Buried ionizable networks are an ancient hallmark of G protein-coupled receptor activation

Daniel G. Isom[a,1] and Henrik G. Dohlman[a,b,1]

Departments of [a]Biochemistry and Biophysics and [b]Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

Edited by Robert J. Lefkowitz, Howard Hughes Medical Institute, Duke University Medical Center, Durham, NC, and approved March 31, 2015 (received for review September 16, 2014)

Seven-transmembrane receptors (7TMRs) have evolved in prokaryotes and eukaryotes over hundreds of millions of years. Comparative structural analysis suggests that these receptors may share a remote evolutionary origin, despite their lack of sequence similarity. Here we used structure-based computations to compare 221 7TMRs from all domains of life. Unexpectedly, we discovered that these receptors contain spatially conserved networks of buried ionizable groups. In microbial 7TMRs these networks are used to pump ions across the cell membrane in response to light. In animal 7TMRs, which include light- and ligand-activated G protein-coupled receptors (GPCRs), homologous networks were found to be characteristic of activated receptor conformations. These networks are likely relevant to receptor function because they connect the ligand-binding pocket of the receptor to the nucleotide-binding pocket of the G protein. We propose that agonist and G protein binding facilitate the formation of these electrostatic networks and promote important structural rearrangements such as the displacement of transmembrane helix-6. We anticipate that robust classification of activated GPCR structures will aid the identification of ligands that target activated GPCR structural states.

7-transmembrane receptor | G protein-coupled receptor | buried charge | structural bioinformatics | molecular evolution

Seven-transmembrane receptors (7TMRs) are present in all domains of life. In archaea and bacteria these receptors convert light energy into transmembrane ion gradients or intracellular signaling cascades. In humans 7TMRs comprise a family of more than 800 G protein-coupled receptors (GPCRs) that detect extracellular signals such as odorants, taste, light, hormones, and neurotransmitters. Although microbial and eukaryotic 7TMRs lack sequence similarity, their 3D folds share a degree of similarity that often is observed in remote evolutionary relationships identified by structure comparison algorithms (1).

Recent breakthroughs in membrane protein crystallography have advanced our understanding of GPCR function by providing models of unactivated and activated receptor conformations (2–8). However, structure-based approaches for systematically quantifying differences between GPCR activation states are lacking. Here we introduce a computational approach that correlates GPCR activation with networks of electrostatic interactions in the receptor core. We show that membrane-spanning networks of ionizable residues, which we find are a common feature of microbial 7TMRs, also represent a unique signature of GPCR activation that is likely to be essential to receptor function.

The molecular changes that accompany 7TMR activation were studied initially in bacteriorhodopsin and rhodopsin, the first microbial 7TMR and GPCR to be crystallized (9–11). More recently, a rapid expansion of GPCR structural information has revealed several molecular switches and structural features that are thought to be indicative of receptor activation. These include the ionic lock, 3–7 lock, tyrosine toggle, nPxxy motif, transmission switch, and the displacement of transmembrane helix-6 (TM6) (12). However, these features are not present in all GPCRs and therefore may not provide a comprehensive description of the cooperative structural changes that coincide with receptor activation.

It is unclear whether microbial 7TMRs and GPCRs share a common genetic origin. However, comparative structure alignment algorithms (1), which can reveal remote evolutionary relationships in the absence of sequence similarity, suggest that all 7TMRs are structurally homologous. Here we show that the structural similarity between microbial 7TMRs and GPCRs extends to structural signatures of receptor activation. These include the presence of membrane-spanning ionizable networks, the displacement of TM6 from transmembrane helix 3 (TM3), and the summed displacement angle of TM6 from its surrounding transmembrane helices. We believe these findings support a remote evolutionary relationship between microbial 7TMRs and GPCRs and provide structural metrics that will complement existing approaches for discriminating activated GPCR conformations.

Results and Discussion

The amino acid side chains of Asp, Glu, His, Cys, Lys, and Arg usually are charged (Fig. 1*A*), and therefore are thought to be incompatible with hydrophobic protein cores. However, buried charged (ionizable) residues do exist, and they often are required for protein function (13, 14). In such cases, the pK_a values of buried residues are tuned by their microenvironment and often exhibit values much different from their pK_a values in solution (Fig. 1*A*) (15–17). When they ionize, buried residues can influence protein structure profoundly. For example, the burial of a Lys residue can depress its pK_a from 10.4 to 7 (16) and cause dramatic pH-dependent changes in protein conformation (18). This pK_a shift results in a protein that is responsive to changes in cellular pH that occur under physiological conditions such as nutrient deprivation, cancer, ischemia, and inflammation. Such electrostatic switches also support a wide range of cellular and physiological

Significance

In the early 1980s, scientists began searching for cell-surface receptors that bind to hormones and neurotransmitters. Among the first was the β-adrenergic receptor, a G protein-coupled receptor (GPCR) that is activated by norepinephrine and epinephrine. Recent breakthroughs have provided more than 100 new GPCR structures, including several in activated conformations. This new structural information presents an opportunity to identify features that distinguish unactivated and activated receptors. Here we use a computational approach to identify structural signatures unique to activated GPCRs. Remarkably, we find that these signatures also are present in distantly related receptors from archaea and bacteria. We propose that these new structural indicators are central to GPCR function and are indicative of GPCR activation.

