

Structural basis for ligand modulation of the CCR2 conformational landscape

Bryn C. Taylor^a, Christopher T. Lee^b, and Rommie E. Amaro^{a,b,1}

^aDepartment of Biomedical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093; ^bDepartment of Chemistry and Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, California, 92093

This manuscript was compiled on November 22, 2018

CC Chemokine Receptor 2 (CCR2) is a part of the chemokine receptor family, an important class of therapeutic targets. These class A G-protein coupled receptors (GPCRs) are involved in mammalian signaling pathways and control cell migration toward endogenous CC chemokine ligands, named for the adjacent cysteine motif on their N-terminus. Chemokine receptors and their associated ligands are involved in a wide range of diseases and thus have become important drug targets. CCR2, in particular, promotes the metastasis of cancer cells and is also implicated in autoimmunity driven type-1 diabetes, diabetic nephropathy, multiple sclerosis, asthma, atherosclerosis, neuropathic pain, and rheumatoid arthritis(1–3). Although promising, CCR2 antagonists have been largely unsuccessful to date(4–6). Here, we investigate the effect of an orthosteric and an allosteric antagonist on CCR2 dynamics by coupling long timescale molecular dynamics simulations with Markov-state model theory. We find that the antagonists shift CCR2 into several stable inactive conformations that are distinct from the crystal structure conformation and disrupt a continuous internal water and sodium ion pathway, preventing transitions to an active-like state. Several metastable conformations present a cryptic drug binding pocket near the allosteric site that may be amenable to targeting with small molecules. Without antagonists, the apo dynamics reveal intermediate conformations along the activation pathway that provide insight into the basal dynamics of CCR2, and may also be useful for future drug design.

CC Chemokine Receptor 2 | CCR2 | GPCR | Markov State Models

The signaling axis of CCR2 and its endogenous ligand, CCL2, is a notable therapeutic target due to its association with numerous diseases, including cancer, autoimmunity driven type-1 diabetes, diabetic nephropathy, multiple sclerosis, asthma, atherosclerosis, neuropathic pain, and rheumatoid arthritis(1–3). Despite much effort that has been devoted to clinical and pre-clinical trials, a successful antagonist has yet to be developed(4–10) (<http://www.clinicaltrials.gov>). Prior to a full-length crystal structure of CCR2, several studies used homology modeling and docking to gain insights into the structure and dynamics of the protein and its associated ligands or small molecule drugs(11–13). However, recently CCR2 was crystallized for the first time(14), opening up new opportunities for rational drug design.

As with most GPCRs, chemokine receptors transmit signals across cell membranes by means of extracellular ligand and intracellular G-protein binding. Distinct conformational states of the receptor are necessary for chemokine/ligand binding, G-protein binding, activation, inactivation, and signal transmission (15–17). GPCRs are no longer considered to be simple on/off molecular switches – instead, they assume a wide range of conformational states, including ligand-specific states, intermediate states, and states that allow for basal (apo) signaling without ligands bound (16, 18–24). Ligands

and small molecule drugs may shift the equilibrium of the receptor's conformational states towards particular states. Effective small molecule antagonists that inhibit CCR2 signaling, potentially by shifting the receptor equilibrium toward inactive conformational states, are desired for treatment of diseases that involve the CCR2/CCL2 axis. Key challenges are to characterize the basal dynamics of CCR2 and to understand how current antagonistic small molecule drugs modulate these dynamics. While crystal structures provide valuable snapshots of proteins and protein complexes, they lack the ability to reveal dynamics at the atomic level. Starting with the newly resolved crystal structure of CCR2 (PDB ID: 5T1A) we performed multi-microsecond all-atom explicitly solvated molecular dynamics (MD) simulations of the receptor in a lipid bilayer in unbound (apo) and dual-antagonist-bound (holo) states (Figure 1). The two antagonists were co-crystallized with CCR2: the orthosteric antagonist, BMS-681(14, 25), and the allosteric antagonist, CCR2-RA-[R](14, 26). Each system was simulated in triplicate on the Anton 2 supercomputer(27) for a total of 260 microseconds (SI Table S1).

While long timescale (10s of microseconds) simulations are useful for analyzing sequential conformational changes, simulations are generally unable to directly probe timescales of biological interest (milliseconds - seconds)(28). One way to bridge this timescale gap is to couple MD simulations with Markov state model theory(29–35) (MSM, described in SI Materials and Methods). Integrating MD simulations with MSMs allowed us to extend the reach of simulated timescales, and identify key differences in the conformational ensembles and dominant slow motions of apo and holo CCR2. We find that the antagonists disrupt a continuous internal water and sodium ion pathway preventing transitions to an active-like state, and shift CCR2 into several stable inactive states that are distinct from the crystal structure conformation, three of

Significance Statement

CCR2 is an important immune system protein that is implicated in many immune and inflammatory diseases, including cancer. Numerous clinical and pre-clinical trials are attempting to target CCR2 to find an effective therapy. To support drug discovery efforts, we use atomic simulations and Markov-state models to investigate the conformational landscape of CCR2; information that is unattainable by experimental methods.

B.C.T., C.T.L., and R.E.A. designed research; B.C.T. performed the research; B.C.T. and C.T.L. analyzed the results; and B.C.T., C.T.L., and R.E.A. wrote the paper.

The authors declare the following competing financial interest(s): REA has equity interest in, and is a cofounder and scientific advisor of, Actavalon, Inc.

¹To whom correspondence should be addressed. E-mail: ramaro@ucsd.edu

which present a cryptic druggable pocket. Without antagonists, intermediate conformations with active-state conformational signatures shed light on the apo dynamics of CCR2 and may also be useful for future drug design.

Results and Discussion

Constructing the Markov State Models. To explore the conformational landscape of both apo and holo CCR2 and identify metastable macrostates and the transitions between the macrostates, we built MSMs from MD simulation data. All simulations, barring initial equilibration runs on local resources, were performed on the Anton 2 supercomputer⁽²⁷⁾ at the Pittsburgh Supercomputing Center. Both the apo and holo systems were simulated in triplicate (SI Table S1, Figure S1). The systems were initiated from the coordinates of the experimentally determined crystal structure⁽¹⁴⁾. For the apo system, the co-crystallized orthosteric and allosteric antagonist ligands were removed. In each case the receptor was placed in a POPC lipid bilayer with 0.15M NaCl and solvated with TIP3P molecules.

