Computational and Structural Biotechnology Journal 21 (2023) 4370–4384

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research article

Protons taken hostage: Dynamic H-bond networks of the pH-sensing GPR68

Bhav Kapur [a,b], Filippo Baldessari [c], Michalis Lazaratos [d], Herbert Nar [a], Gisela Schnapp [a], Alejandro Giorgetti [c,e], Ana-Nicoleta Bondar [e,f,*]

[a] Boehringer-Ingelheim Pharma GmbH & Co. KG, Birkendorfer Straße 65, 88397 Biberach an der Riß, Germany
[b] Christian-Albrechts-University of Kiel, 24118 Kiel, Germany
[c] University of Verona, Department of Biotechnology, 37134 Verona, Italy
[d] Department of Physics, Theoretical Molecular Biophysics Group, Freie Universität Berlin, Arnimallee 14, D-14195 Berlin, Germany
[e] Forschungszentrum Jülich, Institute for Neuroscience and Medicine and Institute for Advanced Simulations (IAS-5/INM-9), Computational Biomedicine, Wilhelm-Johnen Straße, 52525 Jülich, Germany
[f] University of Bucharest, Faculty of Physics, Str. Atomiştilor 405, 077125 Bucharest-Măgurele, Romania

ARTICLE INFO

Keywords:
Proton-sensing G Protein Coupled Receptors
Protocol for structural modeling and model assessment
Graph theory
Dynamic hydrogen-bond networks
Protonation-coupled protein dynamics

ABSTRACT

Proton-sensing G Protein Coupled Receptors (GPCRs) sense changes in the extracellular pH to effect cell signaling for cellular homeostasis. They tend to be overexpressed in solid tumors associated with acidic extracellular pH, and are of direct interest as drug targets. How proton-sensing GPCRs sense extracellular acidification and activate upon protonation change is important to understand, because it may guide the design of therapeutics. Lack of publicly available experimental structures make it challenging to discriminate between conflicting mechanisms proposed for proton-binding, as main roles have been assigned to either an extracellular histidine cluster or to an internal carboxylic triad. Here we present a protocol to derive and evaluate structural models of the proton-sensing GPR68. This approach integrates state-of-the-art homology modeling with microsecond-timescale atomistic simulations, and with a detailed assessment of the compatibility of the structural models with known structural features of class A GPCRs. To decipher structural elements of potential interest for protonation-coupled conformational changes of GPR68, we used the best-compatible model as a starting point for independent atomistic simulations of GPR68 with different protonation states, and graph computations to characterize the response of GPR68 to changes in protonation. We found that GPR68 hosts an extended hydrogen-bond network that inter-connects the extracellular histidine cluster to the internal carboxylic triad, and which can even reach groups at the cytoplasmic G-protein binding site. Taken together, results suggest that GPR68 relies on dynamic, hydrogen-bond networks to inter-connect extracellular and internal proton-binding sites, and to elicit conformational changes at the cytoplasmic G-protein binding site.

1. Introduction

G-Protein Coupled Receptors (GPCRs) form the largest superfamily of the human genome, and about 35% of the approved drugs in the market target GPCRs [1]. GPCRs are integral membrane proteins localized in cell membranes where they sense external messengers such as light, neurotransmitters, drugs, or protons, triggering downstream signaling pathways that elicit a specific cellular response [2,3]. Experimental and computational studies have deciphered key common principles of GPCRs mechanisms of action [4–8] and enabled progress in drug discovery [9–11]. Among GPCRs, a small sub-set function as pH sensors in various cell types [12]. These are the ovarian cancer receptor OGR1, or GPR68 [13], the T-cell death-associated gene 8 (TDAG8), or GPR65, GPR4 [14,15], GPR31 and GPR151 [16–18] and, potentially, the oncogenic G2A receptor, or GPR132 [19]. Proteins can sense pH via protonation changes of titratable sidechains, often histidine (His) and carboxylic sidechains (Asp, Glu) that titrate at physiological pH range [20,21] and, such groups are at the center of two distinct scenarios proposed to explain how pH-sensing GPCRs might work: Early studies assigned to a cluster of extracellular His sidechains a primary role as pH sensors [13], whereas more recent data were interpreted to suggest that pH sensing is accomplished primarily with a triad of carboxylic groups

* Corresponding author at: Forschungszentrum Jülich, Institute for Neuroscience and Medicine and Institute for Advanced Simulations (IAS-5/INM-9), Computational Biomedicine, Wilhelm-Johnen Straße, 52525 Jülich, Germany.
E-mail addresses: nbondar@fizica.unibuc.ro, a.bondar@fz-juelich.de (A.-N. Bondar).

https://doi.org/10.1016/j.csbj.2023.08.034
Received 5 May 2023; Received in revised form 30 August 2023; Accepted 30 August 2023
Available online 2 September 2023

[22]. A fundamental open question is whether and how the external and internal titratable sidechains might communicate with each other along the activation path of a proton-sensing GPCR. To address this question, here we studied the protonation-coupled dynamics of GPR68 with a computational approach that integrates structural modeling, atomistic simulations, and graph-based algorithms.

GPR68, which is one of the best characterized pH-sensing GPCRs, is active at pH ~6.8 and inactive at pH ~7.8 [13]. At its extracellular side GPR68 has a His cluster (Figs. 1, 2) whose role in pH sensing has been investigated via site-directed mutagenesis [13]. Single GPR68 mutants in which H17$^{1.28}$, H20$^{1.31}$, H84$^{2.67}$, or H269$^{7.36}$, were each mutated to Phe, exhibited reduced pH sensitivity of receptor activation at pH ~6.8, and the H245F mutant, significantly decreased pH sensitivity [13]. (Note that the upper scripts indicate the Ballesteros Weinstein, BW, numbers for class A GPCRs [23]). The reduced proton sensitivities of the H20F, H245F and H269F mutants were confirmed by more recent experimental data [24,25]. Intriguingly, the impact on proton sensing was smaller when H245$^{6.52}$ and H269$^{7.36}$ were mutated to Ala than to Phe, which could be interpreted to suggest that these two His sidechains might have structural roles [25]. The double mutants H17F/H84F and H20F/H269F had activity profiles similar to the corresponding single mutants, and even the quadruple mutant, with all four His mutated to Phe, still retained some function. When five [13] or seven extracellular His sidechains [25] were mutated to Phe some activity remained, though it was significantly impaired.

That mutation of selected extracellular His sidechains to Phe decreases proton sensitivity was also observed for TDAG8, whose H243F mutant (H245$^{6.52}$ in OGR1) has reduced proton sensing [26], and for G2A (also known as GPR132), whose H174F mutant (H159$^{4.63}$ in GPR68) had a somewhat reduced pH-dependent activation [27]. But G2A is significantly less pH-sensitive than its homologues [28] and, although single or double His-Phe mutations in GPR4 left the maximal receptor activity largely unchanged, the response of the receptor towards was shifted more acidic values - suggesting that carboxylic groups might be involved [29].

Carboxylic groups of direct interest for proton binding to GPR68 are D67$^{2.50}$, E149$^{4.53}$, and D282$^{7.49}$, which were proposed to establish the so-called carboxylic triad that senses pH in GPCRs [22] (Figs. 1, 2). Proton binding and protonation change at the conserved D2.50 (Table S1) were also proposed before for class A GPCRs such as bovine rhodopsin [30,31], the M2 muscarinic receptor D2.50 [32], and the β1 adrenergic receptor –which favors an inactive conformation when D2.50 is neutral [33]. For GPR68, GPR4 and GPR65, site-directed mutagenesis and pH activation suggested that D2.50 is likely negatively charged at alkaline and intermediate pH values, and neutral at an acidic pH where the receptor is fully active [22]. The same protonation states were suggested for E149$^{4.53}$, whereas D282$^{7.49}$ is thought to be neutral at both intermediate and acidic pH values [22].

The second carboxylic group of the internal triad, E149$^{4.53}$, is conserved as Glu in three of the pH-sensing GPCRs (Table S1); moreover, the finding that mutations at position 4.53 are detrimental to the receptor [40,41] suggests that residue 4.53 could be important for protein folding and conformational dynamics [40]. The third carboxylic group of the triad, D282$^{7.49}$, is part of the DPxxY functional motif that rearranges upon receptor activation [42]. Moreover, in high-resolution experimental structures of class A GPCRs, amino acid residues 2.50 and 7.49 are also part of a common, conserved protein-water hydrogen (H)-bond network with several other groups essential for function [43]. Rearrangements of the conserved motifs are accompanied by an outward movement of the cytoplasmic end of transmembrane (TM) helix 6, and of TM7 towards TM3 [8].

How the internal carboxylic triad might collect protons from the extracellular side of GPR68 and other pH-sensing GPCRs, and how changes in the protonation state of the internal carboxylic triad elicit protein conformational changes that ultimately lead to receptor activation, remains largely unknown. For the internal carboxylic triad to bind protons from the extracellular bulk, or to sense protonation change at an extracellular site, long-distance couplings between the internal carboxylic triad and the extracellular side would be required. Our hypothesis is that dynamic water-mediated H-bond networks are central to such long-distance coupling. This hypothesis is compatible with the key role of internal water molecules and water-mediated H-bond networks in proton transporter function [44–47] and GPCR activation [48–52].

To find out whether protein-water H-bond networks might connect the internal carboxylic triad to the extracellular side of a pH-sensing GPCR, here we study the dynamics of GPR68 for different protonation states. Our computational approach integrates structural modeling of GPR68, 18 atomistic simulations with a total sampling time of 14 μs, a protocol to evaluate GPR68 structural models, and graph-based computations of dynamic protein-water H-bond networks. We find that extracellular His sidechains of GPR68 are part of an extended water-mediated H-bond network with multiple carboxylic groups. This water-mediated histidine-carboxylic cluster connects to the internal carboxylic triad via protein-water H-bond networks whose dynamics depend on the protonation state of GPR68.

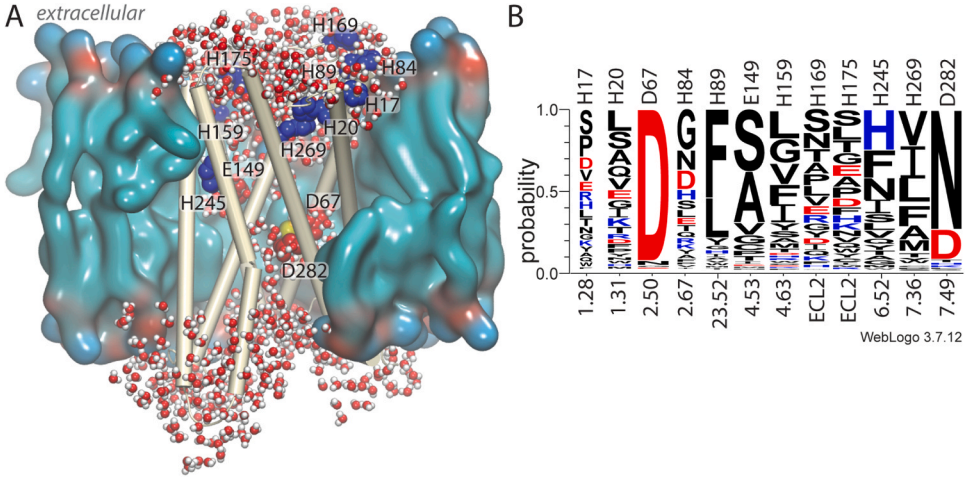

Fig. 1. Location of selected titratable sidechains of GPR68. (A) Cut-away view of membrane-embedded GPR68, based on a simulation presented here. Lipid and water molecules shown are within 10 Å and, respectively, 3 Å of the protein. His, Asp, and Glu sidechains shown are thought important for proton sensing. (B) Conservation of selected His, Asp, and Glu sidechains in human class A GPCRs. The height of each amino acid residue symbol indicates its conservation among the class A GPCR sequences aligned. The corresponding amino acid residue in GPR68 is indicated at the top of the panel; when existing, the corresponding number in the BW scheme is indicated at the bottom of the panel. The thinner one-letter symbols for positions 1.28, 1.31 and 2.67 indicate gaps in some of the aligned GPCR sequences. More detailed conservation analyses for pH-sensing GPCRs are presented in Fig. S1 and Tables S1, S2. The sequence alignment was performed using GPCRdb and is deposited in the Mendeley repository indicated in the Supporting Information that accompanies this work.