Author contributions: D.G.I. designed research; D.G.I. performed research; D.G.I. contributed new reagents/analytic tools; D.G.I. created the pHinder algorithm and computer code; D.G.I. analyzed data; and D.G.I. and H.G.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: disom@unc.edu or hdohlman@med.unc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1417888112/-/DCSupplemental.

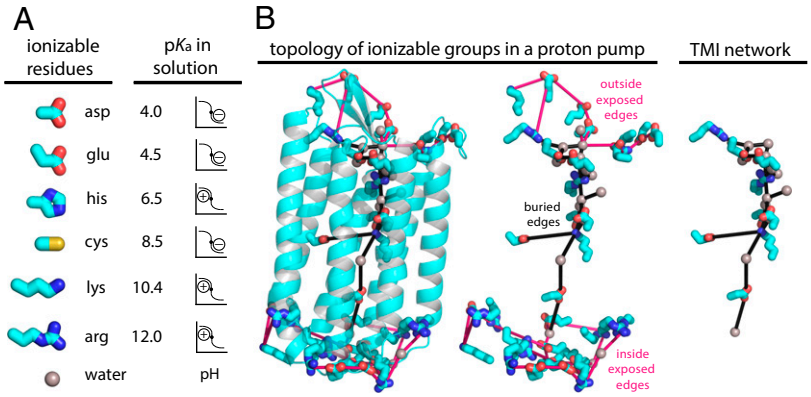

Fig. 1. pHinder results for a representative 7TMR with a TMI network. (*A*) The nodes of a pHinder network consist of ionizable amino acid side chains and water molecules. Solution pK_a values of the ionizable residues, which may shift dramatically when in the protein interior, are listed for reference. Formal charges listed in the simulated titration curves indicate whether the side chain is ionized above (Asp, Glu, Cys) or below (His, Lys, Arg) its usual pK_a value. (*B*) The topological network (edges) of ionizable residues (nodes) calculated for the microbial proton pump archaerhodopsin-1 (AR1). This example shows a network of ionizable groups buried within AR1 (black edges) that connects networks of ionizable groups located inside and outside the cell membrane (pink edges). The displayed structure corresponds to PDB ID code 1VGO.

functions, including oxygen delivery by hemoglobin (13), pH-gated ion transport (19, 20), and signaling by pH-sensing (proton-activated) GPCRs (21) and G proteins (22).

Although pK_a values of buried residues have been measured in model proteins (15–17), this measurement is not currently feasible in more complex systems. However, the relationship between ionizable groups and protein function can be studied computationally using physics-based models of protein electrostatics (23). To accelerate these efforts, we have developed an analytics-based approach, known as "pHinder" (22), for identifying functionally important ionizable interactions in proteins. pHinder differs from other structure-based electrostatics calculations by considering only the spatial organization (i.e., topology) of ionizable groups in a protein. The pHinder model is based on the rationale that topological clusters of acidic residues, basic residues, or networks of ionizable groups buried in the protein core represent electrostatic interactions that can regulate protein function dynamically.

Previously we have used pHinder to identify networks of electrostatic interactions buried in G protein α (Gα) subunits, the principal transducers of GPCR signals (22). Upon GPCR activation, the Gα subunit undergoes substantial structural changes as it exchanges GTP for GDP. The buried network traces a path between the Ras and helical domains and is exposed upon receptor activation (5). Therefore we wondered whether GPCRs also contain buried electrostatic networks that might regulate G protein activation. Indeed, several studies have highlighted the prevalence of ionizable residues, water molecules (24), and ions such as sodium (25, 26) in the cores of GPCRs. Here we used pHinder to analyze a comprehensive list of 27 rhodopsin and 70 ligand-binding GPCR structures. For comparison, we also analyzed structures of 124 microbial opsins, which share the 7TMR architecture but do not couple to G proteins.

The key features of a pHinder calculation are illustrated in Fig. 1. This example shows a transmembrane network of buried ionizable groups that passes through the core of a representative 7TMR. The topological relationships that define this network were calculated by triangulating the ionizable residues and water molecules of the receptor and by removing edges longer than 10.0 Å (Fig. S1), a distance constraint used to model medium-range electrostatic interactions in proteins (27). This procedure generates an interconnected network describing the local and global topology of the ionizable groups. To identify networks of potentially charged residues buried within the protein core, a molecular surface (Fig. S1) is used to classify each residue within the network as buried (>3.0 Å below the surface), margin (<3.0 Å below and <1.0 Å above the surface), or exposed (>1.0 Å above the surface).

Our pHinder calculations on 221 receptor structures revealed recurring network fragments consisting of buried ionizable side chains and water molecules. Using an approach we have termed "consensus network analysis" (CNA), we combined these recurring fragments by overlaying the results of our individual pHinder calculations within each set of microbial 7TMRs, rhodopsins, and ligand-binding GPCRs (Fig. 2). The CNA procedure revealed that each receptor subgroup contains ionizable residues buried in their cores (Fig. S2). These residues are spatially conserved despite the lack of sequence similarity between each 7TMR subgroup (28). We interpret this finding as evidence that spatially conserved buried ionizable groups were acquired early in evolutionary time as a universal and essential feature of 7TMR structure. However, because a common origin for microbial 7TMR and GPCR genes is unknown, we cannot rule out the possibility that these receptors evolved independently.

