Diverse GPCRs exhibit conserved water networks for stabilization and activation

A. J. Venkatarkrishnan, Anthony K. Ma, Rasmus Fonseca, Naomi R. Latorraca, Brendan Kelly, Robin M. Betz, Chaitanya Asawa, Brian K. Kobilka, and Ron O. Dror

Diverse GPCRs exhibit conserved water networks for stabilization and activation

G protein-coupled receptors (GPCRs) have evolved to recognize incredibly diverse extracellular ligands while sharing a common architecture and structurally conserved intracellular signaling partners. It remains unclear how binding of diverse ligands brings about GPCR activation, the common structural change that enables intracellular signaling. Here, we identify highly conserved networks of water-mediated interactions that play a central role in activation. Using atomic-level simulations of diverse GPCRs, we show that most of the water molecules in GPCR crystal structures are highly mobile. Several water molecules near the G protein-coupling interface, however, are stable. These water molecules form two kinds of polar networks that are conserved across diverse GPCRs: (i) a network that is maintained across the inactive and the active states and (ii) a network that rearranges upon activation. Comparative analysis of GPCR crystal structures independently confirms the striking conservation of water-mediated interaction networks. These conserved water-mediated interactions near the G protein-coupling region, along with diverse water-mediated interactions with extracellular ligands, have direct implications for structure-based drug design and GPCR engineering.

GPCR dynamics | water molecules | activation | molecular dynamics simulation | polar network

GPCR protein-coupled receptors (GPCRs) are membrane proteins that act as a control panel for our cells: they recognize the presence of extracellular molecules such as hormones or neurotransmitters and stimulate intracellular signaling pathways in response. Humans have over 800 GPCRs, which share a common architecture comprising seven transmembrane (TM) helices (1). GPCRs transmit signals across the cell membrane by transitioning from an inactive conformational state to an active conformational state. The active conformational state, which is favored by binding of certain ligands to the extracellular side of the GPCR, couples to and stimulates G proteins. Available structures indicate that diverse GPCRs undergo similar conformational changes upon activation (2, 3). However, these GPCRs have evolved to recognize dramatically different ligands: human GPCRs recognize molecules ranging from amines, nucleosides, lipids, and steroids to entire proteins, in most cases with high specificity. This ability to stimulate cellular pathways in response to the presence of specific extracellular ligands makes GPCRs excellent drug targets, and in fact roughly one-third of all drugs act by binding to GPCRs (4). The process by which diverse ligands bind to GPCRs and favor activation is of great interest in molecular biology and drug discovery but remains incompletely understood.

Recent high-resolution crystal structures reveal dozens of water molecules buried in the transmembrane region of GPCRs (5–10). Water has been shown to play a critical role both in the binding of drugs to their targets (11) and in the function of many proteins (12, 13), raising the question of how the water molecules in GPCRs behave and what function they might serve. Indeed, this question was raised even before the high-resolution structures were available (14, 15).

Unfortunately, answering this question from crystal structures alone is difficult, because each crystal structure represents a static snapshot. In reality, water molecules may be highly mobile, and GPCRs themselves are constantly changing conformation (16). Moreover, most of the available GPCR crystal structures were solved at cryogenic temperatures at which water molecules are expected to be much less mobile than at body temperature. Even recovering average positions of water molecules from crystallographic data can be difficult when the corresponding electron density is weak.

Here, we investigate the structural dynamics of water molecules in a diverse set of GPCRs using molecular dynamics (MD) simulations, which allow us to follow the trajectory of each individual water molecule at physiological temperature and in a native-like membrane environment. We compare water dynamics not only between different GPCRs but also between the active and inactive conformational states of the same GPCR. Our results show that the water molecules observed in GPCR crystal structures are not all equal: a few are stable in their crystallographically observed position, but most are highly mobile, spending only a small fraction of their time near that position.

Remarkably, stable water molecules near the G protein-binding site form a network of polar interactions (hydrogen bonds) that rearranges upon receptor activation. This network was almost perfectly conserved across the GPCRs that we examined, in sharp contrast to water-mediated interactions with ligands, which

Significance

G protein-coupled receptors (GPCRs) represent both the largest class of drug targets and the largest family of human membrane proteins. Recent three-dimensional structures have revealed the presence of many water molecules buried inside GPCRs, but the functional role of these water molecules has been unclear. Using extensive atomic-level computer simulations, we find that, although most of these waters are highly mobile, a few are stable. These stable water molecules mediate state-dependent and state-independent polar networks that are conserved across diverse GPCRs. These findings promise to help guide the rational design of GPCR-targeted drugs.

