in simulation (orange in Figure 2). (C) Superimposition of the X-ray structure of GlyR solved in complex with ivermectin (PDB ID: 5VDH) with the AEA upper intersubunit binding mode predicted in simulation (orange in Figure 3).

Material and Methods

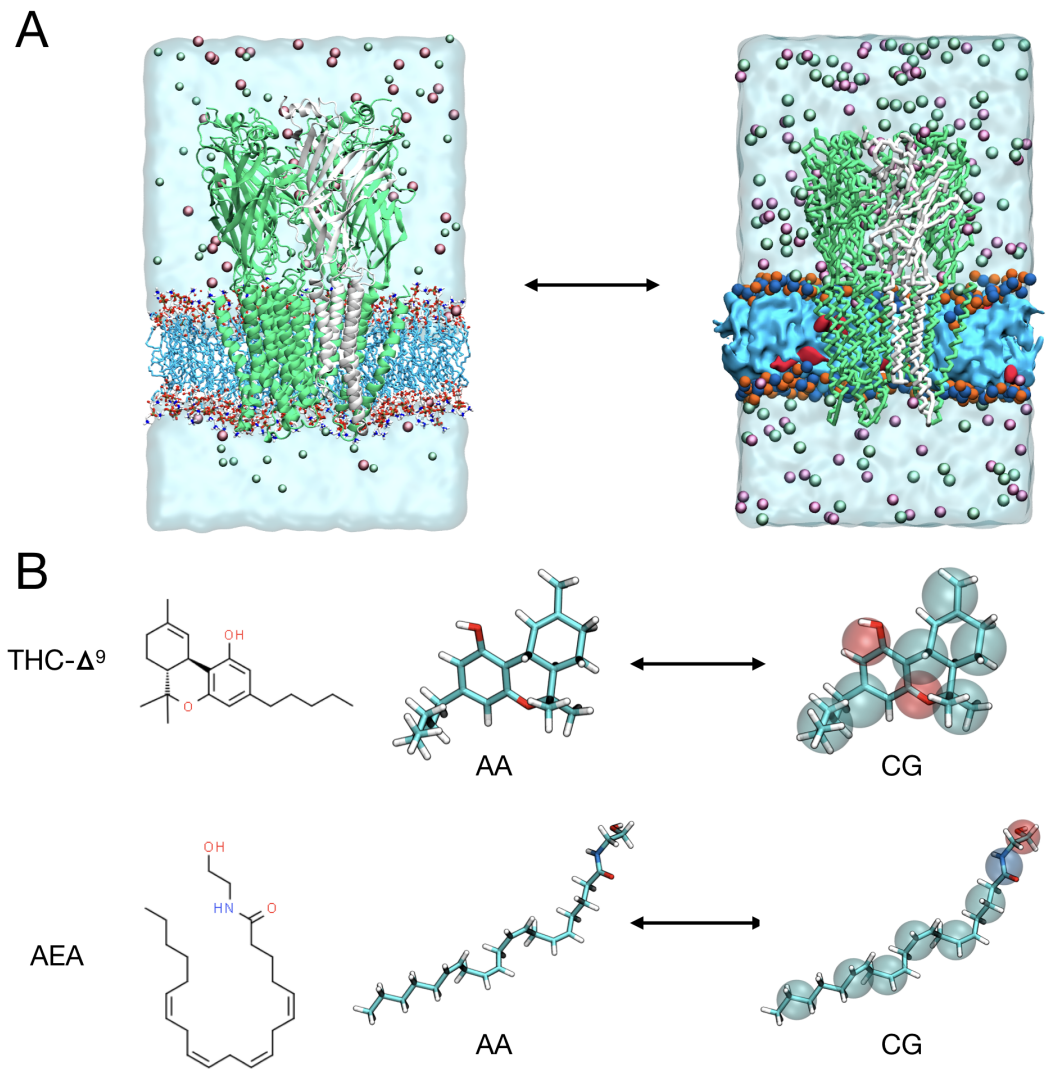

Figure 6: **CG/MD simulation setup**. (A) The GlyR active state (PDB ID: 6PM6) was embedded in a POPC lipid membrane with 5% of cannabinoids. The protein is shown as white and green cartoons. The POPC polar heads are shown as blue and orange spheres with the alkylic tails represented as a cyan surface. The cannabinoid ligands are shown in red. Ions are shown as green (Na^+) and violet (Cl^-) spheres, while water molecules are represented by a light-blue continuum. (B) On the left, the chemical structure of tetrahydrocannabinol (THC) and N-arachidonyl-ethanol-amide (AEA) are shown. All-atom (AA) representations of the two cannabinoids are given in the middle. Corresponding coarse-grained (CG) representations are given on the right with beads represented as semi-transparent colored spheres.