We used time-structure independent component analysis (tICA)^(36–38) starting with all pairwise inter-residue distances to perform dimensionality reduction and identify the features and collective variables (time-structure based independent components (tICs)) that best represent the dominant slow motions in the apo and holo simulations. To reduce the number of pairwise distances to a manageable number, we employed an iterative tICA feature selection approach. First, tICA was run on a curated starting set of hundreds of distances between transmembrane helices. Features with low TIC correlation were subsequently removed from the set and tICA rerun on the resultant new basis. In this iterative fashion, we selected 22 representative features (SI for further details). These resultant features are the 22 sets of distances between residues in the orthosteric and allosteric ligand-binding pockets (Figure 1A).

One MSM was built on the apo data and a second MSM was built on the holo data. In each case, we projected the data into tICA space and clustered the trajectory frames using K-means clustering implemented in PyEMMA⁽³⁹⁾. After clustering, the two MSMs were built and tested. The selected models are Markovian, as indicated by implied timescale plots (Figures S2A, S3A). The Chapman-Kolmogorov test⁽⁴⁰⁾ was used to test the consistency between the MSMs and the MD simulations (Figures S2B, S3B). The apo MSM has a lag time of 14.4 nanoseconds and 665 clusters; and the holo MSM has a lag time of 48 nanoseconds and 790 clusters. The MSMs were coarse-grained using hidden Markov models (HMMs) to identify metastable states and their representative structures, as well as transitions between those states. The apo MSM is coarse-grained into six macrostates; the holo MSM into seven macrostates (Figure 1B,C). A representative structure for each macrostate was selected by taking the centroid of the most populated microstate (Figures S4-6).

For one MD replicate of the holo system, we observed the orthosteric drug dissociate from the protein. Although only one dissociation event is sampled, analyzing the conformations before and after ligand dissociation yields a first glimpse of the allosteric effect of the remaining antagonist on the protein dynamics, and provides a starting point for future rounds of adaptive sampling to obtain robust dissociation statistics (not pursued here).

Comparison of the CCR2 conformational ensemble with other class A GPCRs. We compare representative states from the apo and holo conformational ensemble with other class A GPCRs to establish similarities within the class. We find that the states of apo CCR2 have conformational signatures found in the active or intermediately-active states of GPCRs, suggesting that these states are on a pathway toward activation. Holo CCR2 diverges from the crystal structure to form distinct states that expose putative drug binding pockets and reveal the effect of antagonists on receptor dynamics.

We evaluate the metastable states by comparing helical conformational signatures and conserved groups of structurally neighboring amino acids called 'microswitches'. These include NPXXY (Tyr 305^{7,53}), DRY (Arg 138^{3,50}), Tyr 222^{5,58}, sets of residues in the orthosteric and allosteric binding sites, and the chemokine and G-protein binding pockets to the inactive crystal structure of CCR2 that we used in this study (PDB ID: 5T1A), to an intermediately-active crystal structure of a class A GPCR, A_{2A}AR (PDB ID: 2YDO⁽⁴¹⁾; 25% sequence identity to CCR2), the active crystal structure of a class A GPCR, US28 (PDB ID: 4XT3⁽⁴²⁾; 30% sequence identity to CCR2), and three other chemokine receptors: CCR5 (PDB ID: 4MBS⁽⁴³⁾), CCR9 (PDB ID: 5LWE⁽⁴⁴⁾), and CXCR4 (PDB ID: 4RWS⁽⁴⁵⁾). Signatures of an active GPCR state include: 1) the inward shift of the intracellular part of helix VII toward the helical bundle, 2) the outward shift of the intracellular part of helix VI in concert with helix V, 3) the upward shift and lateral movement of helix III, and 4) the rearrangements of conserved microswitches⁽¹⁸⁾. According to these metrics, the starting crystal structure of CCR2 is in an inactive conformation⁽¹⁴⁾, the crystal structure of A_{2A}AR is in an intermediately-active conformation, and the crystal structure of US28 is in the active conformation.

1) Apo macrostates show an active-like inward shift of the intracellular part of helix VII toward the helical bundle. All of the apo macrostates exhibit an active state hallmark (Figure 2A): the intracellular end of helix VII tilts slightly inward toward the center of the helical bundle. More prominently, the extracellular end of helix VII tilts outward, resembling the active conformation of US28. The holo macrostates show the opposite: the intracellular end of helix VII tilts slightly outward and the extracellular end of helix VII tilts inward, remaining in the crystal structure conformation.

2) Holo macrostates, not apo, show an active-like outward shift of the intracellular part of helix VI in concert with helix V. Helix V and VI in the apo macrostates are not in an active conformation. Instead, it is the holo macrostates that have the intracellular end of helix V and VI tilting outward to resemble the active conformation (Figure 2C,D), suggesting that neither apo nor holo macrostates are in an exclusively inactive or active conformational state, despite starting from a particularly inactive crystal structure. Due to this outward motion of helix VI, holo macrostates K and L exhibit a more open G-protein binding site compared to holo macrostate G which is more closed. The RMSF of the allosteric ligand is larger in macrostates K and L, indicating that the inactive (inward) conformation of helix VI may play a critical role in stabilization of the allosteric ligand (Figure S16).

3) Apo macrostates show an active-like upward shift and lateral movement of helix III. Apo macrostates also resemble the ac-

tive conformation by the slight upward shift of helix III; unlike holo macrostates, which remain in a position similar to the inactive crystal structure (Figure 2B).

4) The rearrangements of conserved microswitches suggest that apo macrostates are active intermediates, and holo macrostates are inactive intermediates. a) *NP_{xx}Y motif (Tyr 305^{7.53})*. In the inactive conformation of GPCRs, Tyr 305^{7.53} points towards helices I, II, or VIII (in CCR2, it points toward II), and in the active state Tyr 305^{7.53} points toward middle axis of helical bundle(18). Each apo macrostate shows Tyr 305^{7.53} pointing downward into the intracellular (G-protein) binding pocket (Figure 2F). This positioning of Tyr 305^{7.53} matches the intermediately-active conformation of A₂AR, which also points down. It does not match the active conformation in US28 that points up, and is also distinct from the inactive crystal structure of CCR2. In six out of the seven the holo macrostates, Tyr 305^{7.53} is stabilized in the inactive state and matches the inactive CCR2 crystal structure conformation as well as the inactive chemokine receptor crystal structures of CCR5 and CCR9.