Author contributions: A.J.V. and R.O.D. designed research; A.J.V., A.K.M., R.F., N.R.L., B.K., R.M.B., C.A., and R.O.D. performed research; A.J.V. and R.F. contributed new reagents/analytic tools; A.J.V., A.K.M., B.K.K., and R.O.D. analyzed data; and A.J.V., A.K.M., B.K.K., and R.O.D. wrote the paper.

Reviewers: D.E.G., University of Copenhagen; and L.P., Universitat Autonoma de Barcelona.

Conflict of interest statement: B.K.K. is a cofounder of and consultant for ConfometRx, Inc.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1 A.J.V. and A.K.M. contributed equally to this work.
2 To whom correspondence may be addressed. Email: kobilka@stanford.edu or ron.dror@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1809251116/-/DCSupplemental.

Published online February 6, 2019.
vary widely from one GPCR to another. This conserved water network appears to play a key role in the GPCR activation mechanism that was preserved asGPCRs evolved to recognize different ligands.

**Results**

We focused on class A GPCRs, as they represent the largest and most diverse class of GPCRs. We chose functionally distinct and evolutionarily distant ligand-activated class A GPCRs for which both inactive-state and active-state structures are available: β2 adrenergic receptor (β2AR) (17, 18), M2 muscarinic receptor (M2R) (19, 20), μ-opioid receptor (MOR) (5, 21), and adenosine A2A receptor (A2A R) (6, 22). We also included δ-opioid receptor (DOR) as it has a particularly high-resolution inactive-state crystal structure (1.8 Å) (7). Collectively, these GPCRs bind to different types of ligands (amines, peptides, nucleosides) and different G proteins (Gs, Gi). All simulations included an explicitly represented lipid bilayer, water, salt ions, and the cocrystallized ligand (SI Appendix, Table S1).

**Only a Minority of Crystallographic Water Molecules Are Stable.** In simulations at physiological temperature, only a few of the waters resolved in high-resolution GPCR crystal structures remain near their crystallographically observed positions (Fig. 1 A and SI Appendix, Fig. S1). The position of a given molecule was tracked using its 3D coordinates after aligning the simulation frames on the residues proximal to the water molecule (SI Appendix, Materials and Methods). Our simulations are consistent with crystallographic data in that the same regions of the receptor are populated by waters in both crystal structures and simulations (SI Appendix, Fig. S1), but in simulation, most of the waters are highly mobile, showing little preference for the specific positions occupied by waters in the crystal structures. A few crystallographic water positions, however, are consistently occupied by water molecules in simulation. We refer to these crystallographic water molecules as “stable,” even though an individual water molecule at such a position is generally replaced by a different water molecule at the same position at least every few hundred nanoseconds. In DOR simulations, for example, the water in the “kink” of transmembrane helix 6 (TM6) was highly stable with a water molecule observed within 1 Å of the crystallographic position for 99% of the time. On the other hand, the waters within the transmembrane helical bundle were highly mobile; for some of these, a water molecule was found within 1 Å of the crystallographic position less than 10% of the time.

To further characterize the behavior of water molecules in the structural framework of the receptor, we investigated the polar interactions (hydrogen bonds) in the receptor formed by the water molecules (Fig. 1 B). We focused on water-mediated interactions in which a pair of receptor residues is connected by one or two water molecules in the network of hydrogen bonds. We estimated the stability of a given water-mediated interaction based on the frequency of the interaction’s presence in simulation (SI Appendix, Materials and Methods).

Only a fraction of the crystallographically observed water-mediated interactions are stable in simulation because a majority of the water molecules are highly mobile, forming many transient, alternative polar interactions. For example, of the 42 water-mediated interactions in the DOR crystal structure, only 19 are formed in 60% or more of simulation frames (Fig. 1 B and SI Appendix, Fig. S2).

**Conserved Networks of State-Independent and State-Dependent Water-Mediated Interactions in GPCR Activation.** Might water molecules play a role in GPCR activation? To address this question, we first compared water-mediated interactions in simulations of inactive and active GPCR states. We also compared the water-mediated interactions across different GPCRs (β2AR, A2A R, MOR, and M2R). For the comparisons, we analyzed interactions between residues in structurally equivalent positions at different receptors. Structurally equivalent residues were identified using GPCRdb numbers (23), which are receptor-independent generic numbers for referencing structurally equivalent positions in GPCRs (SI Appendix, Materials and Methods).