Coarse-grained simulations

Classical coarse-grained molecular dynamics (CG/MD) simulations were performed using the GROMACS package[34] version 2021 with the *Martini 3* force field.[26,28] Initial coordinates of GlyR-α1 from *Zebrafish* in its open, ion-conducting state were taken from the PDB (6PM6).[8] The protein was converted to a CG representation using the martinize 2.0 tool.[35–37] An elastic network with a distance cutoff of 0.8 nm was applied to preserve a symmetric pentameric architecture. The force constant of the elatic network was set to 1500 $kJ\,mol^{-1}\,nm^{-2}$, which yields an RMSD of the backbone beads of 2.4 Å from the initial coordinates without freezing the transmembrane helices; see Table S1. The elastic network was introduced on the receptor upon removing the glycine agonist from the orthosteric site.

The protein was embedded in a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayer (164/173 in upper/lower leaflets) with 5% (16 copies) of THC or AEA using `insane.py`[38] (Figure 6A). Using the same tool, the system was solvated with water, Cl^- counterions, and 0.15 M concentration of NaCl. CG parameters for tetrahydrocannabinol (THC) and N-arachidonyl-ethanol-amide (AEA) were derived as described below.

Ligand parameterization

Initial coordinates for THC and AEA were taken from the GRALL database[39] and converted into CHARMM36-FF[40,41] topologies using the CGENFF python-tool.[42] CG parameters for the ligands were obtained following standard procedures in *Martini 3*[27] as described in SI. The CG mapping is presented in Figure S2. For the bonded interactions, CG parameters were validated by comparing bond distances, angles and dihedrals as well as molecular volume and SASA with all-atom MD results for the same molecules in solution (see SI). The non-bonded parameters were validated by comparing water/octanol partition free energies of the CG models with $\log P_{oct/wat}$ values from the literature (see SI). For the reference simulations in all-atom MD, THC and AEA were simulated with 4000 TIP3P[43] water molecules in the

NPT ensemble at 298K and 1 bar.

CG/MD simulations

For the CG/MD simulations, the molecular systems were energy minimized for 50000 steps using the steepest descent algorithm and then equilibrated for 100 ns MD with position restraints on the protein backbone ($1000\,kJ\,mol^{-1}\,nm^{-2}$ force constant) and for 1 μs without restraints. Production runs of 10 μs were carried out in 50 replicates with different initial velocities for a total sampling of 0.5 ms per molecular system. A 20 fs integration time step t_{step} was used and periodic boundary conditions applied. During the equilibration and production stage, the temperature was maintained constant at 300 K using a modified Berendsen thermostat[44] with a τ_T coupling constant of 1 ps. The pressure was maintained at 1 bar (semiisotropic coupling) using a Berendsen barostat[45] ($\tau_P = 4\,ps$, compressibility of $3 \cdot 10^{-4}\,bar^{-1}$) during the initial equilibration and a Parrinello-Rahman barostat[46] ($\tau_P = 12\,ps$, compressibility of $3 \cdot 10^{-4}\,bar^{-1}$) for the second equilibration and production. Electrostatic interactions were cutoff at 1.1 nm using the reaction-field method[47], while Lennard-Jones interactions were cutoff at 1.1 nm using the potential-shift Verlet method. Bonds between beads were constrained to their equilibrium values using the LINCS algorithm.[48] During the 1 μs MD equilibration and all production runs, an elastic network acting on the protein backbone was introduced to preserve the quaternary structure of the receptor (see above).

Density maps

2D density maps were generated using the `gmx densmap` tool. The upper and lower membrane leaflets were defined as the volume slices (along the axis perpendicular to the membrane) between the center of mass of the membrane and the POPC heads (NC3 and PO4 beads) (see Figure S1). The ligand 3D densities were obtained by computing the occupancy of the ligands in the three-dimensional space using the "`volmap`" plugin of `vmd`[49], with the

protein aligned in the center of the box. The grid points had a distance of 2 Å. All ligand beads were taken into account in their actual size: the radius of regular and small beads is 2.64 Å and 2.30 Å, respectively, while the one of tiny beads 1.91 Å.