The holo macrostate in which Tyr 305^{7.53} is not stabilized in the inactive conformation is accessed after the orthosteric ligand dissociates (State L, Figures 1C, 2F); the allosteric pocket residues rearrange and Tyr 305^{7.53} assumes a downward conformation similar to the apo states and the crystal structure of CXCR4. These concerted events may indicate allosteric cross-talk between the chemokine binding site and the G-protein binding site.

b) *The microswitch residue Trp 256^{6.48}, and the interaction of the DRY motif (Arg 138^{3.50}) with Tyr 222^{5.58}*. Apo and holo macrostates both maintain the same chi angle of the conserved microswitch residue Trp 256^{6.48} which describes an active GPCR when it switches from gauche to trans conformation and facilitates the interaction of Tyr 222^{5.58} and Tyr 305^{7.53}. In the CCR2 crystal structure and the crystal structures of chemokine receptors CCR5 and CXCR4, Trp 256^{6.48} points upward and extends into the helical core. Each apo macrostate shows Trp 256^{6.48} in a single conformation pointing toward helix III (Figure S7). In the holo macrostates, Trp 256^{6.48} access three distinct conformations: one resembling the crystal structure but with the helix shifted slightly outward from the helical core, another that laterally twists toward helix V, and one conformation that points down into the helical core toward the G-protein binding site (Figure S7). This third conformation, in which the orthosteric ligand has dissociated, represented by the holo macrostate L, further suggests cross-talk between the chemokine binding site and the rest of the protein.

The interaction of these two tyrosines and Arg 138^{3.50} also characterizes an active state GPCR (46). In the inactive crystal structure of CCR2, Tyr 222^{5.58} points toward lipids, sterically blocked by Phe 246^{6.38} from interaction with Arg 138^{3.50} and Tyr 305^{7.53} (14). In apo macrostates, Tyr 222^{5.58} remains pointed toward the lipids, never swiveling around to interact with Arg 138^{3.50} or Tyr 305^{7.53} as occurs in activated GPCR states (Figure 2F). Holo macrostates actually show increased range of motion of Tyr 222^{5.58}, diverging from the crystal structure and stabilizing in unique intermediate conformations. The steric obstruction from Phe 246^{6.38} is alleviated in both apo and holo macrostates, as Phe 246^{6.38} swings outward and points toward the lipids. The conforma-

tions of these microswitch residues indicate that both apo and holo macrostates are sampling different conformations.

5) Formation of continuous water pathway suggests movement of apo towards activation. Internal water molecules, which may influence conformational changes in GPCRs by interfering with hydrogen bonding networks of the receptor's backbone and side chains, are postulated to be an integral part of receptor activation in GPCRs(47–50). Work in other GPCRs has additionally shown that activation can allow water and sodium ion flow through GPCR core(51). Furthermore, it has been shown that the activation of GPCRs is voltage sensitive(52). Our simulations enable the direct visualization of water and sodium ion density in both CCR2 ensembles.

A continuous internal water pathway forms in apo CCR2 (Figure 3A). The antagonists disrupt this water pathway, slowing the rate of water entry to and egress from the protein core (Figure 3A). An analysis of the water occupancy per residue (Figure S8A) indicates that several of the high water occupancy residues (e.g. Asp 36^{1.26}, Ser 50^{1.40}, Glu 235^{6.27}, Lys 236^{6.28}, Glu 310^{8.48}, Lys 311^{8.49}) may be exposed to more water in the apo simulations than in the holo simulations simply because the ligands have been removed and the water has access to the binding pockets. The other residues (e.g., Asp 78^{2.40}, Tyr 80^{2.42}, Asp 88^{2.50}, Leu 92^{2.54}, Ile 93^{2.55}, Gly 127^{3.39}, Ile 128^{3.40}, Glu 291^{7.39}, and Phe 312^{8.50}) reside in the protein core, along the continuous pathway (Figure 3A).

Class A GPCRs also possess a conserved sodium binding site at Asp^{2.50} corresponding to Asp 88 in CCR2(53). The dynamics of activation was previously hypothesized to impinge upon the sodium binding pocket, eventually leading to ion permeation into the cytosol(54). A sodium ion occupancy per residue analysis (Figure S8B) indicates that, while no sodium permeation events were observed in the apo trajectories, ions interact with residues Asp 88^{2.50}, Glu 291^{7.39}, and His 297^{7.45} which form and surround the sodium binding site. In holo CCR2, the ligands do not permit sodium ions to enter the protein core; instead, ions interact primarily with residues Glu 270^{6.62} and His 317^{8.55}-Lys 320^{8.58}.

Effects of antagonist binding on CCR2 dynamics

Comparisons of the apo and holo MSMs elucidate the effects of antagonists on CCR2 dynamics. Notably, apo relaxation timescales are an order of magnitude less than holo relaxation timescales (Table S2), indicating that the antagonists greatly perturb CCR2 dynamics. The motions described by apo and holo tIC 0 represent the most striking difference between the two systems' dynamics. In the apo MSM, the inter-residue distances most closely correlated with apo tIC 0 are all a part of the allosteric (G-protein) binding pocket, whereas in the holo MSM, the inter-residue distances most closely correlated with holo tIC 0 are all a part of the orthosteric (chemokine) binding pocket (Figure S9).

Apo tIC 1 represents the flipping of Trp 98^{2.60} into the orthosteric drug binding site (Figure 4A,B, Figure S9C). In the crystal structure Trp 98^{2.60} packs with the tri-substituted cyclohexane of the orthosteric antagonist, BMS-681(14). Without the presence of this ligand, Trp 98^{2.60} assumes three distinct positions. The Trp 98^{2.60} conformation in the cluster at the neutral tIC (boxed in yellow, Figure 4A,B) most closely resembles the conformation of Trp 98^{2.60} in the active GPCR

US28, which is shifted slightly up and in towards the helical core in comparison to the CCR2 crystal structure. The two other conformations are found at the extreme ends of apo tIC 1 in densely populated free energy wells. Of these two conformations, state F assumes the most dramatic conformation and protrudes into the chemokine binding site (Figure S11). In the holo macrostates there is markedly less intrusion into the binding pocket due to the presence of the ligand.