We found that a surprisingly large set of stable water-mediated interactions (i.e., those formed over 60% of the time in simulation) was shared across all receptors examined in a given activation state (Fig. 2 and SI Appendix, Fig. S3): six interactions mediated generally by four waters in the inactive state and 12 interactions mediated generally by five waters in the active state. These conserved interactions represent 20–30% of the stable water-mediated interactions observed in each receptor individually. In contrast, a majority of the remaining stable interactions are found in just one of A2A R, β2AR, M2R, and MOR. For example, of 32 stable water-mediated interactions in active-state M2R, 12 are shared across all four active-state receptors, and 14 of the remaining 20 are unique to M2R.

Among the conserved, stable, water-mediated interactions, three interactions are maintained in a state-independent manner; i.e., they are present in both the inactive states and the active states of A2A R, β2AR, M2R, and MOR. While one of these interactions stabilizes a helical kink of TM6, the other two water molecules are proximal to the G protein-coupling region and mediate interactions between residues on TM1 and TM2 and between residues on TM7 (Fig. 3 A).

Two additional surprising observations link the conserved, stable, water-mediated interactions to the GPCR activation

Venkatakrishnan et al.
process. First, the remaining conserved, stable water-mediated interactions are in the lower half of the GPCR, near the G protein-binding interface, where the largest structural rearrangements occur upon activation. Second, although these interactions are conserved across diverse GPCRs in a given activation state, they differ dramatically between the active and inactive states. Every conserved stable water molecule near the G protein-binding interface changes interaction partners upon activation, and a number of water molecules form conserved, stable water-mediated interactions in the active state but not in the inactive state. Water-mediated interactions between TM1, TM2, and TM7 are stably formed in the inactive states (Fig. 4A). This network is disrupted upon activation and a new stable network of water-mediated interactions is formed between TM3, TM5, TM6, and TM7 (Fig. 4A). Tyr7x53 (GPCRdb number) of the NPxxY motif acts as a switch between the inactive and active water-mediated networks. In the inactive state, Tyr7x53 interacts with a water molecule at the TM1–TM2 interface, and upon activation, Tyr7x53 switches conformation and interacts with water molecules at the interface of TM3, TM5, and TM6. Tyr7x53 has previously been shown to be an activation-associated microswitch in a water-independent role as well (2, 24).

All of the side chains involved in this rewiring of interactions with stable waters upon activation are highly conserved as polar residues across class A GPCRs (>89%) (SI Appendix, Table S2). Taken together, our observations indicate that stable waters near the G protein-coupling interface form a conserved polar interaction network that rearranges upon receptor activation. The...
Materials and Methods

The conserved inactive state-dependent network identified from crystal structures is identical to the conserved, stable water-mediated interaction network found in the MD simulations of inactive states (Fig. 4 A and B). Similarly, we also found a network of three water-mediated interactions formed by Tyr7x53 to be maintained exclusively in the crystal structures of active-state rhodopsin, chemokine receptor US28, and μ-opioid receptor (Fig. 4B). All of the three water-mediated interactions identified from the crystal structures of active states are present in the network of conserved, stable water-mediated interactions found in the MD simulations of the active states (Fig. 4 A and B).

Taken together, the comparative analysis of crystal structures supports the presence of conserved state-dependent and state-independent water-mediated interactions in GPCRs.

Stable Water Molecules in the Ligand-Binding Pocket Vary Between Different GPCRs. A number of the water molecules in high-resolution crystal structures of GPCRs form direct interactions with the ligand, and previous studies have pointed out the importance of water in determining ligand affinities and binding rates (25). We thus also examined the occupancy of waters in the binding pocket and the stability of water-mediated interactions between ligands and receptor residues in simulation.

Most waters in the binding pocket are highly mobile. Certain waters and water-mediated interactions, however, are stable. For example, in simulations of inactive-state DOR, two waters in the binding pocket are highly stable near their crystallographic positions (SI Appendix, Fig. S4A). These two stable waters mediate interactions between a hydroxyl group on the ligand and polar residues on TM5 and TM6.

No stable water-mediated interactions between the ligand and the receptor appear to be conserved across diverse GPCRs. Our simulations showed no such interactions conserved across all of the receptors examined, and only very limited structurally equivalent interactions between subsets of these receptors (SI Appendix, Fig. S4B). This variability in water-mediated interactions across
ligand-binding pockets suggests that, unlike the G protein-coupling region, the binding pocket has been adapted to exploit different water-mediated interactions in different GPCRs as the GPCR family evolved.