Ligand-binding events

Ligand binding events were identified by monitoring the number of protein-ligand contacts and using a dual-cutoff scheme to reduce the statistical noise. For this purpose, a cutoff of 10 contacts (cutoff_{bound}) was used to identify the beginning of a binding event, while a cutoff of 5 contacts ($\text{cutoff}_{unbound}$) to identify ligand unbinding (Figure S11). The total number of frames multiplied by the saving frequency, i.e. the simulation time between two saved frames, yields the ligand residence time per binding event. Binding events lasting longer than 100 ns were annotated as "specific" (Figure S12), extracted and analyzed; they represent approximately 1% of the binding events sampled for THC and AEA (Table S3). Representative structures were then obtained by a cluster analysis with an RMSD cutoff on the ligand coordinates corresponding to the peak of the pair-wise distribution per binding event. The center of the most populated cluster was chosen as representative of the binding event.

Ligand-binding modes

To correct for the pentameric symmetry of the receptor, representative structures of specific ligand-binding events were compared after protein alignment to a common subunit-subunit interface. Upon ranking events by residence time, ligand-binding events were grouped into ligand-binding modes by a leader clustering algorithm using an RMSD cutoff of 0.34 nm and starting from the longest-lived event as reference. The probability per binding mode was then determined by counting the total number of frames associated to it. For THC,

the entire molecule was considered for the RMSD calculation, whereas for AEA, given its higher flexibility, only the head beads were considered. The RMSD cutoff of 0.34 nm was fine-tuned by trials and errors. The characteristic residence time per binding mode was then obtained by fitting the distribution of the residence time of the associated binding events using a single exponential law as $f(t) = A\,e^{-t/\tau_n}$, with τ_n being the residence time of the n-*th* binding mode. Last, atomistic representations of the most populated binding modes were obtained using the "*backward*" tool.[29] For this purpose, GlyR, POPC, THC/AEA and ions were described with the CHARMM36 force field[40,41], while water with the TIP3P[43] model. The backmap files to produce all-atom representations of the ligands were generated from scratch.

Protein-ligand contacts

To pinpoint residues primarily involved in cannabinoid binding, a protein-ligand contact analysis was carried out on the ensemble of snapshots corresponding to the dominant cannabinoid-GlyR binding modes. For this purpose, the `gmx mindist` tool was used to compute the average number of protein-ligand contacts per residue ($N_{contacts}$) using a bead distance cutoff of 0.5 nm. By plotting N_{contacts} along the sequence of the protein per binding mode, the statistically most relevant residues naturally emerge; see Figure 4.

Rendering

`Gnuplot` was used to obtain 2D-maps in Figure 1A, panels A and B in Figure 2 and Figure 3, as well as left panels in Figure 4. `VMD`[49] was used to render Figure 1B and Figure 6. `Pymol`[50] for panel C of Figure 2 and Figure 3, the right panels in Figure 4, and Figure 5.

Author Contributions

MC conceived the study and obtained the funding. AB and MC designed the simulation protocol. AB and PCT parameterized the ligands. AB performed and analyzed the MD simulations. AB and MC wrote the manuscript. AB, PCT and MC analyzed the results and commented on the manuscript. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This work received support from a grant of the "Agence nationale de la Recherche" (ANR–18–CE11–0015) to MC. Support from the French National Center for Scientific Research (CNRS) and funding from a research collaboration agreement with PharmCADD to PCTS are gratefully acknowledged. Computational resources and support at the HPC centers of the University of Strasbourg (Mesocentre) and the University of Reims (ROMEO) are also acknowledged.

Conflict of Interests

The authors declare no conflicts of interest.

References

(1) Lynch, J. W. Molecular structure and function of the glycine receptor chloride channel. *Physiological reviews* **2004**, *84*, 1051–1095.

(2) Dutertre, S.; Becker, C.-M.; Betz, H. Inhibitory glycine receptors: an update. *Journal of Biological Chemistry* **2012**, *287*, 40216–40223.

(3) Lynch, J. W.; Zhang, Y.; Talwar, S.; Estrada-Mondragon, A. Glycine receptor drug discovery. *Advances in Pharmacology* **2017**, *79*, 225–253.

(4) Du, J.; Lü, W.; Wu, S.; Cheng, Y.; Gouaux, E. Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature* **2015**, *526*, 224–229.

(5) Huang, X.; Chen, H.; Shaffer, P. L. Crystal structures of human GlyRα3 bound to ivermectin. *Structure* **2017**, *25*, 945–950.

(6) Huang, X.; Shaffer, P. L.; Ayube, S.; Bregman, H.; Chen, H.; Lehto, S. G.; Luther, J. A.; Matson, D. J.; McDonough, S. I.; Michelsen, K., et al. Crystal structures of human glycine receptor α3 bound to a novel class of analgesic potentiators. *Nature structural & molecular biology* **2017**, *24*, 108–113.