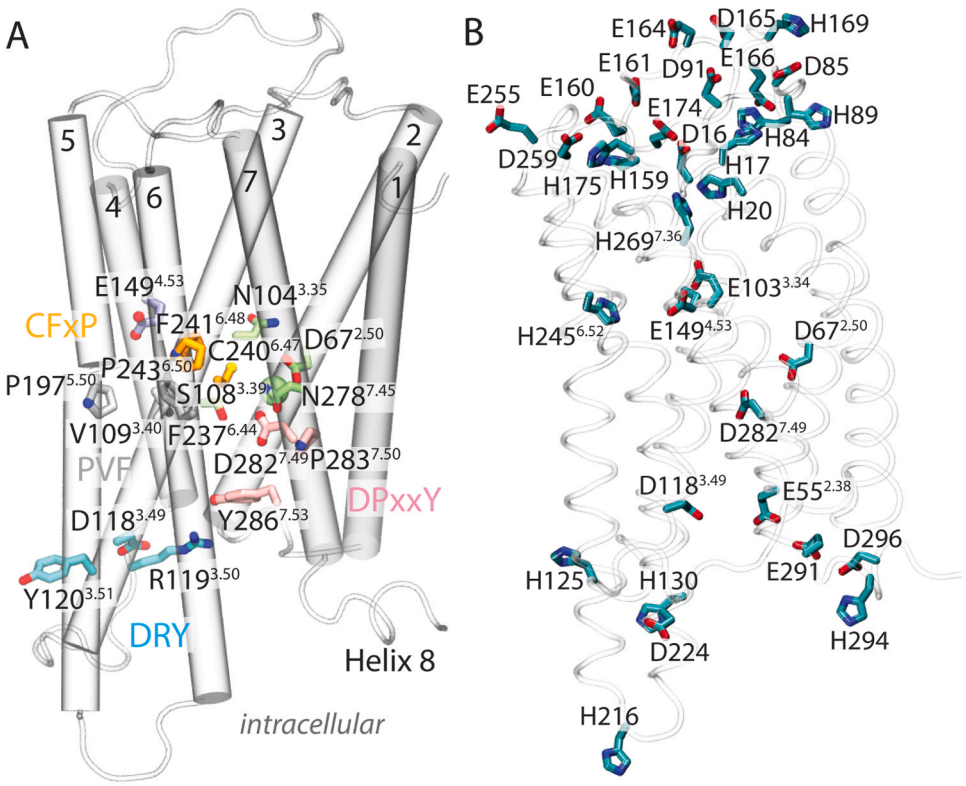

Fig. 2. Location of the conserved functional motifs and of the putative proton-binding sites of GPR68. (A) Amino acid residue sidechains of the conserved motifs of class A GPCRs are shown as bonds colored pink (DPxxY motif, D282$^{7.49}$, P283$^{7.50}$, and Y286$^{7.53}$), cyan, (DRY motif R119$^{3.50}$, D118$^{3.49}$, Y120$^{3.51}$), orange (CFxP transmission switch, C240$^{6.47}$, F241$^{6.48}$, and P243$^{6.50}$), and gray (PVF motif, P197$^{5.50}$, V109$^{3.40}$, and F237$^{6.44}$). DRY is involved in the binding of the cytoplasmic G protein [34,35], CFxP, in the movement of TM5 and TM6 during receptor activation [36,37], PVF, in structural rearrangement of F$^{6.44}$ and I$^{3.40}$ [38]. A sodium ion is a negative allosteric modulator [8,39]. (B) Location of the His, Asp, and Glu groups in GPR68. Note the cluster of carboxylic and His groups at the extracellular side; fewer such groups are in the core and intracellular regions of GPR68. The conservation of amino acid residues in pH-sensing GPCRs is presented in Table S1.

2. Methods

Homology modeling of GPR68 using GOMoDo. The webserver GPCR Online Modeling and Docking, GOMoDo [53], and its newly released version pyGOMoDo [54] have been validated for the state-of-the-art modeling of GPCRs [53]. We used these tools to select templates for GPR68 homology modeling and to perform multiple sequence alignments (MSAs) with the search-and-align algorithm HHblits (profile hidden Markov Models (HMM), HMM-HMM-based lightning-fast iterative sequence search [55–57]). HHblits converts the MSA into an HMM, searches the HMM database, and relies on an expected (*E*) value threshold to query the MSA [57]. To select templates we used as criteria an $E < 10^{-3}$, the higher sequence identity between template and target, and structural conservation of the corresponding BW positions in the protein sequences [23]. An additional comparison was made with GPCRdb [58] to extract, from the GOMoDo hits, those with higher conservation in the region of the extracellular His cluster of GPR68. Templates selected for the GOMoDo homology modeling are presented in Table S3.

Modeling of disulfide bridges and of the N-terminus. There is some uncertainty about the disulfide bridges in GPR68. Class A GPCRs typically have a conserved disulfide bridge between C94$^{3.25}$ and C172$^{45.50}$ [59]. In the case of GPR68, an additional disulfide bond might connect C13$^{1.24}$ to C258^{ECL3} [25]. Given this uncertainty, for each template used in the GOMoDo homology modeling above we performed two separate MODELLER calculations [60,61] distinguished by the restraints imposed on disulfide bridges: one calculation with a C94$^{3.25}$-C172$^{45.50}$ bridge and, separately, one calculation with C94$^{3.25}$-C172$^{45.50}$ and C13$^{1.24}$-C258^{ECL3} disulfide bridges. For each set of restraints on the disulfide bridges we generated 1000 GPR68 models, for a total of 6000 models, which we then ranked according to their calculated DOPE score [62]. For each set of structures generated from a given template, we visually inspected the ten highest-scoring models, discarded models with unstructured regions or other structure artifacts, and kept the structure that had both the highest DOPE score and lacked structure artifacts. Thus, we kept in total six models, one for each of the templates used (Table 1).

As all structural templates used lacked coordinates for the N-terminal segment, we used the loop modeling function of MODELLER to refine the coordinates of the N-terminal 18 amino acid residues of each GPR68 model.

AlphaFold AI-based model and C-terminal palmitoylation of GPR68. The AlphaFold [63] model of GPR68 (as available on alphafold.ebi.ac.uk) is based on the multiple templates listed in Table S3. AlphaFold uses a per-residue confidence metric denoted as predicted local distance difference test, pLDDT, whereby a score of 100 indicates the highest confidence in the accuracy of the model, and scores between 70 and 90 show a confident model. The overall pLDDT score of the AlphaFold

Table 1
Summary of structural models we prepared. We indicate the sequence identity between template and target, when present, the disulfide bridges, and the sequence range of the homology model. The Protein Data Bank (PDB, [66]) codes for the structures used as template are listed in Table S3.

Model#	Proteins used as template (s)	Sequence Identity (%)	Disulfide bridges	Sequence range
Homology models prepared with GOMoDo				
1	Protease	17	C94-C172	1–299
2	Activated Receptor 2 (5NJ6 - Inactive)		C94-C172, C13-C258	
3	Angiotensin II	20	C94-C172	12–299
4	Type 1 Receptor (4ZUD - Inactive)		C94-C172, C13-C258	1–299
5	Angiotensin II		C94-C172	12–290
6	Type 1 Receptor (6DO1- Active)		C94-C172, C13-C258	1–290
AlphaFold AI-based model				
7	Multiple templates	-	C94-C172, C13-C258	1–317
8		-	C94-C172, C13-C258	1–317 C304, C310 -palmitoylated

GPR68 model is 79.69, indicating the model is of reasonable overall quality. Termini and loop regions, which are typically flexible, have lower confidence scores.

A difference between the AlphaFold GPR68 model and the GOMoDo models is that the former predicts a longer amphipathic helix 8 at the C terminus. This is intriguing, because the C-terminus is palmitoylated at C322 and C323 in bovine rhodopsin [64], and at C341 (analogous to C322 in bovine rhodopsin) in the β_2-adrenergic receptor [65]. GPR68 lacks a Cys at the position corresponding to C322, but it has two C-terminal Cys groups at positions 304 and 310 in the sequence. To test the impact of palmitoylation, we performed two independent simulations starting from the AlphaFold AI-based model: one simulation without palmitoylation of the C-terminus, and one simulation with palmitoylated C304 and C310.

Homology modeling of GPR68 using bovine rhodopsin as a template. Earlier site-directed mutagenesis data on GPR68 have been interpreted using a homology model based on the crystal structure of bovine rhodopsin from ref. [67]. To directly compare our GPR68 homology models with this previous model, we used MODELLER v10.1 [61] and the bovine rhodopsin structure PDB ID: 1F88 chain A [67] and generated 100 models of GPR68 with a disulfide bridge between C94$^{3.25}$ and C172$^{45.50}$ (bovine rhodopsin C110 and C187). For each of these 100 initial models, we used the loop refinement tool of MODELLER to generate 3 models, and then ranked the resulting 300 models according to the DOPE and DOPE-HR scores. We inspected visually the 15 best-ranked models and their Ramachandran plots, and the distance between the Sγ atoms of C94$^{3.25}$-C172$^{45.50}$. To prepare a model of bovine rhodopsin based on PDB ID: 1F88 chain A [67], use used CHARMM-GUI [68] and Galaxy [69] to disulfide-bridge C110$^{3.25}$ and C185$^{45.50}$ and to generate coordinates for the missing amino acid residues 236–239 and 328–333. This bovine-rhodopsin-based GPR68 model was not used for molecular dynamics (MD) simulations.

The choice of 15 bovine-rhodopsin based models for visual inspection, as compared to 10 models in the GOMoDo-based protocol above, was made to verify more of the structural models generated with a relatively poorer template; however, using the 10 or the 15 best-ranked bovine rhodopsin-based models for visual inspection leads to the same selection of a rhodopsin-based GPR68 for further consideration.

Sequence analyses of pH-sensing GPCRs. We used GPCRdb [70], Jalview2 [71,72], and WebLogo V.3.7.12 [73,74] to align the sequences of GPR68 downloaded from UniProt [75] (UniProt ID Q15743), GPR4 (P46093), GPR65 (Q8IYL9), GPR132 (Q9UNW8), GPR31 (O00270), and GPR151 (Q8TDV0) (Fig. 1B, Fig. S1); we used the same software packages to generate the logo of the class A GPCR sequence conservation for specific positions of interest (Fig. 1B).

Protonation states. For all simulations performed to verify the structural stability of the homology models of GPR68 (Simulations #1 to #8 in Table 2) we used standard protonation states for all titratable sidechains, i.e., Arg and Lys, positively charged, Asp and Glu, negatively charged, and His groups singly protonated on Nε2.

To explore how protonation states at the external His cluster and the internal carboxylic triad impact conformational dynamics of GPR68, we used the best-ranked GPR68 model (according to the compatibility of MD-equilibrated models with known structural features of class A GPCRs, see ranking protocol below) to perform independent simulations with different protonation states of selected Asp, Glu, and His sidechains (Table 2). These simulations (Sim #10-Sim #18 in Table 2) are distinguished by the protonation states of the carboxylic sidechains of the internal triad (D67$^{2.50}$, E149$^{4.53}$, and D282$^{7.49}$), of the nearby E103$^{3.34}$, and of H269$^{7.36}$ –as the H269F mutation impairs GPR68 function [13].

Force-field parameters and MD simulation protocol. Each GPR68 model was oriented in the membrane using the webserver Orientations of Proteins in Membranes, OPM [76]. We used CHARMM-GUI [68,77] to place GPR68 in hydrated bilayers of POPC (1-palmitoyl-2-oleoyl-sn--glycero-3-phosphocholine) lipids with neutralizing 0.15 M sodium chloride salt. In a separate set of test simulation, we studied the dynamics of GPR68 embedded in a POPC membrane with 10% cholesterol (Sim #9 in Table 2), which is within the value of cholesterol concentrations found in typical plasma membranes [78].

Table 2
Summary of MD simulations performed. Model# is according to Table 1. 'Standard' indicates that all titratable sidechains have standard protonation, i.e., all Asp/Glu sidechains are negatively charged, and all His sidechains are neutral and treated as –Nε2 tautomers. In simulations 10–18, one or two titratable sidechains had non-standard protonation –and we list the protonated oxygen atom of the neutral Asp/Glu (Sims 10, 11, 13–18), and the positively charged His sidechain (Sim 12). Lipid bilayers were composed of POPC, except for Sim9, in which we tested GPR68 with standard protonation states in a POPC membrane with 10% cholesterol.

Sim#	Model#	Protonation	Length (μs)	Lipid Bilayer Composition
1	1	standard	1	POPC
2	2		1	
3	3		1	
4	4		1	
5	5		1	
6	6		1	
7	7		1	
8	8		1	
9	5		0.6	POPC:cholesterol
10	5	E149-Oε1	0.6	POPC
11		E149-Oε2	0.6	
12		H269	0.6	
13		E103-Oε1	0.6	
14		E103-Oε2	0.6	
15		D282-Oδ1	0.6	
16		D67-Oδ2, E149-Oε1, D282-Oδ2	0.6	
17		D67-Oδ2, E149-Oε2, D282-Oδ2	0.6	
18		D67-Oδ2, E103-Oε1 E149-Oε1, D282-Oδ2	0.6	

We used the CHARMM36m force-field parameters for protein and lipids [79–82] with the TIP3P model [83] for water molecules, and NAMD v2.14 [84] to perform the simulations. Initial equilibration and geometry optimization of all systems was performed using the standard CHARMM-GUI protocol for velocity rescaling and constraints. Production runs were performed in the *NPT* ensemble (constant number of particles N, constant pressure $P = 1$ bar, and temperature $T = 303.15$ K) using a Langevin dynamics scheme with a collision frequency of 5 ps^{-1} and a Nosé -Hoover Langevin piston [85,86]. Coulomb interactions were calculated using the smooth-particle mesh Ewald method [87,88] and for real-space interactions we used a switching function between 10 and 12 Å. Lengths of bonds involving hydrogen atoms were fixed [89]. Heating, equilibration and the first 1 ns of the production run were performed with an integration step of 1 fs, and the remaining of each of the productions run, with a step of 2 fs.