Buried ionizable networks within microbial 7TMRs (Fig. 2*A*), such as bacterio- and halorhodopsin, facilitate light-driven proton and ion transport across the cell membrane. These residues exhibit a range of unusual pK_a values that enable them to serve as electrostatic switches essential for light-driven proton pumping (29–31). Moreover, it is known that the ionization state of these residues can be regulated by changes in external pH (30, 32). pH-sensing residues in bacteriorhodopsin trace a path through the core of the receptor that originates at the extracellular proton acceptor Asp85, continues through the buried retinal-linked Lys216 (Schiff base), and extends to the cytoplasmic proton donor Asp96 (Fig. 2*A*). The presence of similar ionizable groups buried within rhodopsin and ligand-binding GPCRs (Fig. 2 *B* and *C*) suggests that protonation also may regulate GPCR structure and function. Although speculative, this idea is consistent with evidence that light-activated (3, 33–35) and ligand-activated (36–38) GPCR complexes are regulated directly by pH.

As shown in Fig. 2 *B* and *C*, the majority of buried ionizable residues are present in TM3, widely regarded as the signaling hub of GPCRs (2). In rhodopsins, these include a Lys residue (Lys296) that is structurally and functionally equivalent to a Lys residue in microbial 7TMRs (Lys216). These lysines are covalently linked to the retinal chromophore. As with microbial 7TMRs, consensus clusters of buried water molecules are found in rhodopsin and ligand-binding GPCRs, consistent with the recent finding that family A GPCRs contain structurally conserved water molecules that mediate structural and functional activation (24). Relative to microbial 7TMRs, rhodopsin and ligand-binding GPCRs contain additional clusters of buried ionizable groups that correspond to the D/ERY (Asp/Glu-Arg-Tyr) motif, a conserved feature that couples GPCRs to G proteins.

Having identified spatially conserved ionizable residues in all three receptor subgroups, we next searched for individual receptor structures that contain complete transmembrane ionizable (TMI)

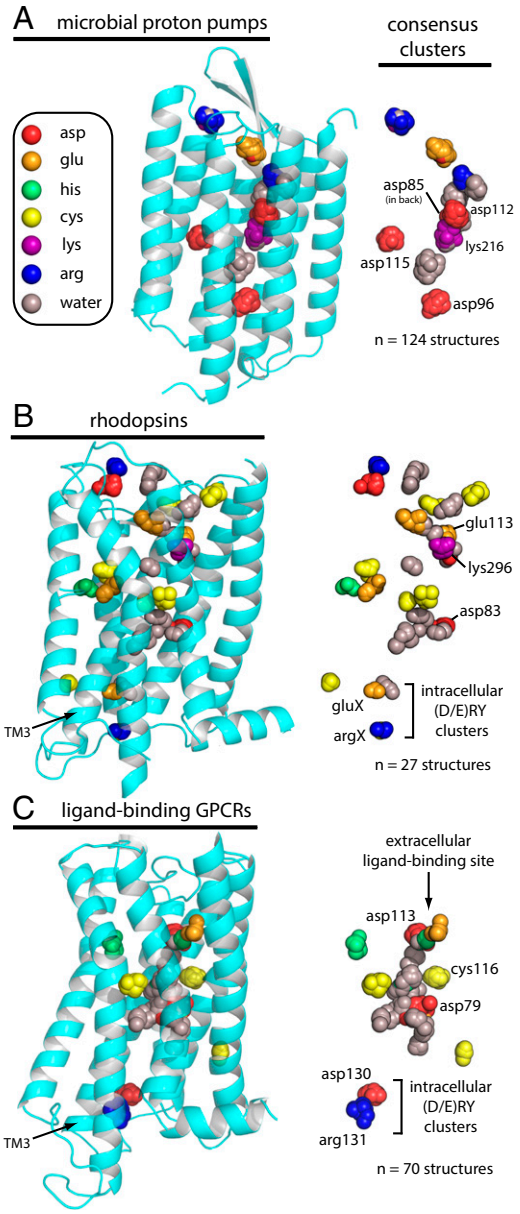

Fig. 2. Buried ionizable groups are conserved in microbial 7TMRs and animal GPCRs. Consensus network analysis of 124 microbial 7TMR (*A*), 27 rhodopsin GPCR (*B*), and 70 ligand-binding GPCR (*C*) structures reveals spatially conserved clusters of buried ionizable residues and water molecules. Colored spheres represent the contribution of individual side chains and water molecules to each consensus cluster. Prominent node clusters are labeled according to the residues of bacteriorhodopsin (*A*), bovine rhodopsin (*B*), and the human β_2-adrenergic receptor (*C*). In *B* and *C* clusters labeled (D/E)RY correspond to the (Asp/Glu-Arg-Tyr) structural motif that is unique to rhodopsins and ligand-binding GPCRs. PDB ID codes for all 221 structures are available in Dataset S1. Displayed structures correspond to PDB ID codes 1M0K (*A*), 1F88 (*B*), and 2KS9 (*C*).

networks. As expected, most microbial 7TMRs (66%, Dataset S2) contained TMI networks (Fig. 1*B*), as is consistent with their function as light-driven proton and ion pumps. In microbial 7TMRs, these networks were not exclusive to activated or unactivated conformations, because they were identified in several photocycle intermediates. We then searched for TMI networks within the eukaryotic 7TMRs (Fig. 3 and Datasets S3 and S4). We were particularly interested in comparing structures of receptors in the unactivated and activated state, defined here as any receptor bound to a G protein or G protein fragment. For example, the two opsin structures compared in Fig. 3*A* are nearly identical but differ by the absence (39) or presence (3) of a G protein fragment. However, only the activated structure contains a TMI network. With the exception of squid rhodopsin, TMI networks were found only in structures of activated rhodopsins (Fig. S2) (4, 6, 40). In opsin and rhodopsin these networks trace a path that originates from extracellular residues above the retinal-binding pocket, passes through the buried Lys296 that couples to the retinal ligand, extends through structurally conserved waters in the receptor core, and terminates at the intracellular D/ERY motif (Fig. 3*A* and Fig. S2). Together, these findings demonstrate that the pHinder algorithm can discern between the activated and unactivated photoreceptor structures.