**Discussion**

The human genome includes hundreds of class A GPCRs, and these receptors have evolved to recognize extremely diverse extracellular ligands—ranging from nucleosides and amines to peptides and even entire proteins—often with very high specificity. On the other hand, all of these GPCRs couple to the same set of intracellular G proteins, which share very similar structures. One might thus imagine that GPCRs would have evolved a common activation mechanism. Indeed, previous studies have identified conserved networks of amino acid contacts and ionic interactions that appear to facilitate such a mechanism (2, 26, 27).

Our results suggest that certain water molecules also play an important, conserved role in GPCR activation. In particular, we discovered a network of highly stable water molecules near the G protein-binding site that is conserved almost perfectly across diverse class A GPCRs and that rearranges in a conserved manner upon GPCR activation. Waters also play a critical role in mediating ligand binding in GPCRs, but the water-mediated interactions in the ligand-binding pocket vary widely among class A GPCRs, likely reflecting the fact that the binding pockets of these different GPCRs have evolved to recognize different ligands.

Our results complement those of previous experimental and computational studies. Recent high-resolution crystal structures have revealed the presence of dozens of waters within GPCRs. Analyses of earlier, lower-resolution crystal structures have suggested the presence of certain conserved structural water molecules (14, 15). Radiolytic footprinting has provided insights into tightly bound waters and dynamic waters in rhodopsin (28). Simulation-based studies have shown the presence of activation-associated internal water pathways in GPCRs (28, 29) and indicated a role for waters in ligand binding (30–32).

A few caveats are in order. First, the force fields used for MD simulation are not perfect, potentially introducing artifacts into our results; simulations, however, have proven useful in elucidating water dynamics that are difficult to characterize experimentally (33, 34). Second, the strengths of hydrogen bonds in proteins can vary depending on factors such as the local structural environment (35), but our current analysis does not take this into account. Third, we lack a direct method to determine the functional importance of a water molecule in a particular position. We thus use stability and conservation across receptors as rough proxies for importance. Fourth, our analysis of stable water molecules and water-mediated interactions does not cover all of the ways in which water may contribute to GPCR function; for example, displacement of certain water molecules from the binding pocket upon ligand binding is a determinant of ligand-binding affinity (36). Finally, our analysis has focused on class A GPCRs as they are the largest and most diverse class. Future work will be necessary to understand the role of water molecules in the activation of other GPCR classes.

Our findings have several implications for GPCR drug discovery and GPCR engineering. First, understanding the behavior of water molecules in the binding pocket can aid structure-based drug discovery (25). Our results show that crystallographically resolved waters cannot all be treated as equal. One needs to characterize the dynamics of water molecules (e.g., through simulations) to identify the stable waters, which can then be exploited for drug design (37).

Second, the conserved stable waters and water-mediated interactions identified in this study can aid in structural modeling. Homology models of GPCRs are widely used for drug discovery, both when no crystal structure of the target is available and when the available structures are not in the desired activation state. The conserved water molecules identified in this study for active-state and inactive-state class A GPCRs can be incorporated into homology models and used to refine those models (see templates provided in Datasets S1 and S2). In both the inactive-state and the active-state models, the conserved water molecules at the
TM6 helical kink, between positions 1x50 and 2x50 and between positions 7x45 and 7x49, should typically be included (Fig. 3). In addition, in the inactive-state models, in the water-mediated interaction of Tyr7x53 with positions 1x50 and 2x50, an additional water molecule may be included unless Tyr7x53 can directly interact with the water molecule interacting with 1x50 and 2x50 (Fig. 4). On the other hand, in the active-state models, the water molecules mediating interactions between positions 3x43, 5x58, and 7x53 and between 6x40 and 7x53 should typically be included (Fig. 4). An additional water molecule may be included between 6x40 and 7x53, as seen in the case of US28 (Fig. 4). While including these waters in the model, one needs to ensure that the chemical nature of the interacting amino acids and their side-chain geometry allow for the formation of a hydrogen bond with the water. Similar to inclusion of conserved water molecules in homology models, these water molecules can also be incorporated into low-resolution experimental GPCR structures, such as cryo-EM structures (38), in which few waters are resolved. Third, mutating GPCR residues to stabilize either the active state or the inactive state has proven useful in GPCR crystallography as well as in functional investigations. The distinct water-mediated interaction networks associated with different activation states provide new avenues for biochemically perturbing the conformational landscape of GPCRs. In particular, residues interacting with the conserved, stable waters identified in this study are excellent candidate sites for introducing mutations to shift the conformational ensemble toward the active or inactive state.