Each simulation was prolonged to 0.6–1 μs, and coordinates were saved each 10 ps. We computed H-bond graphs and average values from the last 600 ns of Simulation #5, and the last 300 ns of Simulations #9–18 (Table 2).

Protocol to rank structural models of GPR68 equilibrated with atomistic MD simulations. As summarized above, six GOMoDo-generated and two AlphaFold-generated structural models were selected for MD simulations (Table 1). To ascertain the relevance of the equilibrated structures obtained from the eight independent MD simulations performed, we first monitored, for each simulation, the structural stability of the TM helical region according to root-mean-squared distance (rmsd) values relative to the starting coordinates, and the percentage of α-helical segments. Separately, each of the eight equilibrated structural models was closely inspected for its compatibility with generally accepted structural features of GPR68.

Graph computations of dynamic H-bond networks. Dynamic H-Bond Networks and water wires were calculated using Bridge2 [90–92]. Briefly, Bridge2 is a set of graph-based algorithms with graphical user interface for highly efficient computations of protein-water H-bond networks. A *graph of H-bonds* consists of nodes, which in our case are the H-bonding sidechains of GPR68, and edges, which are direct or water-mediated H-bonds between sidechains. H-bonds are computed using geometric criteria; here, H-bond computations for the protein-water H-bond were performed with a distance criterion of ≤ 3.5 Å between the H-bond donor and acceptor hetero-atoms, and an H-bond angle of $\leq 60^{\circ}$ [93]. H-Bonds between GPR68 and POPC were computed residue-wise for all phosphate oxygen atoms of a POPC lipid molecule.

Water-mediated bridges between protein sidechains were computed with maximum three H-bonded water molecules per bridge. Thus, the average length of a water-mediated bridge between two protein sidechains can take values between 0.0 (i.e., direct H-bond between the sidechains) and 3.0 (i.e., the sidechains are bridged by three H-bonded water molecules [90,91]. For simplicity, for the sidechain-lipid contacts we considered only the direct H-bonds between protein sidechains and lipid phosphate groups.

An *H-bond cluster* consists of nodes and edges that are all interconnected to each other. An *H-bond path* between two nodes *i* and *j* of the graph consists of the nodes and edges that interconnect *i* and *j*. The *joint occupancy (JO)* of an H-bond path reports the percentage of coordinate sets from the trajectory segment used for analyses in which all intermediate path segments are sampled simultaneously. H-bond paths and H-bond clusters were extracted from the H-bond graphs using Connected Components Analyses [94] as implemented in Bridge/Bridge2.

For clarity, unless otherwise specified, we show H-bond graphs with H-bond occupancies of at least 30%. This minimum H-bond occupancy level was found reasonable for comparisons of the protein H-bond networks sampled in independent simulations of a model soluble protein [95].

Inter-helical distances in GPR68 models and static structures of GPCRs. A geometrical evaluation was introduced in ref. [96] to compare inactive vs. active structures of GPCRs. This geometrical evaluation included a set of inter-helical distances measured between the Cα atoms of G35$^{1.46}$, D67$^{2.50}$, V109$^{3.40}$, W146$^{4.50}$, P197$^{5.50}$, F237$^{6.44}$ and C279$^{7.46}$ (Fig. S2C); ref. [96] used a dataset of 25 static GPCR structures, of which 7 were presented for active conformations, and 18, for inactive conformations. We measured the same inter-helical distances in each of the GPCRs structures used as a template for the GPR68 homology models and monitored the time series of the distances along simulations #1-#8 (Table 1).

Time series and average values of the number of internal water molecules. To monitor the number of internal water molecules in homology models of GPR68 we used a protocol as presented in ref. [43]. Thus, we evaluated the membrane region according to the average location of the phosphate atoms of the lipid bilayer within 10 Å of the protein and monitored the number of water oxygen atoms within $\pm$ 5 Å of these membrane boundaries.

Molecular graphics were prepared with Visual Molecular Dynamics, VMD [97].

3. Results and discussion

We derived eight structural models of GPR68 using different structural templates and two different modeling approaches (GOMoDo, with Modeller used to refine the N-terminus and impose disulfide bridges, vs. AlphaFold, see details for Model #1-#8 in Table 1). Each of these structural models were refined via a prolonged equilibration in a hydrated lipid membrane environment. The ninth model, derived using only Modeller (i.e., a third homology modeling approach) and bovine rhodopsin as a template, was used solely for comparison with a previous homology model. During the MD simulations all models derived with GOMoDo (and Modeller for the N-terminus) or AlphaFold had good structural stability, as indicated by the ~70% of the polypeptide chain being α-helical, Cα RMSD for the TM helical segments mostly within ~2.5–3 Å, and the number of water molecules found within the inter-helical region of the receptor (Table 1, Figs. S3, S4, S5).

To evaluate the compatibility of the GPR68 structural models with known structural features of class A GPCRs, we first tested the usefulness of a set of inter-helical distance pairs that had been proposed before to characterize inactive vs. active conformations of GPCRs [96]. As summarized below, we found that the usage of these distance pairs is inconclusive, as reference values depend on the dataset of experimental static structures. Consequently, for the compatibility tests of the GPR68 models we used instead a set of specific structure features generally accepted for class A GPCRs. The best-ranked GPR68 model was used to verify how internal protein-water H-bond network of potential interest to GPR68 activation respond to changes in the protonation state.

Reference inter-helical distance pairs depend on the dataset of experimental static structures. Dalton and colleagues [96] used inter-helical distances (Fig. S2C) to evaluate structural dynamics in a dataset of 25 static structures of class A GPCRs, and to compare selected inter-helical distances of inactive vs. active-like GPCR structures. They reported that most distance pairs had largely similar values in all static GPCR structures, distances between TM2 and TM3, and between TM3 and TM7, differed by more than 2 Å in active-like vs. inactive GPCRs [96]. We verified these distances for both the static structures we used as templates to model GPR68 (Tables S4-S6), and for conformations sampled during the simulations (Table S7).

The dataset used by Dalton and colleagues [96] and the dataset of static structures we used as templates for GPR68 have in common three entries: the neurotensin receptor in an active-like state [98], the human protease-activated receptor 1 in an inactive conformation [99], and the antagonist-bound CXCR4 chemokine receptor structure [100]. Most average values for inter-helical distances pairs of the two data sets are within relatively good agreement with each other – except for TM1-TM7, TM2-TM7, and TM6-TM7, which differ by ~2–3 Å (Table S6). When accounting for the standard deviation values computed from our dataset, it appears that using the TM2-TM3 and TM3-TM7 distance pairs to characterize inactive vs. active GPCR confirmations is somewhat inconclusive: for some of the helix pairs, differences between distances measured in active vs. inactive structures of the dataset we used are comparable with the distance values within the subset of active-like structures of the current data set vs. the dataset from ref. [96]: average inter-helical distances from the dataset of template structures we used, and the dataset from ref. [96] can differ by as much as 2.6–2.7 Å (Table S6), which is comparable with the values suggested as a characteristic of inactive vs. active GPCR structures [96]. Relative to the corresponding template structures, average values of inter-helical distances computed from the simulations are different by up to ~2–3 Å for several distance pairs in Models #1, #4, and #7 (Table S7, Fig. S6).

To summarize, our analyses suggest that average inter-helical distances in static GPCR structures are dataset-dependent, suggesting limited transferability. Fluctuations of inter-helical distances during MD simulations of a structural model depend on the template used, with Models #1, #4, #6 and #7 exhibiting largest differences between values of specific inter-helical distances measured for the static protein template vs. MD of the corresponding structural model.

A recent study in which TM helix movements were studied for 13 GPCRs whose structures exist for both the inactive and G-protein bound active conformations revealed that, for most of the receptors from the dataset, TM helices rearranged by at least 1 Å at one or both ends; an outward movement and rotation of TM6 was identified as a universal feature of the GPCR activation mechanism [101].

A direct comparison with the TM helix movements recently discussed by Hauser and colleagues [101] would require a study of the activation

pathway of GPR68, including modeling of the GPR68-G protein interactions, which is beyond the scope of our current work. Moreover, none of the static structures analyzed in ref. [101] were used as templates for the GPR68 structural models used here. Nevertheless, we compared inter-helical distances measured for structures of class A GPCRs used in ref. [101] with the corresponding average distances computed for the equilibrated structural models Models #1-#8. Results summarized in Figs. S6 and S7 suggest that a direct comparison of the equilibrated GPR68 structural models with static structures of other GPRCs is indeed somewhat difficult –e.g., mean inter-helical distances computed from simulations of Model #5 tend to have values in between those measured for static structures of inactive vs. active class A GPCRs [101]. This could be interpreted to suggest that, within the structure dataset used for comparisons, Model #5 is overall compatible with structures of class A GPCRs.

Assessment of GPR68 models based on interactions at key functional sites. We inspected closely the equilibrated structural models to ascertain whether specific structural elements have interactions compatible with the current knowledge on class A GPCRs (Table 3).

Model #1 has Y205$^{5.58}$ oriented towards the lipid membrane (Figs. S8A, S9A), whereas static GPCR structures and previous simulations of class A GPCRs suggest that Y205$^{5.58}$ locates inside the TM helical region [102–106]. Consequently, we discarded Model #1 from further analyses. Membrane orientations of Y205$^{5.58}$ are also observed in *Model #2* (Figs. S8B, S9B) and *Model #3* (Fig. S8C, S9C). In addition to the incorrect orientation of Y205$^{5.58}$, Model #2 has TM7 reoriented such that groups of the DPxxY motif move away from the sodium-binding pocket (Fig. 2A, S10B) and lacks sodium ion binding to the internal pocket (Fig. S10B), whereas Model #3 has three sodium ions that bind to the protein at distinct time points along the simulation (Fig. S10C). Consequently, we discarded Model #2 and Model #3. We further discarded *Model #4* as it has an opening of the PVF motif with reorientation of F237$^{6.44}$ and CFxP-F241 (Fig. S10D); Model #4 also has a somewhat unusual sodium binding dynamics, as one sodium ion enters and remains stably bound at the sodium-binding pocket, but two additional sodium ions visit the protein at distinct moments of time along the simulations (Fig. S10D).

Model #5 has good overall stability with α-helical content of ~80% throughout the entire length of the simulation (Fig. S3E), inter-helical distances overall consistent with corresponding values reported for static structure of GPCRs [96] (Table S7), correct orientation of Y205$^{5.58}$ (Fig. S8E, S9E), and a stably-bound single sodium ion that enters the protein from the bulk and then it remains bound at the sodium pocket throughout the entire simulation (Fig. 3). The profile for the number of water molecules in the inter-helical region reaches a plateau value at about 100 water molecules (Fig. S5E). Taken together, the analyses summarized here indicate that the MD simulation of Model#5 has converged, and that Model #5 is compatible with known structural parameters of class A GPCRs, and thus we selected it for further investigation.

Model #6 has two sodium ions bound at the sodium-binding site –one sodium ion interacting closely with D67$^{2.50}$, and the other sodium ion, with D282$^{7.49}$ (Fig. S11B). *Model #7* lacks sodium ion binding to the internal pocket (Fig. S11C), is tilted relative to the membrane normal (Fig. S11C), helix 8 - R301$^{8.58}$ oriented towards interacts D282$^{7.49}$, and Y205$^{5.58}$, towards the lipid membrane. *Model #8*, which differs from Model #7 in that C304 and C310 are palmitoylated, has about the same protein tilt as Models #1-#6 (Fig. S11D), one stably-bound sodium ion at the sodium-binding site (Fig. S11D), the sidechain of Y205$^{5.58}$ oriented towards the membrane (Fig. S8H, S9H), and a D282$^{7.49}$ - R301$^{8.58}$ interaction as observed in Model#7 (Fig. S11D). Models #6–8 were thus discarded from further analyses.

To summarize, our analyses of GPR68 models indicated that Model #5 agrees best with known structural features of class A GPCRs. We thus kept for further analyses only Model #5. In what follows, for simplicity, we refer to Model #5 as GPR68, without the Model index.

An extended protonation-coupled H-bond network inter-connects the extracellular His cluster to the internal carboxylic triad. Numerous water molecules visit the inter-helical region of GPR68 (Fig. 4, S5E): there are, at any given time during the simulations on GPR68 with standard protonation states, about 150 water-mediated wires of up to three H-bonded water molecules between protein sidechains (Fig. 4C). Most of these water wires are very dynamic, with occupancies $< 10\%$ (0.1% in the normalized occupancy plot presented in Fig. 4B). Nevertheless, as discussed below, waters that visit the inter-helical region of GPR68 can participate in dynamic H-bond networks that inter-connect remote regions of the receptor (Fig. 5, S12).

At the extracellular side, the His sidechains proposed to function as pH sensors of GPR68 [13] are part of an extended water-mediated H-bond network that includes several carboxylic sidechains: H89$^{23.52}$ and H169^{ECL2} are close to D91$^{3.22}$, E164^{ECL2}, D165^{ECL2}, and E166^{ECL2}, whereas H17$^{1.28}$, H20$^{1.31}$, and H84$^{2.67}$, are close to D16$^{1.27}$, D85^{ECL1}, E174$^{45.52}$, D255^{ECL3}, and E259^{ECL3} (Fig. 5A). This extracellular cluster of His, Asp, and Glu sidechains connects, via a protein-water H-bond network, to the core of the protein, where the carboxylic groups of the internal triad engage in water-mediated H-bonds with other charged and polar protein sidechains (Fig. 5A). Essentially all local H-bond clusters of the carboxylic and histidine groups have water-mediated bridges between sidechains (Fig. S12).