We then considered the ligand-binding GPCRs and again observed TMI networks within the activated receptor structure (Fig. 3*B*). When water molecules were included in the pHinder calculation, 7 of 70 ligand-binding GPCR structures contained TMI networks (Dataset S4). However, the low atomic resolution of many receptor structures precludes the modeling of buried water molecules. To circumvent this issue, we excluded water from the pHinder calculation and found that only 3 of 70 GPCRs contained TMI networks (Dataset S4). These three structures correspond to the β_2-adrenergic receptor (β_2AR) (5), the only GPCR to be crystalized in complex with a G protein (β_2AR-Gs), and two muscarinic acetylcholine receptors (M2Rs) in complex with G protein-mimetic nanobodies (Fig. 3 *B* and *C*) (8). In β_2AR and M2R, the TMI networks are likely relevant to receptor function, because they originate at the extracellular ligand-binding pocket, continue along TM3 to the intracellular D/ERY motif, and extend into the nucleotide-binding pocket of the stimulatory G protein α subunit (Gαs) (Fig. 3*B*). As expected, mutations within these buried networks alter receptor function (Fig. 3*B*). Examples in β_2AR include Asp79 (41, 42), Asp113 (42, 43), Cys116 (44), Asp130 (45), and Cys327 (46). Asp79 buried in the core of β_2AR is a particularly important site, because mutations of structurally equivalent Asp residues in other GPCRs (α_2-adrenergic, D_2 dopamine, and M1 muscarinic acetylcholine receptors) also impede receptor function (47). We interpret this finding as evidence that agonist and G protein binding stabilize activated receptor conformations, in part by organizing networks of electrostatic interactions that span the membrane.

Further analysis of the β_2AR–Gs complex shows that the appearance of TMI networks in the fully activated state of β_2AR (Fig. 3*B*) coincides with the loss of an internal proton-sensing network located at the interface of the Ras and helical domains of Gαs (Fig. 3*B*) (22). Likewise, the network buried within M2R is linked to a network of ionizable residues residing in its bound G protein-mimetic nanobody (Fig. 3*C*). Notably, portions of the TMI network that extend into the Gαs subunit and nanobody are conserved (Fig. 4*A*). These include a conserved His (Gαs-His387 and nanobody-His106) and acidic residue (Gαs-Glu392 and nanobody-Asp30) that appear necessary for stabilizing the activated receptor conformation. Thus, it appears that receptor activation causes the buried network within Gαs to be replaced by another buried network that links the Gαs with the ligand-binding pocket of β_2AR. Based on these observations, we conclude that the exchange of electrostatic networks within the receptor–Gα complex is a central feature of GPCR functional cycles.

As expected for homologous GPCRs, buried networks in β_2AR and M2R align well with each other (Fig. 4*A*). However, these networks also align remarkably well with those in microbial 7TMRs (Fig. 4*A*). This alignment is most visible in the algal proton pump *Acetabularia* rhodopsin (ARII). Unlike most other microbial 7TMRs, the TMI network in ARII does not require buried water molecules to span the cell membrane (Fig. 4*A*) (48). Compared with TMI networks in β_2AR and M2R, the network in ARII is longer but traces a remarkably similar path through the receptor (Fig. 4*A*). Given the lack of sequence similarity between the GPCRs and ARII (<10%) and the rarity of TMI networks in known membrane protein structures (Fig. 4*B*), the spatial conservation of these networks suggests the importance of internal