Holo tIC 1 represents the concerted movement of 5 pairs of residues in the orthosteric ligand binding site during this event of the orthosteric ligand dissociation (Figure 4C,D; Figure S9D). The separation projected in the first two tICs (Figure 4D) is divided into clusters of frames that occur before (white clusters), during (grey), and after (black) dissociation. The residue pairs identified by tICA that contribute to holo tIC 1 and this ligand dissociation (Figure S9D) were confirmed by analyzing the original simulation data. The key changes are the change in distance between Tyr 49^{1,39} - Thr 292^{7,40}, Trp 98^{2,60} - Tyr 120^{3,32}, Ser 50^{1,40} - Tyr 259^{6,51}, and the chi angle of Glu 291^{7,39}. Notably, four of these residues (Tyr 49^{1,39}, Trp 98^{2,60}, Tyr 120^{3,32}, and Thr 292^{7,40}) are not only involved in binding to the co-crystallized orthosteric antagonist BMS-681 and/or CCL2 binding, but are also critical for GPCR activation (55, 56).

In the CCR2 crystal structure, there is a hydrogen bond between Tyr 49^{1,39} and Thr 292^{7,40}. The gamma-lactam secondary exocyclic amine of the orthosteric ligand forms a hydrogen bond with the hydroxyl of Thr 292^{7,40}, and the carbonyl oxygen of the gamma-lactam forms a hydrogen bond with Tyr 49^{1,39}. During simulation, the distance between Tyr 49^{1,39} and Thr 292^{7,40} averages 0.4 nm until 3 μ s after the ligand dissociates, when it begins fluctuating between 1.3 nm to 0.3 nm (Figure S12). This suggests that the orthosteric ligand dissociation breaks the hydrogen bond between these key ligand binding residues. This motion is captured in the holo MSM: the separation of the two residues is exemplified between States H (pre-dissociation) and L (post-dissociation) in Figure S12A,B. The distance between Ser 50^{1,40} and Tyr 259^{6,51} is also a contributor to holo tIC 1, and shows the same outward movement of helix I. In Figure S12C, we see a slight decrease in distance between the residue pair, followed by the same lag time of 3 μ s as discussed above, and finally an increase in distance as the extracellular end of helix I bends away from the helical bundle.

The positioning of the orthosteric ligand and the conformation of Trp 98^{2,60} are closely linked (Figure 4C, S13). After ligand dissociation, in holo states K and L (purple and white, respectively), Trp 98^{2,60} turns towards helix III, bending slightly inward toward the chemokine binding site. Prior to ligand dissociation, Trp 98^{2,60} has two distinct conformations. In the first conformation (states I and G, black and yellow), the ligand positions itself between helices I and VII, in the same conformation as the crystal structure. Trp 98^{2,60} is constrained in a downward position, pointing intracellularly, also resembling the CCR2 crystal structure conformation and the crystal structure of chemokine receptor CCR5. In the second conformation (States H and J, cyan and grey), Trp 98^{2,60} flips up and out of the binding pocket, pointing extracellularly, and the ligand moves between helices I and II. This conformation of Trp 98^{2,60} closely resembles the conformation of chemokine receptor CCR9. The third conformation of Trp 98^{2,60} is found

in State M (red), and is the most prominent position of the residue as it extends deeper into the chemokine binding site toward helix III. In this case, the ligand interacts with helices II, IV, and V, and there are no transitions from this state to a dissociated state.

As in the apo MSM, the absence of the orthosteric ligand causes a shift in the position of Trp 98^{2,60}. In the holo simulations shown in Figure S14, the dissociation event is preceded by a doubling of the distance between Trp 98^{2,60} and Tyr 120^{3,32}, and 3 μ s after dissociation the distance returns to its previous 0.4 nm. This increase in distance may be required for the ligand to begin the process of dissociating. Another drastic change during the dissociation event is the switch of Glu 291^{7,39} from a constrained chi angle of -50 to -100 degrees to an unconstrained 150 to -150 degrees (Figure S14). After dissociation, this angle more closely resembles the conformation in all apo simulations.

Finally, the faster dominant motions (tICs 2, 3, 4) in the holo MSM consist of rearrangements in the allosteric ligand binding site, suggesting that an allosteric rearrangement must first happen in order for the orthosteric ligand to dissociate. Further evidence for this is the observed correlated motion of the downward flip of the conserved residue Tyr 305^{7,53} in the G-protein binding site with the dissociation of the orthosteric ligand from the chemokine binding site.

Overall, holo macrostates show more helical tilting, bending, and binding site expansion, which increases the solvent-accessible surface area (SASA) when compared to the crystal structure and the apo macrostates. However, the apo simulations have greater overall motion and residue fluctuation, suggesting that the antagonist ligands dampen CCR2 dynamics (Figure S15).

Opening of a cryptic druggable pocket. A dramatic expansion of the extracellular (chemokine) binding site is exhibited in the holo macrostates. The expansion is caused by a pronounced outward tilting of helix VI and slight outward tilting of helix II in the holo macrostates, whereas the apo macrostates show the opposite, with a slight inward tilting of both helices VI and II (Figure 2C,E). The intracellular (G-protein) binding site also enlarges in the holo macrostates due to the outward shift of the intracellular ends of helices V and VI, but remains obstructed in all apo and holo states. In the crystal structure, this obstruction occurs by the interaction of Arg 138^{3,50} with Asp 137^{3,49} and with Thr 77^{2,39} (14), which are maintained throughout all the simulations. The outward movement of the intracellular end of helix VI and the movement of helix V toward helix VI in states L, J, and H in the holo MSM create a putative site for novel allosteric antagonists; this pocket also transiently appears in the apo simulations (Figure 5). Computational solvent mapping(57) of this novel site indicates that the pocket presents surfaces that are amenable to ligand binding due to its ability to bind clusters of multiple different drug-like probes (Figure 5C). The pocket can be accessed through the lipid bilayer between helices IV and V, or from the G-protein binding site, as a deeper extension of the current allosteric binding site of CCR2-RA-[R], and may be useful for rational drug design or modification of current antagonists.