(7) Kumar, A.; Basak, S.; Rao, S.; Gicheru, Y.; Mayer, M. L.; Sansom, M. S.; Chakrapani, S. Mechanisms of activation and desensitization of full-length glycine receptor in lipid nanodiscs. *Nature communications* **2020**, *11*, 3752.

(8) Yu, J.; Zhu, H.; Lape, R.; Greiner, T.; Du, J.; Lü, W.; Sivilotti, L.; Gouaux, E. Mechanism of gating and partial agonist action in the glycine receptor. *Cell* **2021**, *184*, 957–968.

(9) Cerdan, A. H.; Sisquellas, M.; Pereira, G. P.; Barreto Gomes, D. E.; Changeux, J.-P.; Cecchini, M. The Glycine Receptor Allosteric Ligands Library (GRALL). *Bioinformatics* **2020**, *36*, 3379–3384.

(10) Ahmed, S.; Ross, S.; Slade, D.; Radwan, M.; Zulfiqar, F.; Matsumoto, R. Global Burden of Disease Study 2013 Collaborators (2015). Global Burden of Diseases, Injuries, and Risk Factors Study 2013. Lancet. 386; 9995: 743–800. Madras BK (2015). Update of cannabis and its medical use. Report to the WHO Expert Committee on Drug Dependence. *Lancet 386*, 743–800.

(11) Campbell, C. T.; Phillips, M. S.; Manasco, K. Cannabinoids in pediatrics. *The Journal of Pediatric Pharmacology and Therapeutics* **2017**, *22*, 176–185.

(12) Hejazi, N.; Zhou, C.; Oz, M.; Sun, H.; Ye, J. H.; Zhang, L. Δ9-Tetrahydrocannabinol and endogenous cannabinoid anandamide directly potentiate the function of glycine receptors. *Molecular pharmacology* **2006**, *69*, 991–997.

(13) Yévenes, G. E.; Zeilhofer, H. U. Molecular sites for the positive allosteric modulation of glycine receptors by endocannabinoids. *PLoS One* **2011**, *6*, e23886.

(14) Foadi, N.; Leuwer, M.; Demir, R.; Dengler, R.; Buchholz, V.; de la Roche, J.; Karst, M.; Haeseler, G.; Ahrens, J. Lack of positive allosteric modulation of mutated α1S267I glycine receptors by cannabinoids. *Naunyn-Schmiedeberg's archives of pharmacology* **2010**, *381*, 477–482.

(15) Xiong, W.; Cui, T.; Cheng, K.; Yang, F.; Chen, S.-R.; Willenbring, D.; Guan, Y.; Pan, H.-L.; Ren, K.; Xu, Y., et al. Cannabinoids suppress inflammatory and neuropathic pain by targeting α3 glycine receptors. *Journal of Experimental Medicine* **2012**, *209*, 1121–1134.

(16) Kumar, A.; Kindig, K.; Rao, S.; Zaki, A.-M.; Basak, S.; Sansom, M. S.; Biggin, P. C.; Chakrapani, S. Structural basis for cannabinoid-induced potentiation of alpha1-glycine receptors in lipid nanodiscs. *Nature Communications* **2022**, *13*, 1–14.

(17) Chatzigoulas, A.; Cournia, Z. Rational design of allosteric modulators: Challenges and successes. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2021**, *11*, e1529.

(18) Hocker, H. J.; Cho, K.-J.; Chen, C.-Y. K.; Rambahal, N.; Sagineedu, S. R.; Shaari, K.; Stanslas, J.; Hancock, J. F.; Gorfe, A. A. Andrographolide derivatives inhibit guanine

nucleotide exchange and abrogate oncogenic Ras function. *Proceedings of the National Academy of Sciences* **2013**, *110*, 10201–10206.

(19) Lu, S.; Ji, M.; Ni, D.; Zhang, J. Discovery of hidden allosteric sites as novel targets for allosteric drug design. *Drug discovery today* **2018**, *23*, 359–365.

(20) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *The journal of physical chemistry B* **2007**, *111*, 7812–7824.

(21) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The power of coarse graining in biomolecular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 225–248.

(22) Salassi, S.; Simonelli, F.; Bartocci, A.; Rossi, G. A Martini coarse-grained model of the calcein fluorescent dye. *Journal of Physics D: Applied Physics* **2018**, *51*, 384002.

(23) Souza, P. C. T.; Limongelli, V.; Wu, S.; Marrink, S. J.; Monticelli, L. Perspectives on high-throughput ligand/protein docking with Martini MD simulations. *Frontiers in Molecular Biosciences* **2021**, *8*, 199.