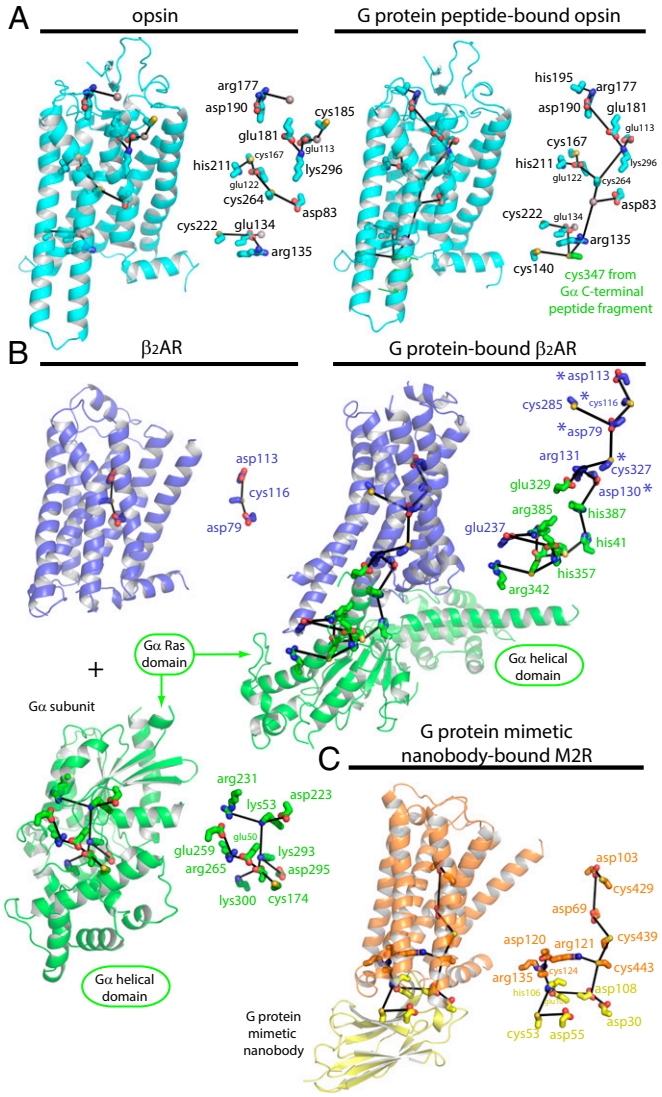

Fig. 3. Buried TMI networks are unique to GPCR structures activated by physiological stimuli. (*A* and *B*) A comparison of unactivated and activated opsin (*A*) and β_2AR (*B*) structures. Activated structures contain complete TMI networks. In the β_2AR-Gs complex, these networks extend into the Ras domain of the Gαs. Gray spheres in *A* indicate water molecules. In *B*, the purple asterisks within the TMI network of β_2AR indicate residues that, when mutated, alter receptor function. (*C*) A complete TMI network also was observed in the structure of M2R in complex with G protein-mimetic nanobody. In *B* and *C* the colors of the node label correspond to the protein chain that contributes the side chain. Displayed structures correspond to PDB ID codes 3CAP and 3DQB (opsin) (*A*); 1AZT (Gαs subunit), and 3D4S and 3SN6 (β_2AR) (*B*); and 4MQT (M2R) (*C*).

electrostatic interactions in the structure, function, and evolution of the 7TMR superfamily. However, we cannot rule out the possibility that these networks evolved independently in microbial 7TMRs and GPCRs, despite the apparent structural and functional similarities we have quantified.

In addition to TMI networks, the spatial conservation of buried charge in prokaryotic and eukaryotic 7TMRs is reflected in other structural similarities (Fig. 4 *C* and *D*). As shown in Fig. 4*C*, the distribution of buried network residues throughout transmembrane helices is similar in microbial 7TMRs, rhodopsin, and ligand-binding GPCRs. Most notably, TM3, widely considered the signaling hub of GPCRs, contains the largest proportion of TMI network residues. Other than in TM3, there are differences in the distribution of buried charge between the more distantly related microbial 7TMRs and ligand-binding GPCRs. However, based on the distribution of buried charged residues (Fig. 4*C*), we speculate that animal rhodopsin is positioned between microbial 7TMRs and ligand-binding GPCRs in evolutionary time.

Our calculations also revealed a remarkable similarity in transmembrane helix displacement angles among microbial 7TMRs, rhodopsin, and GPCRs. As shown in Fig. 4*D*, a 15° displacement angle is strongly conserved between TM6 and TM3 in microbial 7TMRs and GPCRs, even though microbial 7TMRs do not couple to G proteins. With the exception of TM4, the displacement angles between TM6 and the other TM helices also were found to be largely conserved (Fig. S3). To be comprehensive, we expanded our structure-based analyses to include 15,000 GPCR homology models acquired from the Protein Model Portal (49). These calculations reinforced our findings presented in Fig. 4 *C* and *D* and provided a broader estimate for the rarity of TMI networks in unactivated GPCR structural models (Fig. 4*B*). We believe that together these observations provide further structural evidence for a remote evolutionary relationship between microbial 7TMRs and animal GPCRs.

The ability to discriminate activated GPCR structures is important for understanding receptor structure and function and for identifying ligands that interact specifically with different states of receptor activation. In this work we have shown that TMI networks are an indicator of GPCR activation. However, other structural features, such as TM6 displacement and the ionic lock, are used widely to interpret GPCR structures. As shown in Fig. 4*D*, our calculations show that the displacement of TM6 correlates with activated rhodopsin and ligand-binding GPCR structures. Furthermore, the wide displacement angle of TM6 in the activated β_2AR–Gs complex is particularly striking. This structural difference suggests that the observed TM6 displacement angles in receptors activated by G protein peptides or in complex with G protein-mimetic nanobodies could be even larger. Ionic locks (i.e., conserved ionic interactions between an Arg in TM6 and Glu in TM3) are disrupted in activated rhodopsin structures. However, ionic locks also were disrupted in many unactivated rhodopsin structures. Based on these observations, we conclude that TM6 displacement and the presence of TMI networks serve as robust indicators of receptor activation.

At the dawn of microbial life, 7TMRs emerged as molecular machines for transducing sunlight into biologically useful energy. The process of evolution expanded the superfamily of 7TMRs, including members with little or no sequence similarity and with diverse cellular functions. Although the evolutionary history of 7TMRs is unclear, there now is compelling evidence for structural homology between microbial 7TMRs and animal GPCRs. Most significantly, we have shown that 7TMRs from all three domains of life contain networks of ionizable residues buried in their cores and that these networks are both structurally conserved and functionally important. In particular, these networks correlate with receptor–G protein assembly and therefore serve as an indicator of GPCR activation. Moving forward, we believe that this structural signature could be used to rank-order GPCR activation states within ensembles of receptor structures, whether they are determined experimentally or generated by molecular modeling. Last, our studies reveal the importance of buried charged networks in GPCRs. These networks are tunable, and their behavior can be modulated by dynamic changes in membrane potential and electrostatic interactions. These properties likely were acquired early in evolutionary time as a universal feature of the 7TMR superfamily. As new structures become available, we now have the opportunity to discern the structural basis for GPCR function.