Conclusions

To characterize the basal dynamics of CCR2 and understand how small molecule antagonists modulate these dynamics, we coupled long timescale atomic simulations and MSM theory to compare the metastable states accessed by apo and holo CCR2 in its native membrane-embedded form. Several intermediate states reveal a novel cryptic binding site that could be targeted with small molecule inhibitors. The kinetics, dynamics, and conformational ensembles of the apo and holo systems differ greatly: antagonists dampen CCR2 dynamics, slow transitions between metastable states, and are associated with different key motions and residue rearrangements. Without antagonists, CCR2 is able to access other distinct metastable states that are likely sampling along an activation pathway. These intermediate states inform on the basal dynamics of CCR2, and may be useful for modification of previously unsuccessful drugs.

Digital Figures.

Materials and Methods

See the Supporting Information for full Materials and Methods. MD trajectories and MSM construction scripts will be available for download.

System Preparation and Molecular Dynamics Simulations. Two systems were simulated: CCR2 holo, with both co-crystallized antagonist ligands bound, and CCR2 apo, without ligands bound. CCR2-RA-[R] and BMS 681(14) were deleted to build the apo system. Each all-atom system is embedded in a biologically similar POPC bilayer, explicitly solvated with TIP3P, and simulated with 150mM NaCl, at pH 7.4, at 310K and 1 bar. The initial coordinates were taken from the experimental crystal structure(14) and simulated for 50ns MD simulations on local resources before simulation on Anton2. The Anton2 simulations were run in the NPT ensemble, using Anton's Nose-Hoover thermostat-barostat, at 310K and 1 bar with a 2.5-fs time-step and particle mesh Ewald electrostatic approximation. The two systems were simulated for an aggregate total of 260 microseconds.

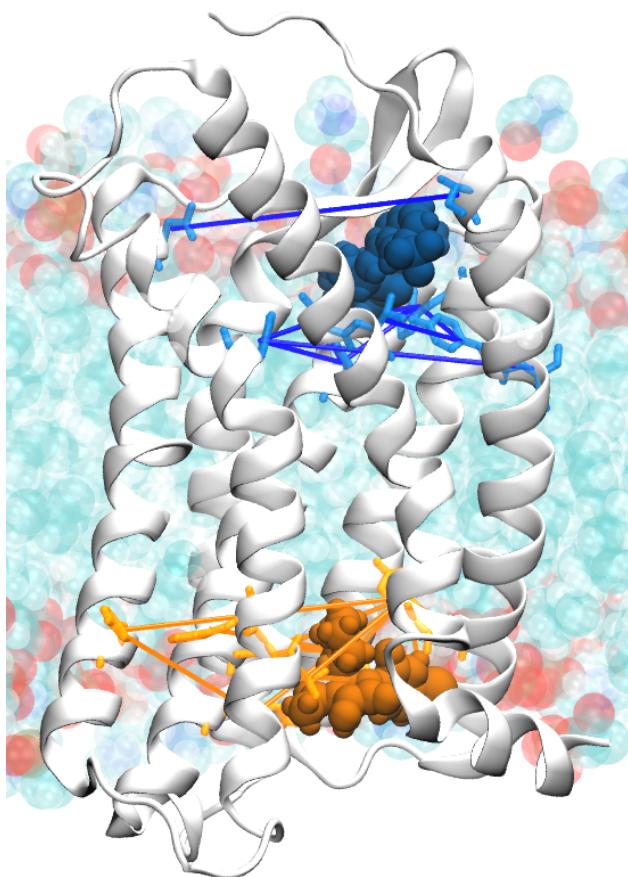
Building the Markov State Models. PyEMMA(39) was used for feature selection, dimensionality reduction, clustering, MSM building, and MSM analysis. Dimensionality reduction using TICA(36, 37) was performed on all trajectories. The apo and holo systems were clustered separately using K-means. The models were selected based on implied timescale plots and Chapman-Kolmogorov tests. MSMs were coarse-grained with hidden Markov models (HMMs). Representative states were selected by identifying the most populated cluster in each metastable state, and using MSMBuilder(58) to find the centroid of that cluster.

ACKNOWLEDGMENTS. The authors thank Tracy Handel and Irina Kufareva for their valuable input and discussions as well as initial modification of CCR2 coordinates. We also thank the organizers and participants of the 2018 Workshop on Free energy Methods, Kinetics and Markov State Models in Drug Design for helpful comments and discussion. This work was supported in part by the Director's New Innovator Award Program NIH DP2-OD007237, the National Biomedical Computation Resource (NBCR) NIH P41-GM103426, and the National Science Foundation through XSEDE supercomputing resources provided via TG-CHE060073 to R.E.A. C.T.L. also acknowledges support from the NIH Molecular Biophysics Training Program (T32-GM008326) Anton 2 computer time was provided by the Pittsburgh Supercomputing Center (PSC) through Grant R01GM116961 from the National Institutes of Health. The Anton 2 machine at PSC was generously made available by D.E. Shaw Research.