E149$^{4.53}$ of the internal carboxylic triad water-bridges to E103$^{3.34}$ at all times during the simulation with standard protonation states (Fig. 5 B-F); this is a short, mostly one-water mediated bridge (Fig. S12B). The E103$^{3.34}$-E149$^{4.53}$ water bridge connects transiently to D67$^{2.50}$ and D282$^{7.49}$ of the internal carboxylic triad (Fig. 5D, S12D), as well as to H-bond paths leading to H245$^{6.52}$ and H269$^{7.36}$ (Fig. 5B, S12B), H175^{ECL2} (Fig. 5C, S12D) or to D85^{ECL1}/E174$^{45.52}$ (Fig. 5E, S12E), and to the local H-bond cluster of D91$^{3.22}$, E164^{ECL2} and E166^{ECL2} (Fig. 5F, S12F). D85^{ECL1}, D91$^{3.22}$, and E174$^{45.52}$ are part of local water-mediated H-bond clusters with H175^{ECL2}, H245$^{6.52}$ and H269$^{7.36}$ (Fig. 5G-J, S12G-J).

The emerging picture is that water-mediated H-bond paths, typically mediated by 2–3 H-bonded water molecules, with only a handful of bridges mediated by one water molecule on the average (Fig. S12), establish dynamic long-distance connections between internal carboxylic groups, and between internal carboxylic groups and the extracellular histidine-carboxylic cluster.

Some of the water-mediated H-bond clusters that inter-connect the internal carboxylic groups to the extracellular cluster contain H-bonds

Table 3
Summary of structural parameters used to evaluate the conformational dynamics of the GPR68 homology models with standard protonation states for all titratable sidechains. We report, for each model, the average RMSD values of the TM helices computed every 10 ps from the last 600 ns of each simulation.

Model#	RMSD (Å)	Observations	Figures
1	2.6 ± 0.2	Y205$^{5.58}$ faces the membrane instead of the interior of the protein	S8A, S9A, S10A
2	2.5 ± 0.2	Y205$^{5.58}$ faces the membrane, PVF motif distorted, helix 8 unstructured	S8B, S9B, S10B
3	2.4 ±0.2	Y205$^{5.58}$ faces the membrane, TM3 partially unstructured, three sodium ions enter the receptor	S8C, S9C, S10C
4	3.0 ±0.3	TM4 partially unstructured, one sodium ion at the binding pocket, two more sodium ions visit, PVF motif distorted	S8D, S9D, S10D
5	2.3 ±0.1	Good structural stability, one sodium ion bound, correct orientation of Y205$^{5.58}$	S8E, S9E, S11A
6	2.3 ±0.2	2 sodium ions interact with D67$^{2.50}$ and D282$^{7.49}$, TM4 partially unstructured	S8F, S9F, S11B
7	2.1 ±0.2	Y205$^{5.58}$ faces the membrane. TM7 partially unstructured, helix 8 tilted, D282$^{7.49}$ interacts with R301	S8G, S9G, S11C
8	1.6 ±0.2	Y205$^{5.58}$ faces the membrane, D282$^{7.49}$ interacts with R301	S8H, S9H, S11D

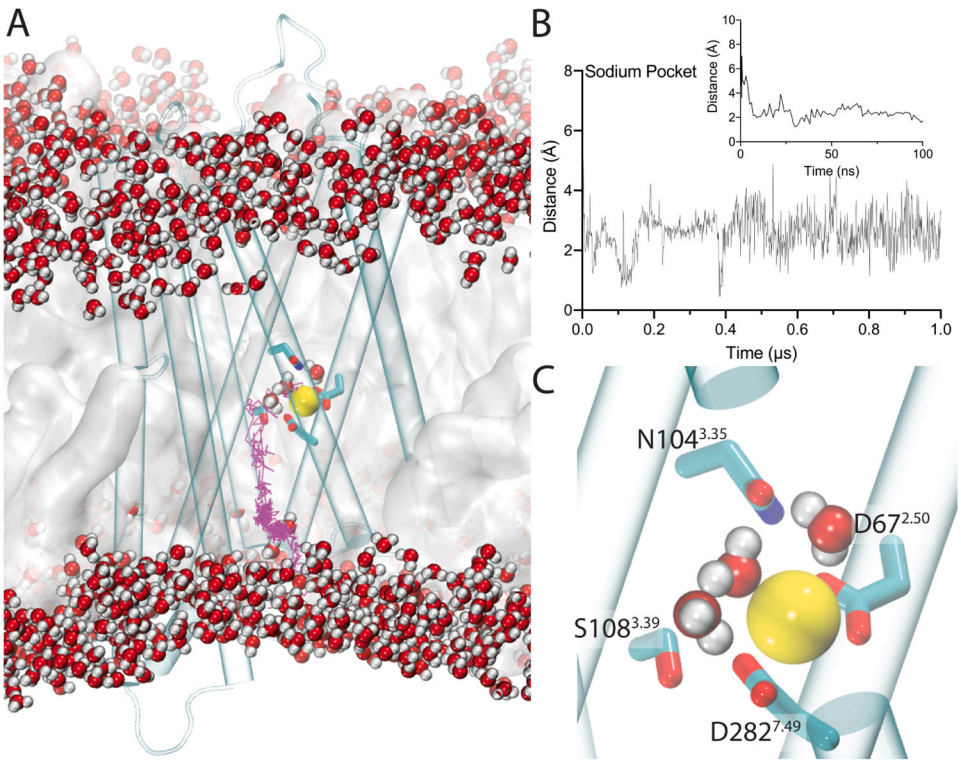

Fig. 3. Dynamics of the sodium entry in the sodium-binding pocket of GPR68 with standard protonation states. (A) The trajectory of the sodium ion entry in the sodium pocket of GPR68 Model#5 is shown in purple. Internal water molecules shown are within 3.5 Å of the sodium ion, which is shown as a yellow sphere. (B) Time series of the sodium ion entry to the sodium pocket. We report the distance (in Å) between the geometric center of the Cα coordinates of D67$^{2.50}$, N104$^{3.35}$, S108$^{3.39}$, and D282$^{7.49}$; the insert illustrates the rapid binding of the sodium ion within ~30 ns. The sodium ion remains bound at the sodium-binding site through the end of the 1 μs simulation. (C) Close view of water and protein interactions the sodium-binding pocket.

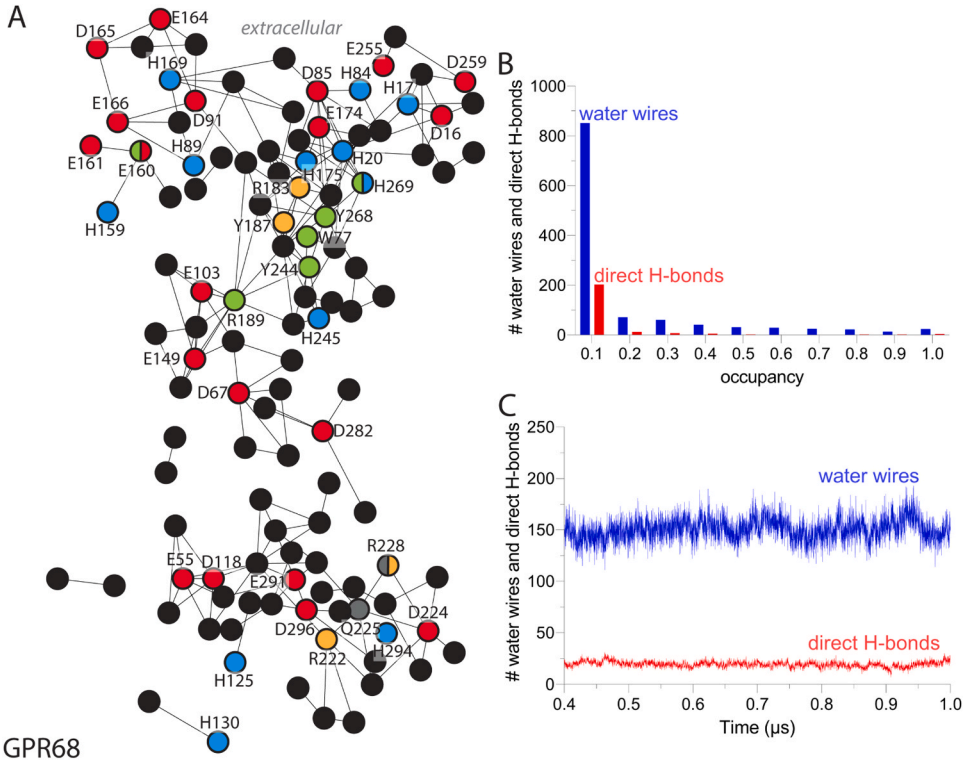

Fig. 4. Dynamic water mediated protein-water H-bond network of GPR68. (A) H-bond network of GPR68. Graph nodes representing Asp/Glu sidechains are colored red, His sidechains, blue, groups involved in allosteric modulator binding, green, and all other H-bonding sidechains, black. Note that H269$^{7.36}$ and E160$^{4.64}$ are also a part of the allosteric binding site. For clarity, only selected nodes are labeled. The minimum H-bond occupancy shown is 30%. (B) Histogram of the average occupancy of all H-bond connections. (C) Time series of all unique H-bond connections sampled at least once during the trajectory segment used for data analyses. The blue profile shows the number of water wires between protein side chains, and red, direct H-bonds between protein sidechains.

that are rather dynamic, with occupancies < 50%; but, in the same local H-bond cluster, distinct water-mediated H-bond bridges may have rather different occupancy values. For example, in the H-bond cluster that connects E149$^{4.53}$ to D85^{ECL1}, the water-mediated bridges between E149$^{4.53}$ and R189$^{5.42}$, and between R189$^{5.42}$ and N186$^{5.39}$, have occupancies of, respectively, 97% and 70%; the two water-mediated bridges that connect N186$^{5.39}$ to E174$^{45.52}$ via H175^{ECL2} have occupancies of 54–55%; the water bridge from E174$^{45.52}$ to D85^{ECL1} is present during 30% of the time. This long-distance path between E149$^{4.53}$ and D85^{ECL1} thus passes via 4 sidechains (and about 10 water molecules, Fig. S12E) has a JO value of 13% (Fig. 5E), that is, 13% of time during the last 600 ns of the simulation, all H-bonds needed to connect E149$^{4.53}$

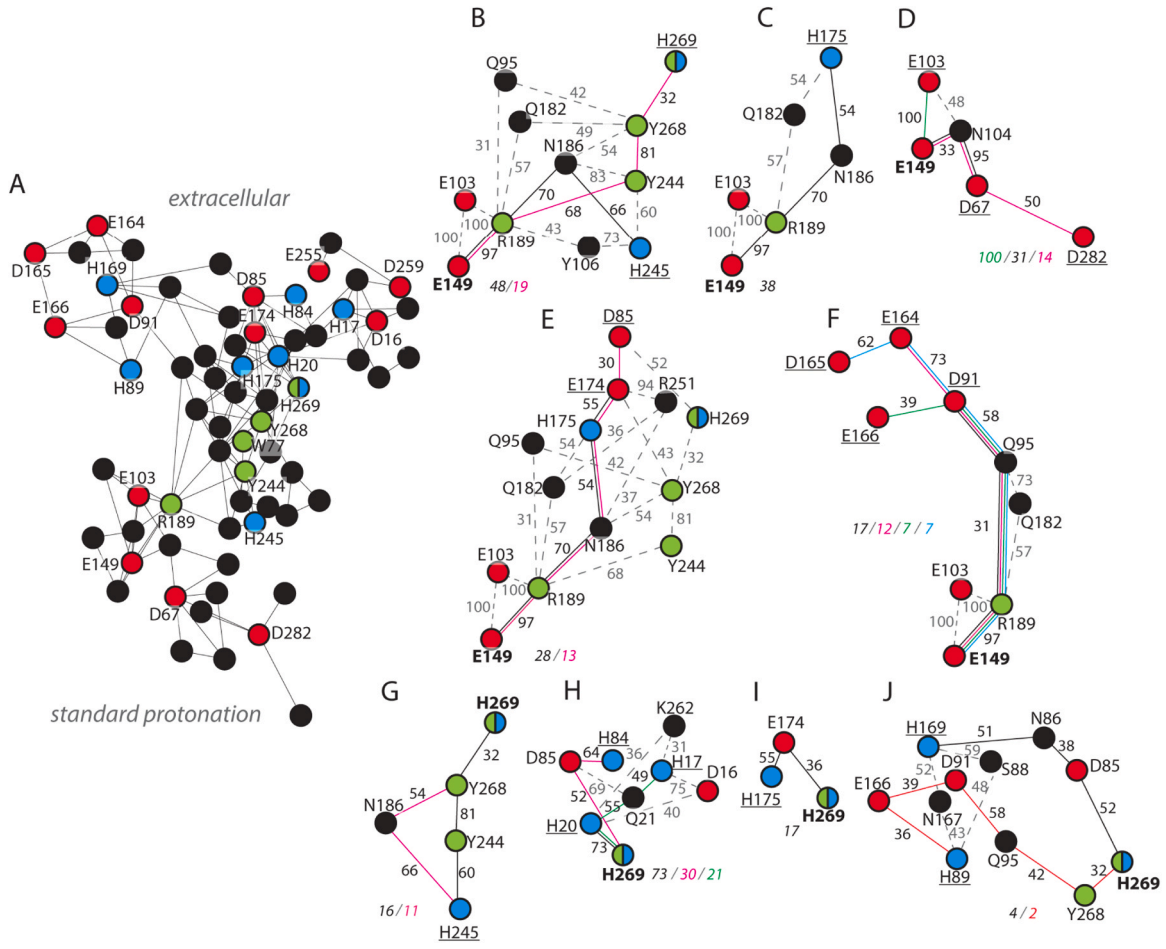

Fig. 5. Water-mediated H-bond network of GPR68. The graph computation and color codes are the same as detailed for Fig. 4. Note that H269 is a part of the allosteric binding site. For clarity, only selected nodes are labeled. (A) The H-bond cluster of E149. (B-J) Local H-bond clusters that inter-connect selected His, Asp, and Glu sidechains. Color-coded numbers in italics are JO values of the corresponding H-bond paths. The average number of water molecules in each bridge depicted in panels B-J are shown in Fig. S12.