Methods

Acquisition of 7TMR Structures. Microbial 7TMR and GPCR structures were downloaded from the Protein Data Bank (PDB). The query parameters used to collect type I microbial opsin structures were "TRANSMEMBRANE is ALPHA-HELICAL and TRANSMEMBRANE is Bacterial and Algal Rhodopsins." Query parameters used to collect type II eukaryotic opsin structures were "TRANSMEMBRANE is ALPHA-HELICAL and TRANSMEMBRANE is G Protein-Coupled Receptors (GPCRs)." The set of structures procured is detailed in Dataset S1.

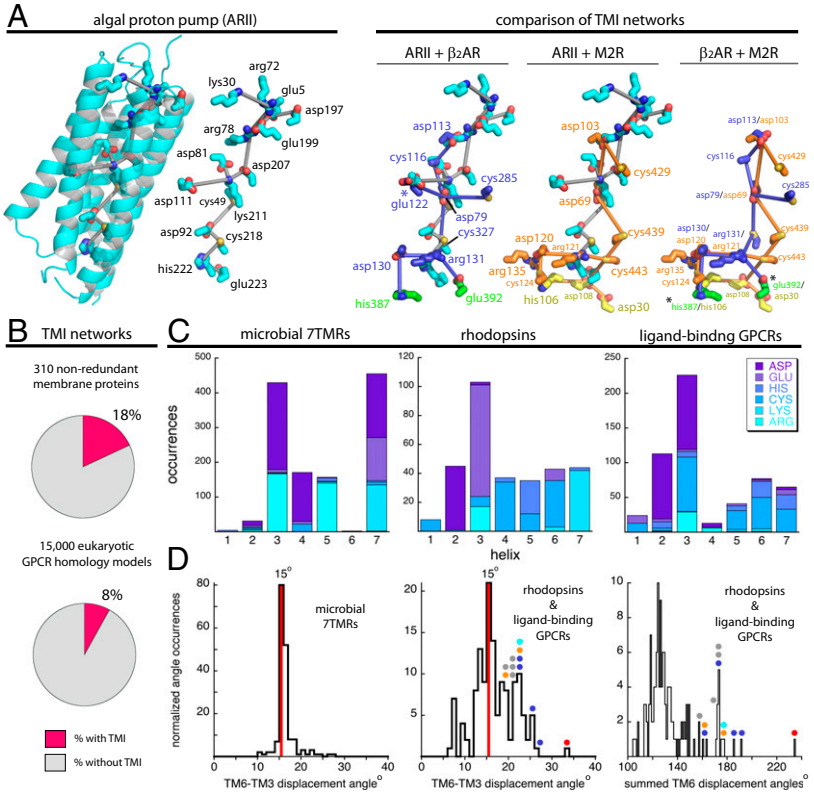

Fig. 4. Ancient TMI networks from microbial proton pumps are conserved in activated GPCR conformations. (*A*) The TMI network within the microbial proton pump from the marine algae *Acetabularia acetabulum* (ARII) (PDB ID code 3AM6) superimposed on to networks in activated β_2AR (ARII + β_2AR) and M2R (ARII + M2R). A comparison of homologous TMI networks in β_2AR and M2R (β_2AR + M2R) is shown for reference. Label colors for β_2AR and M2R are the same as in Fig. 3. The TMI networks for β_2AR and M2R have been truncated at the first nodes of the Gαs subunit (β_2AR) or G protein-mimetic nanobody (M2R). Black asterisks indicate spatially conserved TMI network nodes shared by the Gαs subunit and G protein-mimetic nanobody. The purple asterisk in the TMI network of β_2AR (ARII + β_2AR) indicates an additional node that appears when the pHinder network edge constraint is increased to 10.5 Å. (*B*) The percentage of 310 non–7TM-containing membrane proteins and 15,000 GPCR homology models that contained TMI networks. (*C*) The distribution of buried network ionizable residues throughout the TM helices of microbial 7TMR, rhodopsin, and ligand-binding GPCR structures. (*D*) Breakdown of the displacement angle between TM6 and TM3 and the summed displacement angles between TM6 and the six other TM helices. Colored circles indicate G protein-bound structures (red, PDB ID code 3SN6); nanobody-bound β_2AR structures (purple, PDB ID codes 4LDE, 4LDL, 4LDO, and 3P0G); nanobody-bound M2R structures (orange, PDB ID codes 4MQT and 4MQS); Gα peptide-bound opsin (cyan, PDB ID code 3DQB); and rhodopsin structures (gray, PDB ID codes 3PQR, 4A4M, 2X72, and 4BEY).