References

1. Ben-Baruch A (2006) The multifaceted roles of chemokines in malignancy. *Cancer and Metastasis Reviews* 25(3):357–371.
2. O'Connor T, Borsig L, Heikenwalder M (2015) CCL2-CCR2 signaling in disease pathogenesis. *Endocrine, Metabolic and Immune Disorders - Drug Targets* 15(2):105–118.
3. Solomon M, Balasa B, Sarvetnick N (2010) CCR2 and CCR5 chemokine receptors differentially influence the development of autoimmune diabetes in the NOD mouse. *Autoimmunity* 43(2):156–163.
4. Pease JE, Horuk R (2009) Chemokine receptor antagonists: Part 1. *Expert Opinion on Therapeutic Patents* 19(1):39–58.
5. Pease JE, Horuk R (2009) Chemokine receptor antagonists: part 2. *Expert opinion on therapeutic patents* 19(2):199–221.
6. Horuk R (2009) Chemokine receptor antagonists: overcoming developmental hurdles. *Nature reviews. Drug discovery* 8(1):23–33.
7. Scholten DJ, et al. (2012) Pharmacological modulation of chemokine receptor function. *Br. J. Pharmacol* 165(6):1617–1643.
8. Lim S, Yuzhalin A, Gordon-Weeks A, Muschel R (2016) Targeting the CCL2-CCR2 signaling axis in cancer metastasis. *Oncotarget* 7(19):28697–710.
9. Solari R, Pease JE, Begg M (2015) Chemokine receptors as therapeutic targets: why aren't there more drugs? *Eur. J. Pharmacol.* 746:363367.
10. (2018) <https://www.clinicaltrials.gov/>.
11. Shahlaei M, Fassihai A, Papaleo E, Pourfarzam M (2013) Molecular dynamics simulation of chemokine receptors in lipid bilayer: a case study on C-C chemokine receptor type 2. *Chemical biology & drug design* 82(5):534–545.
12. Chavan S, Pawar S, Singh R, Sobhia ME (2012) Binding site characterization of G protein-coupled receptor by alanine-scanning mutagenesis using molecular dynamics and binding free energy approach: application to C-C chemokine receptor-2 (CCR2). *Molecular Diversity* 16(2):401–413.
13. Kothandani G, Gadhe CG, Cho SJ (2012) Structural insights from binding poses of CCR2 and CCR5 with clinically important antagonists: a combined in silico study. *PLoS One* 7(3):e32864.
14. Zheng Y, et al. (2016) Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. *Nature* 540(7633):458–461.
15. Latorraca NR, Venkatakrishnan AJ, Dror RO (2017) GPCR Dynamics: Structures in Motion. *Chemical Reviews* 117(1):139–155.
16. Venkatakrishnan AJ, et al. (2013) Molecular signatures of G-protein-coupled receptors. *Nature* 494:185–194.
17. Zhang Q, Zhou M, Zhao L, Jiang H, Yang H (2018) Dynamic States of the Ligand-Free Class A G Protein-Coupled Receptor Extracellular Side. *Biochemistry* 0(0):null.
18. Katritch V, Cherezov V, Stevens R (2013) Structure-Function of the G-protein-Coupled Receptor Superfamily. *Annual review of pharmacology and toxicology* 53:531–556.
19. Manglik A, et al. (2015) Structural Insights into the Dynamic Process of beta(2)-Adrenergic Receptor Signaling. *Cell* 162:1431–1431.
20. Katritch V, Cherezov V, Stevens RC (2012) Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol. Sci.* 33:17–27.
21. Malik RU, et al. (2013) Detection of G protein-selective G protein-coupled receptor (GPCR) conformations in live cells. *J. Biol. Chem.* 288:17167–17178.
22. Yao XJ, et al. (2009) The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex. *Proc. Natl. Acad. Sci. U. S. A.* 106:9501–9506.
23. Nygaard R, et al. (2013) The Dynamic Process of beta(2)-Adrenergic Receptor Activation. *Cell* 152:532–542.
24. Bockenhauer S, Furstenberg A, Yao XJ, Kobilka BK, Moerner WE (2011) Conformational dynamics of single G protein-coupled receptors in solution. *J. Phys. Chem. B* 115:13328–13338.
25. Carter PH, et al. (2015) Discovery of a Potent and Orally Bioavailable Dual Antagonist of CC Chemokine Receptors 2 and 5. *ACS Med Chem Lett.* 6(4):439444. PMID: 25893046.
26. Dassa OA, et al. (2007) Novel, Acidic CCR2 Receptor Antagonists: Lead Optimization. *Letters in Drug Design and Discovery* 4(4):263–271.
27. (2014) Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer in *International Conference for High Performance Computing, Networking, Storage and Analysis*, SC, No. January, pp. 41–53.
28. Bowman GR, Voelz Va, Pande VS (2011) Taming the complexity of protein folding. *Current Opinion in Structural Biology* 21(1):4–11.
29. Swope WC, Pitera JW (2004) Describing Protein Folding Kinetics by Molecular Dynamics Simulations. *J. Phys. Chem. B* 108:6571–6581.
30. Swope W, et al. (2004) Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 2. Example Applications to Alanine Dipeptide and a beta-Hairpin Peptide. *J. Phys. Chem. B*.
31. Singhal N, Snow CD, Pande VS (2004) Using path sampling to build better Markovian state models: Predicting the folding rate and mechanism of a tryptophan zipper beta hairpin. *Journal of Chemical Physics*.
32. Malmstrom RD, Lee CT, Van Wart AT, Amaro RE (2014) Application of molecular-dynamics based markov state models to functional proteins. *Journal of Chemical Theory and Computation*.
33. Amaro RE, Mulholland AJ (2018) Multiscale methods in drug design bridge chemical and biological complexity in the search for cures. *Nature Reviews Chemistry* 2:0148.
34. Amaro RE, et al. (2018) Ensemble Docking in Drug Discovery. *Biophysical Journal* 114(10):2271–2278.
35. Bowman GR, Pande VS, Noe F (2014) *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*. (Springer Netherlands) Vol. 797.
36. Schwantes CR, Pande VS (2013) Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. *J. Chem. Theory Comput.* 9:2000–2009.
37. Perez-Hernandez G, Paul F, Giorgino T, De Fabritis G, Noe F (2013) Identification of slow