to D85^{ECL1} over distances of ~24–28 Å are sampled simultaneously. The slightly shorter path from E149$^{4.53}$ to E174$^{45.52}$ has a larger JO value, 28% (Fig. 5E), such that E149$^{4.53}$ connects to the extracellular bulk, via a continuous path via R189$^{5.42}$, N186$^{5.39}$, and H175^{ECL2}, for almost one third of the time. E149$^{4.53}$ may also connect to carboxylic sidechains of the E164^{ECL2}-D165^{ECL2}-E166^{ECL2} cluster – the JO values of these paths are 7–17% (Fig. 5F, Table S8), suggesting a somewhat more dynamic path. H269$^{7.36}$ connects to H84$^{2.67}$ via D85^{ECL1} and ~5 water molecules during ~30% of the time (Fig. 5H, S12H).

To verify how results of the H-bond cluster analyses might depend on the length of the trajectory segment used for data analyses, we repeated the analyses summarized above by using, instead of the last 600 ns, only the last 300 ns of the simulations of GPR68 with standard protonation states. Results summarized in Fig. S13 show that H-bond graphs computed from the last 300 ns vs. the last 600 ns of the simulation have overall qualitatively similar connections of E149$^{4.53}$ to different regions of the protein; average H-bond occupancies are very similar, with variations within 12% (for more details see Fig. S13). Based on this finding, H-bond-graphs for simulations in which we probe the response of GPR68 to changes in protonation are computed from the last 300 ns of each of these simulations.

The analyses summarized above suggest that GPR68 with standard protonation states of all His and carboxylic sidechains hosts an extensive water-mediated H-bond network that connects the extracellular histidine-carboxylic cluster to internal carboxylic sidechains, including the carboxylic triad. In spite of their length, some of the H-bond paths that connect internal carboxylic sidechains to the extracellular side via multiple water-bridged sidechains can nevertheless have all intermediate H-bond segments sampled simultaneously. We suggest that these H-bond networks might help mediate long-distance conformational couplings between remote regions of the protein.

Protonation of the internal acidic triad impacts the H-bond connections of the extracellular histidine-carboxylic cluster. GPR68 with standard protonation states for all titratable sidechains lacks H-bond connections between the DRY motif and the core H-bond network (Fig. 5A). By contrast, when both E149$^{4.53}$ and E103$^{3.34}$ are protonated, the core network H-bond network of GPR68 includes shortest-distance H-bond paths that inter-connect H245$^{6.52}$ and H269$^{7.36}$ to D67$^{2.50}$, to DPxxY-Y286$^{7.53}$, and to DRY R119$^{3.50}$ (Figs. 6A, 6C). But the sampling of these long-distance H-bond paths depends on details of the protonation of E103$^{3.34}$ and E149$^{4.53}$ – the paths are found when either of the Glu sidechain is –Oε1 protonated (Figs. 6A, 6C), and not –Oε2 (Figs. 6B, 6D).

Negatively charged D282$^{7.49}$ is part of the long-distance H-bond paths that extend from DRY-R119$^{3.50}$ to H269$^{7.36}$ when E103$^{3.34}$ or E149$^{4.53}$ are –Oε1 protonated (Figs. 6A, 6C), or when only H269$^{7.36}$ is positively charged (Fig. 6F); these paths include 6–7 sidechains, and about 15–17 H-bonded water molecules (Figure S17 A,C,F). When only D282$^{7.49}$ is neutral, its local H-bond cluster with D67$^{2.50}$, S108$^{3.39}$ and H245$^{6.52}$ connects to the extracellular side, but lacks bridges to the cytoplasmic DRY motif (Fig. 6E). Instead, DRY-R119$^{3.50}$ bridges to DPxxY-Y286$^{7.53}$ via mostly one water molecule; D282$^{7.49}$ bridges to H245$^{6.52}$ via ~2.8 water molecules, and it H-bonds directly to D67$^{2.50}$

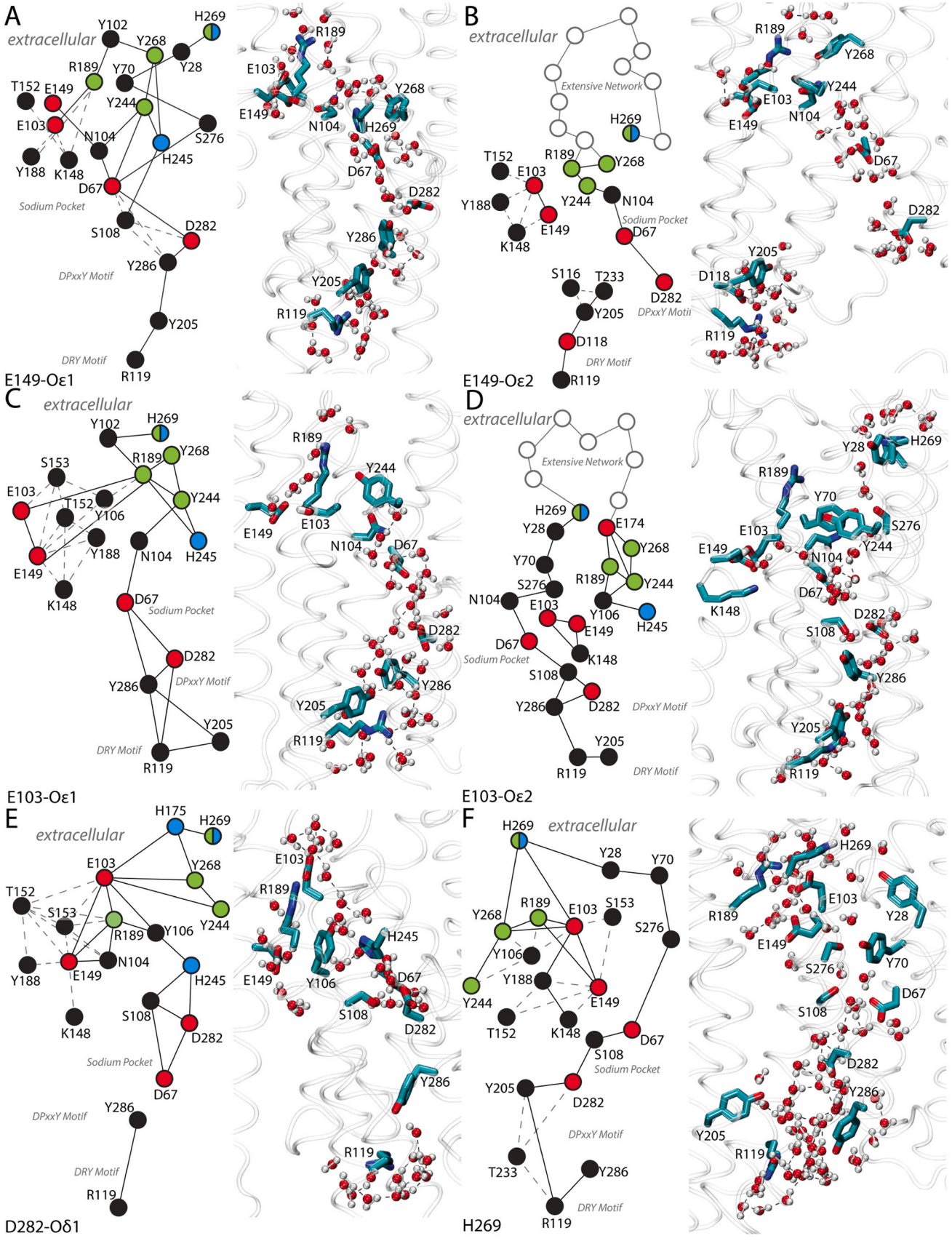

Fig. 6. Protonation-coupled dynamics of the H-bond network of GPR68. Empty gray nodes indicate additional H-bond connections as shown in Figs. S15 B&D. (A-F) Core network of GPR68 with protonated E149-Oε1 (panel A), protonated E149-Oε2 (panel B), protonated E103-Oε1 (panel C), protonated E103-Oε2 (panel D), protonated D282-Oδ1 (panel E), and positively charged H269 (panel F). Note that these simulations lack transient sodium ion binding at the sodium-binding site. Additional information, including the average number of water molecules in bridges, are presented in Figs. S14-S15, and S17.

(Fig. 6E, S17E).

Two independent simulations with all three carboxylic sidechains of the internal triad protonated suggest that such protonation state appears incompatible with extensive local connections involving all three carboxylic sidechains, which are instead involved in distinct H-bonds or H-bond clusters (Figs. 7A, 7B). A long-distance H-bond path from the DRY motif to the extracellular histidine-carboxyl cluster via D282$^{7.49}$ and D67$^{2.50}$ is sampled when, in addition to the three carboxylic sidechains of the internal triad, E103$^{3.34}$ is also protonated –but, in this protonation state, E103$^{3.34}$ and E149$^{4.53}$ are no longer a part of the core H-bond network (Fig. 7C). GPR68 with positively charged H269$^{7.36}$ rearranged its internal H-bond network such that H269$^{7.36}$ directly connects to E103$^{3.34}$ and has an extended H-bond network that connects all core areas of GPR68 (Fig. 6F).

Membrane lipid interactions of GPR68. Overall, regardless of the protonation state, there tend to be 3–5 sidechains that sample H-bonds to lipid phosphate groups from the surrounding membrane, with an average occupancy ≥ 70% (Table S9, Figures S19-S23) –these are typically Arg and Lys sidechains, though Thr sidechains can also H-bond to lipid phosphate groups. In each simulation two, sometimes three protein sidechains, interact with the same lipid phosphate group (Table S9). The protein sidechains H-bonding to the same lipid are typically relatively close in sequence – such as T134$^{4.38}$ and K136$^{4.40}$, R210$^{5.63}$ and R213$^{5.66}$, R183$^{5.36}$ and Y187$^{5.40}$, or H125$^{3.56}$ and R128$^{34.52}$ (Table S9, Figures S19-S23). Though details of H-bonding occupancies differ, the same amino acid residues or sequence regions tend to bind membrane lipids. High-occupancy lipid binding sites at the membrane-exposed surface of GPR68 are located mainly at the intracellular side, except for T14$^{1.25}$, R183$^{5.36}$ and Y187$^{5.40}$, which are at the extracellular side of GPR68 (Fig. 8A). R183$^{5.36}$ and Y187$^{5.40}$ are well connected to the H-