Alignment of 7TMR Structures. The complete set of 7TMR structures was aligned to rhodopsin (PDB ID code 1F88) chain A by a two-step process using customized Python code. First, the primary amino acid sequences of the target (i.e., 1F88), and query structures were aligned using the Smith–Waterman algorithm (50). This primer alignment was used to identify the maximum span of homologous sequence shared by the target and query structures. Second, these regions of homologous sequence were aligned structurally using the Gaussian-weighted rmsd algorithm (51). Structures that could not be aligned using this procedure are listed in Dataset S1.

pHinder Calculation. The pHinder algorithm was used to calculate the topology of ionizable groups in 3D structures of proteins using a two-step triangulation procedure. First, a Delaunay triangulation was calculated for terminal side-chain atoms of all ionizable residues (Asp, Glu, His, Cys, Lys, Arg) and water molecules (if included) of the protein. The triangulation then was minimized by removing edges longer than 10 Å (the default cutoff distance used in all pHinder calculations). In Fig. 4*A*, the cutoff distance was extended to 10.5 Å (relative to a cutoff distance of 10 Å in Fig. 3*B*) to capture an additional Asp side chain that is spatially conserved between the β_2AR and *Acetabularia* rhodopsin II. After removal of each long edge, the triangulation was simplified further by removing redundant network connections. Using a molecular surface, each network node was classified as buried (>3.0 Å below the surface), margin (<3.0 Å below and <1.0 Å above the surface), or exposed (>1.0 Å above the surface). Depth of burial was determined by measuring the minimum distance between the ionizable group and the triangular facets of the surface. Buried networks were identified as contiguous runs of buried nodes. Buried network edges that crossed the molecular surface were truncated at the first margin node encountered. Thus, by convention, the terminal nodes of buried networks may not be buried. Select aspects of the triangulation procedure and an example molecular surface are illustrated in Fig. S1.

CNA of 7TMRs. Consensus networks of internal ionizable side chains and water molecules (if included) were calculated for each 7TMR subgroup in a three-step procedure. First, the network nodes from the individual pHinder calculations were combined and clustered using a distance constraint (1.0 Å for microbial 7TMRs, 2.0 Å for rhodopsins, and 2.0 Å for ligand-binding GPCRs) and minimum cluster size (40 for microbial 7TMRs, 8 for rhodopsins, and 10 for ligand-binding GPCRs). Second, a Delaunay triangulation was calculated for the set of clustered nodes using a 10.0-Å distance constraint. Third, the triangulated cluster nodes were subjected to a second round of iterative clustering using a refined distance constraint (0.5 Å for microbial 7TMRs, 2.0 Å for rhodopsins, and 2.0 Å for ligand-binding GPCRs) and minimum cluster size (20 for microbial 7TMRs, 5 for rhodopsins, and 3 for ligand-binding GPCRs).

Comparative Structural Analysis of 7TMRs. The detailed approaches that comprise our comparative structural analysis of microbial 7TMRs, rhodopsins, and ligand-binding GPCRs are included in *SI Methods*.

1. Gibrat JF, Madej T, Bryant SH (1996) Surprising similarities in structure comparison. *Curr Opin Struct Biol* 6(3):377–385.