a



b

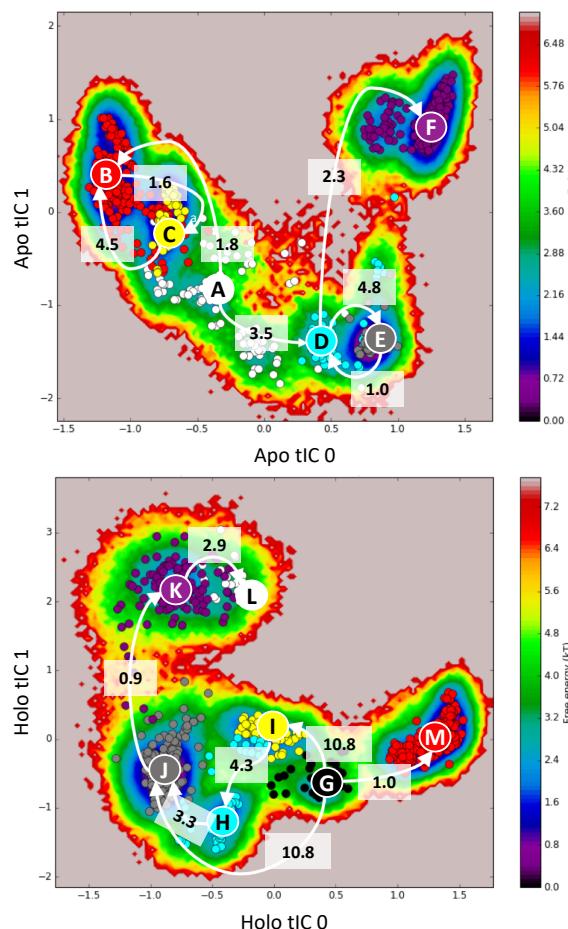


Fig. 1. MD simulations of CCR2 in a lipid bilayer were performed for unbound (apo) and dual-antagonist-bound (holo) systems. A) Sets of residue pairs surrounding the two ligand binding pockets were used as tICA inputs. The protein is shown in white cartoon. Lipids are teal, red, and blue. The orthosteric and allosteric ligands are shown in blue and orange, respectively, with interresidue pair distances denoted by similarly colored lines. The free energy and HMMs of B) apo and C) holo CCR2, projected here in the first two tICA components. Coarse-grained states are labeled and colored. Transition rates between macrostates are represented by white arrows and are in units of μs^{-1} .

- molecular order parameters for Markov model construction. *J. Chem. Phys.* 139:015102.
38. Nüske F, Keller BG, Pérez-Hernández G, Mey ASJS, Noé F (2014) Variational approach to molecular kinetics. *Journal of Chemical Theory and Computation* 10(4):1739–1752. PMID: 26580382.
 39. Scherer MK, et al. (2015) PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *Journal of Chemical Theory and Computation* 11(11):5525–5542.
 40. Prinz JH, et al. (2011) Markov models of molecular kinetics: Generation and validation. *Journal of Chemical Physics* 134(17).
 41. Lebon G, et al. (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* 474:521–525.
 42. Burg JS, et al. (2015) Structural basis for chemokine recognition and activation of a viral g protein-coupled receptor. *Science* 347(6226):1113–1117.
 43. Tan Q, et al. (2013) Structure of the CCR5 Chemokine Receptor-HIV Entry Inhibitor Maraviroc Complex. *Science* 341(6152):1387–1390.
 44. Oswald C, et al. (2016) Intracellular allosteric antagonism of the CCR9 receptor. *Nature* 540:462–465.
 45. Qin L, et al. (2015) Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science* 347(6226):1117–1122.
 46. Caliman A, Swift S, Wang Y, Miao Y, McCammon J (2015) Investigation of the conformational dynamics of the apo A2A adenosine receptor. *Protein Science: A Publication of the Protein Society* 24:1004–1012.
 47. Jastrzebska B, Palczewski K, Golczak M (2011) Role of bulk water in hydrolysis of the rhodopsin chromophore. *The Journal of biological chemistry* 286(21):18930–7.
 48. Angel TE, Chance MR, Palczewski K (2009) Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proceedings of the National Academy of Sciences of the United States of America* 106(21):8555–8560.
 49. Choe HW, et al. (2011) Crystal structure of metarhodopsin II. *Nature* 471:651655.
 50. Huang W, et al. (2015) Structural insights into μ -opioid receptor activation. *Nature* 524:315321.
 51. Yuan S, Filipek S, Palczewski K, Vogel H (2014) Activation of G-protein-coupled receptors correlates with the formation of a continuous internal water pathway. *Nature communications* 5:4733.
 52. Rinne A, Birk A, Bünnemann M (2013) Voltage regulates adrenergic receptor function. *Proceedings of the National Academy of Sciences of the United States of America* 110(4):1536–41.
 53. Katritch V, et al. (2014) Allosteric sodium in class A GPCR signaling. *Trends in biochemical sciences* 39(5):233–44.
 54. Vickery ON, et al. (2018) Intracellular Transfer of Na^+ in an Active-State G-Protein-Coupled Receptor. *Structure* 26(1):171–180.
 55. Berkhouit TA, et al. (2003) CCR2: characterization of the antagonist binding site from a combined receptor modeling/mutagenesis approach. *Journal of medicinal chemistry* 46(19):4070–86.
 56. Hall SE, et al. (2009) Elucidation of binding sites of dual antagonists in the human chemokine receptors CCR2 and CCR5. *Molecular pharmacology* 75(6):1325–36.
 57. Kozakov D, et al. (2015) The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nature Protocols* 10(5):733–755.
 58. Beauchamp KA, et al. (2011) MSMBuilder2: Modeling conformational dynamics on the picosecond to millisecond scale. *Journal of Chemical Theory and Computation* 7(10):3412–3419.