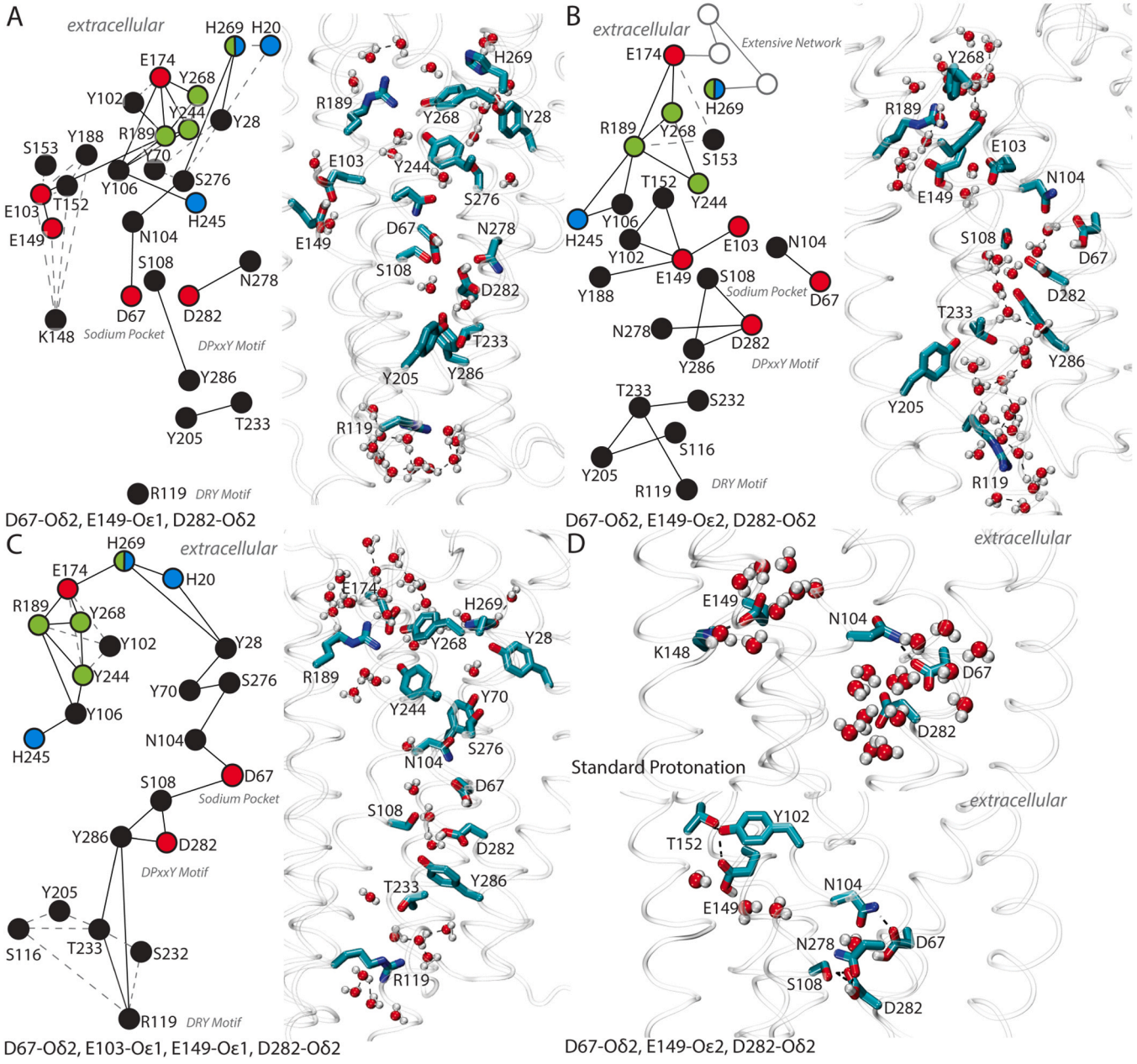

Fig. 7. Protonation of all three sidechains of the internal carboxylic triad associates with perturbed local interactions. (A-E) Core protein-water H-bond network of GPR68 with protonated D67-Oδ2, E149-Oε1, D282-Oδ2 (panel A), protonated D67-Oδ2, E149-Oε2, D282-Oδ2 (panel B), protonated D67-Oδ2, E149-Oε1, D282-Oδ2 (panel C). (D) Direct-H bonding partners and waters within 3.5 Å of the internal carboxylate triad (D67, E149, D282) when the receptor has standard protonation and when the triad is D67-Oδ2, E149-Oε2, D282-Oδ2 protonated. Additional information is presented in Figs. S16, S18.

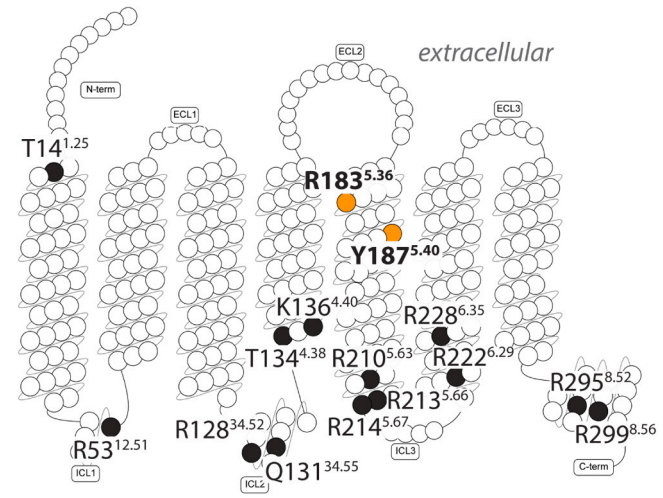

Fig. 8. Summary of high-occupancy POPC binding sites sampled in simulations of GPR68 in a hydrated POPC membrane, for all protonation states considered in this work. (A) Snake diagram of GPR68 showing selected lipid-binding sites. Amino acid residues that sample H-bonding to lipids are shown as black dots, except for R183 and Y187, which are colored peach to indicate them as persistent lipid-binding sites (H-bond occupancy of at least 70%). Note that R183 and Y187 are also part of the GPR68H-bond network shown in Fig. 4A. In the test simulation performed for GPR68 in a hydrated POPC:cholesterol membrane, R183 and Y187H-bond to the same POPC phosphate group with occupancies of 64% and 27%, respectively; transient H-bonds between cholesterol and GPR68 (H-bond occupancies <5%) are sampled by R53, R210, and R213. Details on lipid interactions of GPR68 in hydrated POPC membranes are presented in Table S9 and Figs. S19-S23.

bond network of GPR68 (Fig. 4A).

All simulations discussed above were performed with GPR68 embedded in a hydrated POPC membrane, as a simpler membrane system facilitated comparisons of the motions of the distinct structural model we derived here, and of the response of GPR68 to changes in the protonation state. Given the important potential role of cholesterol as 'modulator of receptor function' [107], we performed an additional test simulation of GPR68, with standard protonation states, embedded in a POPC membrane with 10% cholesterol (Table 2). We computed from this simulation the H-bond graph and local H-bond clusters corresponding to those presented in Fig. 5 for GPR68 in POPC. As illustrated in Figure S24B, on the timescale of the test simulation with the POPC: cholesterol membrane we find residue N104$^{3.35}$ of the sodium-binding site lack connections to E149$^{4.53}$ and E103$^{3.34}$ as sampled in the reference simulation of GPR68 in POPC (Fig. 5 A,D). Whereas such an apparent sensitivity of the internal H-bond dynamics to the presence of cholesterol might indeed appear compatible with a putative role of cholesterol in shaping GPR68 dynamics, is unclear at the moment; we suggest that further simulations and experimental validation would be needed.

Lipid interactions at the groups corresponding to GPR68 K136$^{4.40}$, R213$^{5.66}$, and R228$^{6.35}$ were found in simulations of the μ-opioid receptor [108]. Likewise, groups corresponding to GPR68 R222$^{6.29}$ and R228$^{6.35}$ interact with lipids in simulations of the adenosine A1 receptor [109]. That is, in three distinct GPCRs, the group at position 6.35 H-bonds to lipids. As R228$^{6.35}$ and the nearby Q225$^{6.32}$ are thought important for interactions between GPRCs and the G-Protein [110], and G proteins themselves interact with lipid membranes [111], the question arises as to whether, in a GPCR-G protein complex, R228$^{6.35}$ and Q225$^{6.32}$ (Fig. 9B) could be part of more complex lipid-protein H-bond clusters together with G protein groups.

A test GPR68 homology model with bovine rhodopsin as a template. Early site-directed mutagenesis work on GPR68 used for interpretation a homology model generated based on a structure of bovine rhodopsin [13]; based on that homology model, H245$^{6.52}$ was suggested to be important for the structural integrity of GPR68, and to be oriented toward the lipid membrane [13]. By contrast, in the GPR68 model we selected here (Model#5), the sidechain of H245$^{6.52}$ is oriented towards the interior of the protein, where it connects to E149$^{4.53}$ (Fig. 5A) and to a cluster of amino acid residues implicated in the binding of positive allosteric modulators (see R189$^{5.42}$, Y244$^{6.51}$, Y268$^{7.35}$ and H269$^{7.36}$ in Fig. 5B) [112].

To verify whether the membrane orientation of H245$^{6.52}$ proposed in the earlier work [13] might have been due to the template used, we

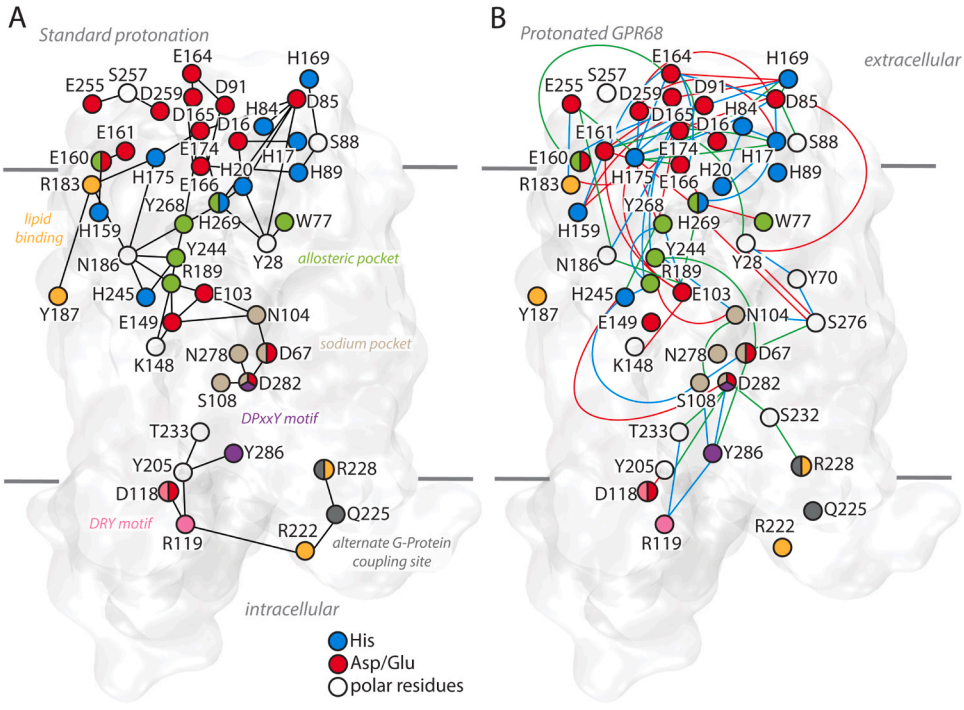

Fig. 9. Summary of main results on the protonation-coupled dynamics of GPR68. (A) Schematic representation of the H-bond network of GPR68 with standard protonation states for all titratable sidechains. An H-bond network with multiple Asp, Glu, and His sidechains extends from D282$^{7.49}$ to the extracellular side. The DRY motif lacks persistent H-bond connections to this network. (B) Response of the internal H-bond network of GPR68 to changes in the protonation state. Lines colored green and red indicate H-bond connections that depend on the protonation state, and cyan, H-bond connections sampled in simulations of GPR68 with non-standard protonation states, as follows: green –protonated E149-Oε1, protonated E103-Oε1, positively charged H269, protonated D67-Oδ2, protonated E103-Oε1, and protonated E149-Oε1 and D282-Oδ2; red –protonated E149-Oε2, protonated E103-Oε2, protonated D282-Oδ1, and protonated D67-Oδ2, E103-Oε1, E149-Oε1, D282-Oδ2. Note that, depending on the protonation state, transient connections are sampled between DPxxY-Y286 and DRY-R119, and between DPxxY-Y286 and the extracellular H-bond network, such that the H-bond network of GPR68 extends from the extracellular pH-sensing region to the cytoplasmic G-protein binding site.

generated a test homology model of GPR68 using as a template the same bovine rhodopsin structure. And, indeed, when we bovine rhodopsin as a template, we obtained for GPR68 a homology model in which H245$^{6.52}$ orients towards the membrane (Figure S25).

Given the absence of independent data on the most likely orientation of H245$^{6.52}$, we inspected more closely the test homology model of GPR68 obtained using the bovine rhodopsin structure as a template. We found that W77$^{2.60}$, Y205$^{5.58}$ and Y268$^{7.35}$, are oriented towards the membrane, instead of the interior of the receptor, as indicated by docking and site-directed mutagenesis studies (for W77$^{2.60}$ and Y268$^{7.35}$) [113], and (for Y205$^{5.58}$) by analyzing the previously solved Class A GPCR structures and their residue contact frequency maps, and mutagenesis studies [77–81]. Moreover, in the homology model based on bovine rhodopsin, both E103$^{3.34}$ and E149$^{4.53}$ are oriented towards the membrane core.

Pursuant to the considerations above we suggest that, overall, the homology model of GPR68 based on bovine rhodopsin might be insufficiently reliable to support the initial suggestion [13] that H245$^{6.52}$ would have a structural role. Instead, based on the structural modeling and simulation work pursued here, we suggest that H245$^{6.52}$ takes part in the extracellular histidine-carboxyl cluster.

4. Conclusions

Changes in extracellular pH may characterize cancer tissue and could be exploited to develop treatment strategies [114]. Proton-sensing GPCRs, used by specialized cells to respond to changes in the extracellular pH, are of interest as drug targets for solid tumors associated with an acidic extracellular environment. The mechanism by which proton-sensing GPCRs work is unclear, and conflicting scenarios exist: Early site-directed mutagenesis data and homology modeling of GPR68 were interpreted to suggest that a cluster of His sidechains at the extracellular side of the receptor binds protons at acidic pH [13]. By contrast, more recent work argued in favor of an internal carboxylic triad as the main pH-sensing unit [22].