2. Venkatakrishnan AJ, et al. (2013) Molecular signatures of G-protein-coupled receptors. *Nature* 494(7436):185–194.

3. Scheerer P, et al. (2008) Crystal structure of opsin in its G-protein-interacting conformation. *Nature* 455(7212):497–502.
4. Standfuss J, et al. (2011) The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* 471(7340):656–660.
5. Rasmussen SG, et al. (2011) Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.
6. Choe HW, et al. (2011) Crystal structure of metarhodopsin II. *Nature* 471(7340):651–655.
7. Deupi X, et al. (2012) Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proc Natl Acad Sci USA* 109(1):119–124.
8. Kruse AC, et al. (2013) Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* 504(7478):101–106.
9. Grigorieff N, Ceska TA, Downing KH, Baldwin JM, Henderson R (1996) Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J Mol Biol* 259(3):393–421.
10. Pebay-Peyroula E, Rummel G, Rosenbusch JP, Landau EM (1997) X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* 277(5332):1676–1681.
11. Palczewski K, et al. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289(5480):739–745.
12. Trzaskowski B, et al. (2012) Action of molecular switches in GPCRs—theoretical and experimental studies. *Curr Med Chem* 19(8):1090–1109.
13. Perutz MF (1978) Electrostatic effects in proteins. *Science* 201(4362):1187–1191.
14. Perutz MF (1992) What are enzyme structures telling us? *Faraday Discuss* 93:1–11.
15. Isom DG, Cannon BR, Castañeda CA, Robinson A, García-Moreno B (2008) High tolerance for ionizable residues in the hydrophobic interior of proteins. *Proc Natl Acad Sci USA* 105(46):17784–17788.
16. Isom DG, Castañeda CA, Cannon BR, García-Moreno B (2011) Large shifts in pKa values of lysine residues buried inside a protein. *Proc Natl Acad Sci USA* 108(13):5260–5265.
17. Isom DG, Castañeda CA, Cannon BR, Velu PD, García-Moreno E B (2010) Charges in the hydrophobic interior of proteins. *Proc Natl Acad Sci USA* 107(37):16096–16100.
18. Chimenti MS, et al. (2012) Structural reorganization triggered by charging of Lys residues in the hydrophobic interior of a protein. *Structure* 20(6):1071–1085.
19. Zhou M, Morais-Cabral JH, Mann S, MacKinnon R (2001) Potassium channel receptor site for the inactivation gate and quaternary amine inhibitors. *Nature* 411(6838):657–661.
20. Chang Y, et al. (2014) Structural basis for a pH-sensitive calcium leak across membranes. *Science* 344(6188):1131–1135.
21. Ludwig MG, et al. (2003) Proton-sensing G-protein-coupled receptors. *Nature* 425(6953): 93–98.
22. Isom DG, et al. (2013) Protons as second messenger regulators of G protein signaling. *Mol Cell* 51(4):531–538.
23. Nielsen JE, Gunner MR, García-Moreno BE (2011) The pKa Cooperative: A collaborative effort to advance structure-based calculations of pKa values and electrostatic effects in proteins. *Proteins* 79(12):3249–3259.
24. Angel TE, Chance MR, Palczewski K (2009) Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proc Natl Acad Sci USA* 106(21):8555–8560.
25. Fenalti G, et al. (2014) Molecular control of δ-opioid receptor signalling. *Nature* 506(7487):191–196.
26. Katritch V, et al. (2014) Allosteric sodium in class A GPCR signaling. *Trends Biochem Sci* 39(5):233–244.
27. Lee KK, Fitch CA, Garcia-Moreno EB (2002) Distance dependence and salt sensitivity of pairwise, coulombic interactions in a protein. *Protein Sci* 11(5):1004–1016.
28. Strotmann R, et al. (2011) Evolution of GPCR: Change and continuity. *Mol Cell Endocrinol* 331(2):170–178.
29. Rousso I, Friedman N, Sheves M, Ottolenghi M (1995) pKa of the protonated Schiff base and aspartic 85 in the bacteriorhodopsin binding site is controlled by a specific geometry between the two residues. *Biochemistry* 34(37):12059–12065.
30. Morgan JE, et al. (2012) Structure changes upon deprotonation of the proton release group in the bacteriorhodopsin photocycle. *Biophys J* 103(3):444–452.
31. Száraz S, Oesterhelt D, Ormos P (1994) pH-induced structural changes in bacteriorhodopsin studied by Fourier transform infrared spectroscopy. *Biophys J* 67(4):1706–1712.
32. Lanyi JK (2006) Proton transfers in the bacteriorhodopsin photocycle. *Biochim Biophys Acta* 1757(8):1012–1018.
33. Arnis S, Fahmy K, Hofmann KP, Sakmar TP (1994) A conserved carboxylic acid group mediates light-dependent proton uptake and signaling by rhodopsin. *J Biol Chem* 269(39):23879–23881.
34. Mahalingam M, Martínez-Mayorga K, Brown MF, Vogel R (2008) Two protonation switches control rhodopsin activation in membranes. *Proc Natl Acad Sci USA* 105(46): 17795–17800.
35. Lüdeke S, Mahalingam M, Vogel R (2009) Rhodopsin activation switches in a native membrane environment. *Photochem Photobiol* 85(2):437–441.
36. Ghanouni P, et al. (2000) The effect of pH on beta(2) adrenoceptor function. Evidence for protonation-dependent activation. *J Biol Chem* 275(5):3121–3127.
37. Deupi X, Kobilka BK (2010) Energy landscapes as a tool to integrate GPCR structure, dynamics, and function. *Physiology (Bethesda)* 25(5):293–303.
38. Dror RO, et al. (2011) Activation mechanism of the β2-adrenergic receptor. *Proc Natl Acad Sci USA* 108(46):18684–18689.
39. Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP (2008) Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* 454(7201):183–187.
40. Singhal A, et al. (2013) Insights into congenital stationary night blindness based on the structure of G90D rhodopsin. *EMBO Rep* 14(6):520–526.
41. Barak LS, Ménard L, Ferguson SS, Colapietro AM, Caron MG (1995) The conserved seven-transmembrane sequence NP(X)2,3Y of the G-protein-coupled receptor superfamily regulates multiple properties of the beta 2-adrenergic receptor. *Biochemistry* 34(47):15407–15414.
42. Strader CD, et al. (1988) Conserved aspartic acid residues 79 and 113 of the beta-adrenergic receptor have different roles in receptor function. *J Biol Chem* 263(21):10267–10271.
43. Strader CD, et al. (1987) Identification of residues required for ligand binding to the beta-adrenergic receptor. *Proc Natl Acad Sci USA* 84(13):4384–4388.
44. Zuscik MJ, Porter JE, Gaivin R, Perez DM (1998) Identification of a conserved switch residue responsible for selective constitutive activation of the beta2-adrenergic receptor. *J Biol Chem* 273(6):3401–3407.
45. Rasmussen SG, et al. (1999) Mutation of a highly conserved aspartic acid in the beta2 adrenergic receptor: Constitutive activation, structural instability, and conformational rearrangement of transmembrane segment 6. *Mol Pharmacol* 56(1):175–184.
46. O'Dowd BF, et al. (1988) Site-directed mutagenesis of the cytoplasmic domains of the human beta 2-adrenergic receptor. Localization of regions involved in G protein-receptor coupling. *J Biol Chem* 263(31):15985–15992.
47. Savarese TM, Fraser CM (1992) In vitro mutagenesis and the search for structure-function relationships among G protein-coupled receptors. *Biochem J* 283(Pt 1): 1–19.
48. Wada T, et al. (2011) Crystal structure of the eukaryotic light-driven proton-pumping rhodopsin, Acetabularia rhodopsin II, from marine alga. *J Mol Biol* 411(5):986–998.
49. Haas J, et al. (2013) The Protein Model Portal–a comprehensive resource for protein structure and model information. *Database* 2013:bat031.
50. Smith TF, Waterman MS (1981) Identification of common molecular subsequences. *J Mol Biol* 147(1):195–197.
51. Damm KL, Carlson HA (2006) Gaussian-weighted RMSD superposition of proteins: A structural comparison for flexible proteins and predicted protein structures. *Biophys J* 90(12):4558–4573.