(24) Kjølbye, L. R.; Pereira, G. P.; Bartocci, A.; Pannuzzo, M.; Albani, S.; Marchetto, A.; Jiménez-García, B.; Martin, J.; Rossetti, G.; Cecchini, M., et al. Towards design of drugs and delivery systems with the Martini coarse-grained model. *QRB Discovery* **2022**, *3*, e19.

(25) Fábián, B.; Thallmair, S.; Hummer, G. Small ionic radii limit time step in Martini 3 molecular dynamics simulations. *The Journal of Chemical Physics* **2022**, *157*, 034101.

(26) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M.; Wassenaar, T. A., et al. Martini 3:

a general purpose force field for coarse-grained molecular dynamics. *Nature methods* **2021**, *18*, 382–388.

(27) Alessandri, R.; Barnoud, J.; Gertsen, A. S.; Patmanidis, I.; de Vries, A. H.; Souza, P. C. T.; Marrink, S. J. Martini 3 Coarse-Grained Force Field: Small Molecules. *Advanced Theory and Simulations* **2022**, *5*, 2100391.

(28) Souza, P. C. T.; Thallmair, S.; Conflitti, P.; Ramírez-Palacios, C.; Alessandri, R.; Raniolo, S.; Limongelli, V.; Marrink, S. J. Protein–ligand binding with the coarse-grained Martini model. *Nature communications* **2020**, *11*, 1–11.

(29) Wassenaar, T. A.; Pluhackova, K.; BoÌckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models. *Journal of chemical theory and computation* **2014**, *10*, 676–690.

(30) Laverty, D.; Thomas, P.; Field, M.; Andersen, O. J.; Gold, M. G.; Biggin, P. C.; Gielen, M.; Smart, T. G. Crystal structures of a GABAA-receptor chimera reveal new endogenous neurosteroid-binding sites. *Nature structural & molecular biology* **2017**, *24*, 977–985.

(31) Marchetto, A.; Si Chaib, Z.; Rossi, C. A.; Ribeiro, R.; Pantano, S.; Rossetti, G.; Giorgetti, A. CGMD platform: integrated web servers for the preparation, running, and analysis of coarse-grained molecular dynamics simulations. *Molecules* **2020**, *25*, 5934.

(32) Hilpert, C.; Beranger, L.; Souza, P. C. T.; Vainikka, P. A.; Nieto, V.; Marrink, S. J.; Monticelli, L.; Launay, G. Facilitating CG simulations with MAD: The MArtini database server. *Journal of chemical information and modeling* **2023**, *63*, 702–710.

(33) Empereur-Mot, C.; Pesce, L.; Doni, G.; Bochicchio, D.; Capelli, R.; Perego, C.; Pavan, G. M. Swarm-CG: automatic parametrization of bonded terms in MARTINI-based coarse-grained models of simple to complex molecules via fuzzy self-tuning particle swarm optimization. *ACS omega* **2020**, *5*, 32823–32843.

(34) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25.

(35) de Jong, D. H.; Singh, G.; Bennett, W. D.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved parameters for the martini coarse-grained protein force field. *Journal of chemical theory and computation* **2013**, *9*, 687–697.

(36) Kroon, P. C.; Grünewald, F.; Barnoud, J.; van Tilburg, M.; Souza, P. C.; Wassenaar, T. A.; Marrink, S.-J. Martinize2 and Vermouth: Unified Framework for Topology Generation. *arXiv preprint arXiv:2212.01191* **2022**,

(37) *https://github.com/marrink-lab/vermouth-martinize*

(38) Wassenaar, T. A.; Ingólfsson, H. I.; Bockmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. *Journal of chemical theory and computation* **2015**, *11*, 2144–2155.

(39) *https://ifm.chimie.unistra.fr/grall*

(40) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell Jr, A. D.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *The journal of physical chemistry B* **2010**, *114*, 7830–7843.

(41) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell Jr, A. D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. *Journal of chemical theory and computation* **2012**, *8*, 3257–3273.

(42) Vanommeslaeghe, K.; Raman, E. P.; MacKerell Jr, A. D. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *Journal of chemical information and modeling* **2012**, *52*, 3155–3168.

(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics* **1983**, *79*, 926–935.

(44) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of chemical physics* **2007**, *126*, 014101.

(45) Berendsen, H. J.; Postma, J. v.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics* **1984**, *81*, 3684–3690.

(46) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics* **1981**, *52*, 7182–7190.

(47) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *The Journal of chemical physics* **1995**, *102*, 5451–5459.

(48) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry* **1997**, *18*, 1463–1472.

(49) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38.

(50) DeLano, W. L.; Bromberg, S. PyMOL user's guide. *DeLano Scientific LLC* **2004**, *629*.