Until now, three-dimensional structures of pH-sensing GPCRs have proved difficult to determine with structural biology. Only one such structure has been determined, for GPR65 (as accessed on December 5, 2022) [115], but it remains proprietary with no public access to the structural information. To derive insight into intra-molecular interactions potentially important for the molecular mechanism of a pH-sensing GPCR, we generated multiple structural models of GPR68, refined their structures with atomistic simulations in hydrated membrane environments, and ascertained their compatibility with generally accepted structural features of class A GPCRs. We selected the model best compatible with these structural features and used it to evaluate the protonation-coupled dynamics of GPR68.

Dynamic protein-water H-bond networks are thought essential for propagation of agonist-induced conformational changes leading to the formation of an active receptor state [43, 49, 50, 92]. The internal protein-water H-bond network of a pH-sensing GPCR such as GPR68 would thus need to respond to changes in the protonation state. Given the uncertainties about the identity of the pH-sensing unit that changes protonation state at acidic pH, we performed independent atomistic simulations distinguished by the protonation states of selected amino acid residues located at sites thought important for pH sensing. To evaluate the response of the dynamic H-bond networks of GPR68 to changes in the protonation state, we used graph-based algorithms with which we identified dynamic water-mediated H-bond networks.

We found that, regardless of the protonation state of GPR68, extracellular His sidechains are part of a more extended bulk-exposed cluster that also contains several Asp and Glu sidechains (Figs. 5, 9). Such an extended water-mediated H-bond cluster of His and carboxylic sidechains is reminiscent of the proton-binding antennas thought to prolong the dwell time of the proton at the surface of proton-binding proteins [116–122]. In GPR68 with standard protonation states, His and carboxylic sidechains of the extracellular cluster inter-connect to the internal carboxylic triad via long-distance water-mediated H-bond paths (Figs. 6, 7, 9); though extended throughout much of the receptor, this internal protein-water H-bond network lacks connections to the DRY motif (Fig. 9A). Transient connections of the internal H-bond network and the DRY motif are sampled during simulations with distinct protonation states (Fig. 9B). That is, long-distance H-bond paths extending from the extracellular histidine-carboxyl cluster to the core of the protein are sampled during independent simulations (Fig. 9), suggesting robust connections between sites directly relevant to proton binding. The finding that the internal H-bond network responds to changes in protonation (Fig. 9) suggests conformational couplings that could contribute to the protonation-dependent conformational change of GPR68 along its activation path.

The participation of an extended protein-water H-bond network in the protonation-coupled conformational dynamics of GPR68, as we suggest here, could help explain why the pH sensitivity of GPR68 could not be completely abolished by single amino acid residue mutations at His sidechains, or even when several extracellular His sidechains are mutated simultaneously [13, 22, 112]: a large network with multiple titratable sites (Fig. 9) may remain to ensure robustness of the receptor function, even though moderately less efficiently than the wild-type protein.

The sequence identity among the six pH-sensing receptors is relatively low, ~11–38% (Figure S1, Table S2). Though structural elements that are used by each of these pH-sensing GPCRs are largely unclear, their sequences contain carboxylic and His sidechains that could change protonation at physiological pH: for example, GPR68 and GPR4 have 9–10 His groups each, whereas GPR132 has 3 His groups (Figure S26). At the region where E149$^{4.53}$ water bridges to E164^{ECL2}/E166^{ECL2}, three of the five pH-sensing GPCRs have 4–6 His and carboxylic groups within ECL2 (Table S1). Moreover, titratable sidechains of GPR68, including E149$^{4.53}$ and D282$^{7.49}$ may be mutated in cancer (Figure S27).

The functional implications of the similarities and differences in the content of titratable sidechains of pH-sensing GPCRs, and the cancer-associated mutations of titratable sidechains, are currently unclear, and the fact that distinct pH-sensing GPCRs have somewhat different pH sensing range suggests that the local electrostatic environment of the pH-sensing groups could be GPCR specific; likewise, we think that assigning to specific titratable sidechains cancer-related functionalities would be premature. In the future, methodological developments in the structural biology of pH-sensing GPCRs, spectroscopy measurements to identify putative proton-binding sites, and molecular simulations to probe the protonation-coupled receptor dynamics, will be needed for a complete view of the general physical-chemical principle used by pH-sensing GPCRs. In the future, the structural modeling presented here could provide the foundation for constant pH simulations, as reported recently for the opioid receptor [123], would be of interest to explore likely protonation states for titratable sidechains of GPR68, and simulations with polarizable force fields would be of interest to probe the water-mediated H-bond networks that inter-connect titratable sidechains of GPR68.

CRediT authorship contribution statement

Bhav Kapur: Conceptualization, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing. **Filippo Baldessari:** Investigation. **Michalis Lazaratos:** Validation; Visualization; Writing – review & editing. **Herbert Nar:** Project administration, Resources, Supervision, Funding acquisition, Writing – review & editing. **Gisela Schnapp:** Project administration, Resources, Supervision, Writing – review & editing. **Alejandro Giorgetti:** Methodology, Validation, Writing – review & editing. **Ana-Nicoleta Bondar:** Conceptualization, Methodology, Supervision, Funding acquisition; Writing – review & editing.

Declaration of Competing Interest

G.S. and H.N. are employees of Boehringer Ingelheim Pharma GmbH & Co. KG.

Acknowledgements

This work was supported by the European Union's Horizon 2020 research and innovation program (PROTON-ITN) under Marie Skłodowska Curie Actions Grant 860592 (to HN and A-NB). ML was supported in part by the DFG Collaborative Research Center SFB 1078. We thank the High-Performance Computing cluster based at Ingelheim am Rhein, Boehringer Ingelheim GmbH, for the computing time used to carry out the simulations.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2023.08.034.

References

[1] Sriram K, Insel PA. G Protein-Coupled Receptors as targets for approved drugs: how many targets and how many drugs? Mol Pharm 2018;93:251–8.
[2] Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. Annu Rev Pharmacol Toxicol 2013;53:531–56.
[3] Wacker D, Stevens RC, Roth BL. How ligands illuminate GPCR molecular pharmacology. Cell 2017;170:414–27.
[4] Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. Nature 2013;494:185–94.
[5] Latorraca NR, Venkatakrishnan AJ, Dror RO. GPCR dynamics: structures in motion. Chem Rev 2017;117:139–55.
[6] Wooten D, Chrisopoulos A, Marti-Solano M, Babu MM, Sexton PM. Mechanisms of signalling and biased agonism in G protein coupled receptors. Nat Rev Mol Cell Biol 2018;19:638–53.
[7] Weis WI, Kobilka BK. The molecular basis of G Protein-Coupled Receptor activation. Annu Rev Biochem 2018;87:897–919.
[8] Zhou Q, Yang D, Wu M, Guo Y, Guo W, Zhong L, et al. Common activation mechanisms of class A GPCRs. eLife 2019;8:e50279.
[9] Hauser A, Attwood MM, Rask-Andersen M, Schiöth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. Nat Rev Drug Discov 2017;16:829–42.
[10] Yang D, Zhou Q, Labroska V, Qin S, Darbalaei S, Wu Y-T, et al. G protein-coupled receptors: structure- and function-based drug discovery. Signal Transduct Target Ther 2021;6:7.
[11] Odoelman CS, Percival B, Wallis H, Chang M-W, Ahmad Z, Scholey D, et al. G-Protein coupled receptors: structure and function in drug discovery. RSC Adv 2020;10:36337.
[12] Silva PHI, Wagner CA. Physiological relevance of proton-activated GPCrs. Pflüg Arch - Eur J Physiol 2022;474:487–504.
[13] Ludwig M-G, Vanek M, Guerini D, Gasser JA, Jones CE, Junker U, et al. Proton-sensing G-protein-coupled receptors. Nature 2003;425:93–8.
[14] Sin WC, Zhang Y, Zhong W, Adhikarakunnathu S, Powers S, Hoey T, et al. G protein-coupled receptors GPR4 and TDAG8 are oncogenic and overexpressed in human cancers. Oncogene 2004;23:6299–303.
[15] Yu M, Cui R, Huang Y, Luo Y, Qin S, Zhong M. Increased proton-sensing receptor GPR4 signaling promotes colorectal progression by activating the hippo pathway. EBioMedicine 2019;48:264–76.
[16] Mashiko M, Kurosawa A, Tani Y, Tsuji T, Takeda S. GPR31 and GPR151 are activated under acidic conditions. J Biochem 2019;166.
[17] Fehrenbacher N, da Silva IT, Ramirez C, Zhou Y, Cho H-J, Kuchay S, et al. The G protein-coupled receptor GPR31 promotes membrane association of KRAS. J Cell Biol 2017;216:2329–38.
[18] Jiang B-C, Zhang W-W, Yang T, Guo C-Y, Cao D-L, Zhang Z-J, et al. Demethylation of G-protein coupled receptor 151 promoter facilitates the binding of Krüppel-like factor 5 and enhances neuropathic pain after nerve injury in mice. J Neurosci 2018;38:10535–51.
[19] Zohn D, Klinger M, Karp X, Kirk H, Symons M, Chrzanowska-Wodnicka M, et al. G2A is an oncogenic G protein-coupled receptor. Ocogene 2000;19:3866–77.
[20] Schönichen A, Webb BA, Jacobson MP, Barber DL. Considering protonation as a posttranslational modification regulating protein structure and function. Annu Rev Biophys 2013;42:289–314.
[21] Bondar A-N. Proton-binding motifs of membrame-bound proteins: from bacteriorhodopsin to spike protein S. Front Chem 2021;9:685761.
[22] Rowe JB, Kapolka NJ, Taghon GJ, Morgan WM, Isom DG. The evolution and mechanism of GPCR proton sensing. J Biol Chem 2020;296:1–13.
[23] Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G Protein-Coupled Receptors. Methods Neurosci 1995;25:366–428.
[24] Huang X-P, Karpiak J, Kroeze WK, Zhu H, Chen Y, Moy SS, et al. Allosteric ligands for the phrarmacologically dark receptors GPR68 and GPR65. Nature 2015;527:477–83.
[25] Huang X-P, Kenakin TP, Gu S, Shoichet BK, Roth BL. Differential roles of extracellular histidine residues of GPR68 for proton-sensing and allosteric modulation by divalent metal ions. Biochemistry 2020;59:3594–614.
[26] Wang J-Q, Kon J, Mogi C, Tobo M, Damirin A, Sato K, et al. TSAG8 is a proton-sensing and psychosine-sensitive G-protein-coupled receptors. J Biol Chem 2004;279:45626–33.
[27] Murakami N, Yokomizo T, Okuno T, Shimizu T. G2A is a proton-sensing G-protein coupled receptor antagonized by lysophosphatidylcholine. J Biol Chem 2004;279:42484–91.
[28] Radu CG, Nijagal A, McLaughlin J, Wang J, Witte ON. Differential proton sensitivity of related G protein-couled receptors T cell death-associated gene 8 and G2A expressed in immune cells. Proc Natl Acad Sci 2004;102:1632–7.
[29] Liu J-P, Nakakura T, Tomura H, Tobo M, Mogi C, Wang J-Q, et al. Each one of certain histidine residues in G-protein-coupled receptor GPR4 is critical for extracellular proton-induced stimulation of multuple G-protein-signaling pathways. Pharmacol Res 2010;61:499–505.
[30] Fahmy K, Jäger F, Beck M, Zvyaga TA, Sakmar TP, Siebert F. Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site-directed mutants. Proc Natl Acad Sci USA 1993;90:10206–10.
[31] Rath P, DeCaluwé LLJ, Bovee-Geurts PHM, DeGrip WJ, Rotschschild KJ. Fourier Transform Infrared Difference spectroscopy of rhodopsin mutants: light activation of rhodopsin causes hydrogen-bonding change in residue aspartic acid-83 during Meta II formation. Biochemistry 1993;32:10277–82.
[32] Vickery ON, Carvalheda CA, Zaidi SA, Pisliakov AV, Katritch V, Zachariae U. Intracellular transfer of Na^+ in an active-state G-Protein-Coupled Receptor. Structure 2018;26:171–80.
[33] Vanni S, Neri M, Tavernelli I, Rothlisberger U. A conserved protonation-induced switch can trigger "ionic-lock" formation in adrenergic receptors. J Mol Biol 2010;397:1339–49.
[34] Deupi X, Edwards P, Singhal A, Nickle B, Oprian D, Schertler G, et al. Stabilized G protein binding site in the structure of constitutively active metharhodopsin-II. Proc Natl Acad Sci 2011;109:119–24.
[35] Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung AY, Kobilka TS, et al. Crystal Structure of the β2 Adrenergic Receptor-Gs Protein Complex. Nature 2011;477:549–55.
[36] Deupi X, Edwards P, Singhal A, Nickle B, Oprian D, Schertler G, et al. Stabilized G protein binding site in the structure of constitutively active metharhodopsin-II. Proc Natl Acad Sci 2012;109:119–24.
[37] Olivella M, Caltabiano G, Cordomí A. The role of cysteine 6.47 in class A GPCRs. BMC Struct Biol 2013;13:3.
[38] Wacker D, Wang C, Katritch V, Han GW, Huang X-P, Vardy E, et al. Structural features for functional selectivity at serotonin receptors. Science 2013;340:615–9.
[39] Katritch V, Fenalti G, Abola EE, Roth BL, Cherezov V, Stevens RC. Allosteric sodium in class A GPCR signaling. Trends Biochem Sci 2014;39:233–44.
[40] Sanchez-Reyes OB, Cooke ALG, Tranter DB, Rashid D, Eilers M, Reeves PJ, et al. G Protein-coupled receptors contain two conserved packing clusters. Biophys J 2017;112:2315–26.
[41] Chelikani P, Hornak V, Eilers M, Reeves PJ, Smith D, RajBhandary UL, et al. Role of group-conserved residues in the helical core of β2-adrenergic receptor. PNAS 2007;104:7027–32.
[42] Fritze O, Filipek S, Kuksa V, Palczewski K, Hoffmann KP, Ernst OP. Role of the conserved $NPxxY(x)_{5,6}F$ motif in the rhodopsin ground state and during activation. PNAS 2003;100:2290–5.
[43] Bertalan É, Lešnik S, Bren U, Bondar A-N. Protein-water hydrogen-bond networks of G protein-coupled receptors: Graph-based analyses of static structures and molecular dynamics. J Struct Biol 2020;212:107634.
[44] del Val C, Bondar L, Bondar A-N. Coupling between inter-helical hydrogen bonding and water dynamics in a proton transporter. J Struct Biol 2014;186:95–111.
[45] Kandori H. Role of internal water molecules in bacteriorhodopsin. Biochim Biophys Acta 2004;1460:177–91.
[46] Kouyama T, Nishikawa T, Tokuhisa T, Okumura H. Crystal structure of the L intermediate of bacteriorhodopsin: evidence for vertical translocation of a water molecule during the proton pumping cycle. J Mol Biol 2004;335:531–46.
[47] Bondar A-N, Smith JC. Protonation-state coupled conformational dynamics in reaction mechanisms of channel and pump rhodopsins. Photochem Photobiol 2017;93:1336–44.
[48] Bondar A-N, Alfonso-Priesto M. Hydrogen-bond networks for proton couplings in G-protein coupled receptors. Front Phys 2022;10:963716.
[49] Venkatakrishnan AJ, Ma AK, Fonseca R, Latorraca NR, Kelly B, Betz RM, et al. Diverse GPCRs exhibit conserved water networks for stabilization and activation. Proc Natl Acad Sci 2019;116:3288–93.
[50] Murakami M, Kouyama T. Crystal structure of squid rhodopsin. Nature 2008;453:363–7.
[51] Fried SDE, Hewage KSK, Eitel AR, Struts AV, Weerasinghe N, Perera SMDC, et al. Hydratio-mediated G-protein-coupled receptor activation. PNAS 2022;119:e2117349119.
[52] Lesnik S, Bren U, Domratcheva T, Bondar A-N. Fentanyl and the fluorinated fentanyl derivative NFEPP elicit distinct hydrogen-bond dynamics of the opioid receptor. J Chem Inf Model 2023;63:4732–48.
[53] Sandal M, Duy TP, Cona M, Zung H, Carloni P, Musiani F, et al. GoMoDo: A GPCR online modeling and docking webserver. PloS One 2013;8:e74092.