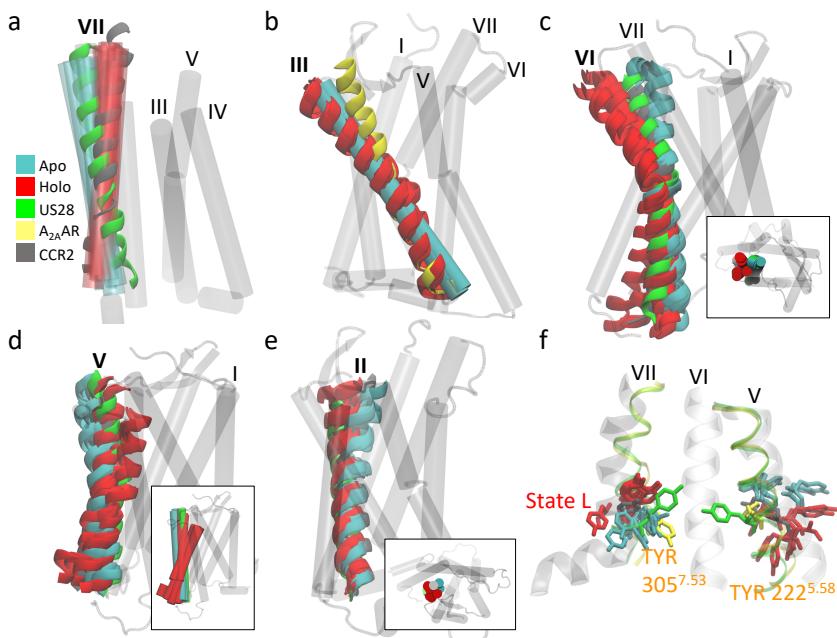


Fig. 2. Apo and holo macrostates are compared to the active crystal structure of US28 (green, PDB ID 4XT3) or the active crystal structure of A_{2A}AR (yellow, PDB ID 2YDO) and the crystal structure of CCR2 (grey, PDB ID 5T1A). A) Helix VII of apo macrostates resemble the active conformation of US28; holo macrostates resemble the CCR2 crystal structure. B) The conformation of helix III in apo macrostates subtly resemble active A_{2A}AR; holo macrostates are tilted away from the center of the helical bundle. C,D) Helix V and VI of apo macrostates straighten or tilt in toward the center of the protein, similar to the active conformation of US28 and the inactive CCR2 crystal structure; helix V and VI of several holo macrostates tilt away from the binding sites, accessing more active-like conformations than the apo macrostates or even the active state of US28. E) Helix II in the apo macrostates shifts inward; in the holo macrostates it shifts outward. F) In licorice are conserved motifs TYR 305^{7.53} and TYR 222^{5.58}. All six apo metastable state assume a new conformation for TYR 305^{7.53}, pointing intracellularly and in a similar conformation to active A_{2A}AR. Six out of the seven holo metastable states have TYR 305^{7.53} in the same conformation as the equilibrated crystal structure. Post-ligand-dissociation holo state L assumes a new position of TYR 305^{7.53}, more similar to the dominant apo conformation. Apo metastable states sample a narrower range of conformations for TYR 222^{5.58} than holo.

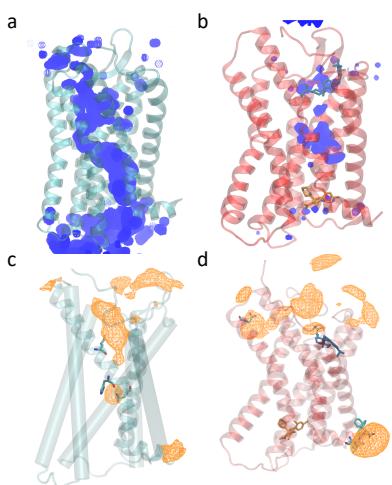


Fig. 3. Ligands disrupt a continuous internal water and sodium ion pathway. Average water density over a 50 microsecond simulation of A) Apo (teal) and B) Holo (red). The orthosteric ligand is shown in blue and the allosteric ligand is shown in orange. Total average sodium ion density in C) Apo and D) Holo. Highest occupancy residues depicted in cyan licorice and plotted in Figure S8.

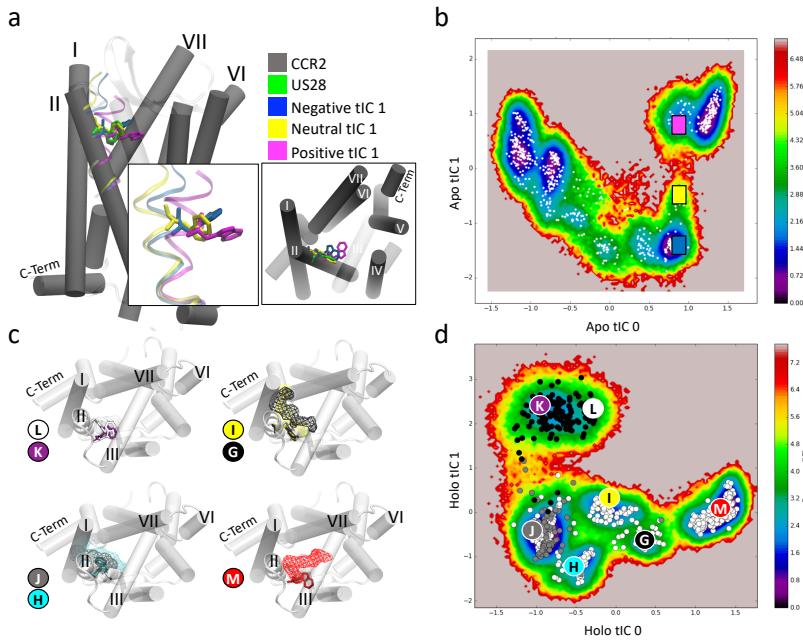


Fig. 4. A) In apo CCR2, tICA 1 represents TRP 98^{2,60} in three distinct positions. In gray is the crystal structure; in green is the active crystal structure of US28; in blue and yellow are transitions, and in magenta is the most dramatic conformation. Each conformation is plotted on the free energy in tICA space in B). C) In holo CCR2, the positioning of the orthosteric ligand and the conformation of TRP 98^{2,60} is closely linked. Shown in light silver cartoon is CCR2 5T1; TRP 98^{2,60} is displayed as purple in state K, white in state L, gray in state J, cyan in state H, yellow in state I, black in state G, and red in state M. D) Holo CCR2. White circles are clusters of frames before any ligand dissociation. Grey circles are clusters of frames during the event. Black circles are clusters of frames after the event.

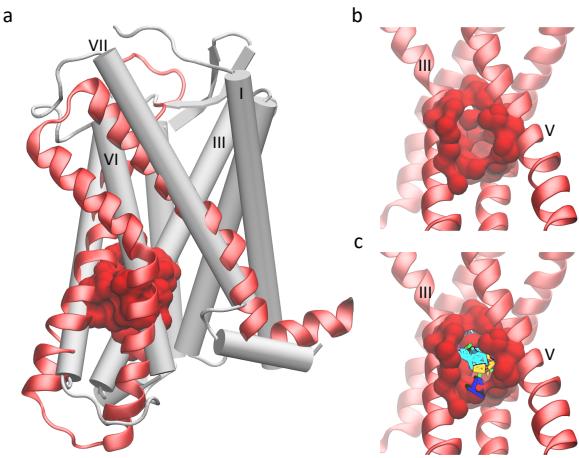


Fig. 5. A putative allosteric drug binding pocket is revealed by three holo macrostates. A) A comparison of the CCR2 crystal structure (white cartoon) with helices V, VI, and VII (red new cartoon) of one holo macrostate. The pocket is shown in red surface. B) A closer view of the pocket from the other side of the protein, between helices III and V. C) Small organic probes used for computational fragment mapping are multicolored.