More generally, our results reveal a “hidden” structural feature of GPCRs, a network of stable water-mediated interactions that is not evident from crystal structures alone. Our study also highlights the untapped potential of comparing structural dynamics across a drug target family to identify shared and distinct molecular features.

Materials and Methods

The MD simulation trajectories of inactive β2AR, M2R, and MOR and active β2AR were obtained from previously published studies (SI Appendix, Table S1). New MD simulations were performed for inactive DOR and A2aR and active A2aR, M2R, and MOR (SI Appendix, Materials and Methods and Table S1). The simulations were all performed using CHARMM force fields (SI Appendix, Table S1). We verified that the simulations initiated from the inactive-state and active-state structures were maintained in their respective states. Water-mediated interactions were computed for every frame of the MD simulations based on geometric criteria (SI Appendix, Materials and Methods). The stability of a water-mediated interaction between a pair of residues was defined as the fraction of frames in which either a direct or extended water-mediated interaction is formed. The occupancy of water molecules near cryo-EM crystallographic waters was calculated as the fraction of simulation frames in which a water molecule is within 1 Å of the crystallographic water position (SI Appendix, Materials and Methods).

ACKNOWLEDGMENTS. This work was supported by National Institutes of Health (NIH) Grant R01GM127359 (to R.O.D.); NIH Grants R37DA036246 and R101GM118318 (to B.K.K.); a Stanford ChEM-H seed grant (to A.J.V.); Grant NNF15SC0015268 from Novo Nordisk Foundation and Stanford Bio-X programme (to R.F.); and National Science Foundation Graduate Research Fellowships (to N.R.L. and R.M.B.). B.K.K. is a Chan-Zuckerberg Biohub Investigator.

1. Fredriksson R, Lagerström MC, Lundin L-G, Schölmerich H (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol 63:1256–1272.

2. Venkatashiraman AJ, et al. (2014) Diverse activation pathways of GPCRs. Nature 510:6470–6479.

3. Venkatashiraman AJ, et al. (2013) Molecular signatures of G-protein-coupled receptors. Nature 494:185–194.

4. Santos R, et al. (2017) A comprehensive map of molecular drug targets. Nat Rev Drug Discov 16:19–34.

5. Huang W, et al. (2015) Structural insights into μ-opioid receptor activation. Nature 524:315–321.

6. Segala E, et al. (2016) Controlling the dissociation of ligands from the adenine A2AR receptor through modulation of salt bridge strength. J Med Chem 59:6470–6479.

7. Fenali G, et al. (2014) Molecular control of μ-opioid receptor signalling. Nature 506:191–196.

8. Suno R, et al. (2018) Crystal structures of human orexin 2 receptor bound to the subtype-selective antagonist EMPA. Structure 26:7–19.e5.

9. Wang S, et al. (2017) D2 dopamine receptor high-resolution structures enable the discovery of selective antagonists. Science 358:381–386.

10. Miller-Gallacher JL, et al. (2014) The 2.1 Å resolution structure of cyanopindolol-bound human dopamine D2 receptor. Nature 513:5707–5708.

11. Isberg V, et al. (2016)GPCRs: An information system for G-protein-coupled receptors. Nucleic Acids Res 44:D356–D364.

12. Carpenter B, Tate CG (2017) Active state structures of G-protein-coupled receptors highlight the similarities and differences in the G protein and arrestin coupling interfaces.Curr Opin Struct Biol 45:124–132.

13. Jazayeri A, Dias JM, Marshall FH (2015) From G-protein-coupled receptor structure resolution to rational drug design. J Biol Chem 290:19489–19495.

14. Pardo L, Deupi X, Dölker N, López-Rodríguez ML, Campillo M (2007) The role of in-conserved residues in the activation of the inactive state of rhodopsin. Proc Natl Acad Sci USA 104:15580–15584.

15. Angel TE, Chance MR, Palczewski K (2009) Conserved waters mediate structural and functional landscape of GPCRs. In particular, residues interacting with the conserved, stable waters identified in this study are excellent candidate sites for introducing mutations to shift the conformational ensemble toward the active or inactive state.