[54] Ribeiro RP, Giorgetti A. pyGoMoDo: GPCRs modeling and docking with Python. In: Cowen L, editor. Bioinformatics. Oxford University Press, Place Published; 2023.
[55] Söding J. Protein homology detection by HMM-HMM comparison. Bioinformatics 2005;21:951–60.
[56] Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction, Nucleic Acid. Research 2005;33: W244–8.
[57] Remmert M, Biegert A, Hauser A, Söding J. HHblits: lightning-fast iterative protein sequence by HMM-HMM alignment. Nat Methods 2012;9:173–5.
[58] Kooistra AJ, Mordalski S, Pándy-Szekeres P, Esguerra M, Mamyrbekov A, Munk C, et al. GPCRdb in 2021: integrating GPCR sequence, structure and function. Nucl Acid Res 2021;49:D335–43.
[59] Wheatley M, Wootten D, Conner MT, Simms J, Kendrick R, Logan RT, et al. Lifting the lid on GPCRs: the role of extracellular loops. Br J Pharmacol 2012;165: 1688–703.
[60] Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, et al. Comparative structure modeling with MODELLER. Curr Protoc Bioinforma 2006; 15:1–30.
[61] Webb B, Sali A. Comparative protein structure modeling using MODELLER. Curr Protoc Bioinforma 2016;54.
[62] Shen J, Sali A. Statistical potential for assessment and prediction of protein structures. Prot Sci 2006;15:2507–24.
[63] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596:583–9.
[64] Ovcinnikov YA, Abdulaev NG, Bogachuk AS. Two adjacent cysteine residues in the C-terminal cytoplasmic fragment of bovine rhodopsin are palmitylated. FEBS J 1988;230:1–5.
[65] O'Dowd BF, Hnatowich M, Caron MG, Leftkowitz RJ, Bouvier M Palmitoylation of the human β2-adrenergic receptor Mutation of Cys341 in the carboxyl tail leads to an uncoupled nonpalmytoylated form of the recepto. J Biol Chem 1989;264: 7564–9.
[66] Rose Y, Duarte JM, Lowe R, Segura J, Bi C, Bhikadiya C, et al. RCSB Protein Data Bank: Architectural advances towards integrated searching and efficient access to macromolecular structure data from the PDB archive. J Mol Biol 2021;433: 166704.
[67] Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, et al. Crystal structure of rhodopsin: a G protein-coupled receptor. Science 2000;289: 739–45.
[68] Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 2008;29:1859–65.
[69] Community TG. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses. Nucleic Acid Res 2022;50:W345–51.
[70] Isberg V, de Graaf C, Bortolato A, Cherezov V, Katrich V, Marshall FH, et al. Generic GPCR residue numbers - aligning topology maps minding the gaps. Trends Pharmacol Sci 2015;36:22–31.
[71] Clamp M, Cuff J, Searle SM, Barton GJ. The Jalview Java alignment editor. Bioinformatics 2004;20:426–7.
[72] Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics 2009;25:1189–91.
[73] Crooks GE, Hon G, Chadonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res 2004;14:1188–90.
[74] Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences, Nucleic Acid. Research 1990;18:6097–100.
[75] UniProt, UniProt: a hub for protein information, Nucleic Acid Research, 43 (2015) D204-D212.
[76] Lomize M, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM server: resources for positioning of proteins in membranes, Nucleic Acid. Research 2011;40:D370–6.
[77] Wu EL, Cheng X, Jo S, Rui H, Song KC, Dávila-Contreras EM, et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. J Comput Chem 2014;35:1997–2004.
[78] Subczynski WK, Pasenkiewicz-Gierula M, Widomska J, Mainali L, Raguz M. High cholesterol/low cholesterol: effects on biological membranes: a review. Cell Biochem Biophys 2017;75:369–85.
[79] Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983;4:187–217.
[80] MacKerell Jr AD, Bashford D, Bellot M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–616.
[81] Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J Phys Chem B 2010;114:7830–43.
[82] Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods 2016;40:71–3.
[83] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparisonofsimple potential functions for simulations of liquid water. J Chem Phys 1983;79:926–35.
[84] Kalé L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, et al. NAMD2: Greater scalability for parallel molecular dynamics. J Comput Phys 1999;151: 283–312.
[85] Feller SE, Zhang Y, Pastor RW, Brooks B. Constant pressure molecular dynamics simulation: the Langevin piston method. J Chem Phys 1995;103:4613–21.
[86] Martyna GJ, Tobias DJ, Klein ML. Constant-pressure molecular-dynamics algorithms. J Chem Phys 1994;101:4177–89.
[87] Darden T, York D, Pedersen L. Particle mesh Ewald: an N x log(N) method for Ewald sums in large systems. J Chem Phys 1993;98:10089–92.
[88] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103:8577–93.
[89] Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints. Molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–41.
[90] Siemers M, Lazaratos M, Karathanou K, Guerra F, Brown LS, Bondar A-N. Bridge: a graph-based algorithm to analyze dynamic H-bond networks in membrane proteins. J Chem Theory Comput 2019;15:6781–98.
[91] Siemers M, Bondar A-N. Interactive interface for graph-based analyses of dynamic H-bond networks: application to spike protein S. J Chem Inf Model 2021;61: 2998–3014.
[92] Bondar A-N. Graphs of hydrogen-bond networks to dssect protein conformational dynamics. J Phys Chem B 2022;126:3973–84.
[93] Karathanou K, Lazaratos M, Bertalan E, Siemers M, Buzar K, Schertler GFX, et al. A graph-based approach identifies dynamic H-bond communication networks in spike protein S of SARS-CoV-2. J Struct Biol 2020;212:107617.
[94] Cormen TH, Leiserson CE, Rivest RL, Sten C. Introduction to algorithms. third ed. Massachusetts Institute of Technology,; 2009.
[95] Bondar A-N. Interplay between local protein interactions and water bridging of a proton antenna carboxylate cluster. BBA - Biomembranes 2022;1864:184052.
[96] Dalton JAR, Lans I, Giraldo J. Quantfying conformational changes in GPCRs: glimpse of a common functional mechanism. BMC Bioinformatics 2015;16:124.
[97] Humphrey W, Dalke W, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–8.
[98] White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, et al. Structure of the agonist-bound neurotensin receptor. Nature 2012;490:508–13.
[99] Zhang C, Srinivasan Y, Arlow DH, Fung JJ, Palmer DJ, Zheng Y, et al. High-resolution crystal structure of human protease-activated receptor 1. Nature 2012; 492:387–92.
[100] Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antaganists. Science 2010;330:1066–71.
[101] Hauser AS, Kooistra AJ, Munk C, Heydereich FM, Veprintsev DB, Bouvier M, et al. GPCR activation mechanisms across classes and macro-microscales. Nat Struct Mol Biol 2021;28:879–88.
[102] Do HN, Haldane A, Levy RM, Miao Y. Unique features of different classes of G-protein-coupled receptors revealed from sequence coevolutionary and structural analysis. Proteins Struct Funct Bioinf. 2022;90:601–14.
[103] Hedderich JB, Persechino M, Becker K, Heydenreich FM, Gutermuth T, Bouvier M, et al. The pocketome of G-protein-coupled receptors reveals previously untargeted allosteric sites. Nat Comm 2022;13:2567.
[104] Hulme EC. GPCR activation: a mutagenic spotlight on crystal structures. TRENDS Pharmacol Sci 2013;34:67–84.
[105] Rosenbaum DM, Rasmussen SGF, Kobilka BK. The structure and function of G-protein-coupled receptors. Nature 2009;459:356–63.
[106] Zhang XC, Cao C, Zhou Y, Zhao Y. Proton transfer-mediated GPCR activation. Protein Cell 2015;6:12–7.
[107] Gimpl G, Burger K, Fahrenholz F. Cholesterol as modulator of receptor function. Biochemistry 1997;36:10959–74.
[108] Angladon M-A, Fosségré M, Vercauteren DP. Interactions of POPC, DPPC, and POPE with the μ opioid receptor: a coarse-grained molecular dynamics study. PLoS One 2019;14:e0123646.
[109] Bhattarai A, Wang J, Miao Y. G-Protein-Coupled Receptor-membrane interactions depend on the receptor activation state. J Comput Chem 2020;41:460–71.
[110] Vohra S, Taddese B, Conner AC, Poyner DR, Hay DL, Barwell J, et al. Similarity between class A and class B G-protein-coupled receptors exemplified through calcitonin gene-related peptide receptor modeling and mutagenesis studies. J R Soc Interface 2013;10:20120846.
[111] Escribá PV, Ozaita A, Ribas C, Miralles A, Fodor E, Farkkas T, et al. Role of lipid polymorphism in G protein-membrane interactions: Nonlamellar-prone phospholipids and peripheral protein binding to membranes. Proc Natl Acad Sci USA 1997;94:11375–80.
[112] Huang X-P, Karpiak J, Kroetze W, Zhu H, Chen X, Moy SS, et al. Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65. Nature 2015;527: 477–83.
[113] Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn AL, et al. Structural insights into μ-opioid receptor activation. Nature 2015;524:315–21.
[114] Gerweck LE, Seetheraman K. Cellular pH gradient in tumor versus normal tissue: potential exploitation for the treatment of cancer. Cancer Res 1996;56:1194–8.
[115] leadXpro, leadXpro has just elucidated the world's first structure of a proton-sensing GPCR target, 2022.
[116] Checover S, Marantz Y, Nachliel E, Gutman M, Pfeifer M, Tittor J, et al. Dynamics of the proton transfer reaction on the cytoplasmic surface of bacteriorhodopsin. Biochemistry 2001;40:4281–92.
[117] Sacks V, Marantz Y, Aagaard A, Checover S, Nachliel E, Gutman M. The dynamic feature of the proton collecting antenna of a protein surface. Biochim Biophys Acta 1998;1365:232–40.
[118] Ädelroth P, Brzezinski P. Surface-mediated proton-transfer reactions in membrane-bound proteins. Biochim Biophys Acta 1655;2004:102–15.

[119] Lorch S, Capponi S, Pieront F, Bondar A-N. Dynamic carboxylate/water networks on the surface of the PsbO subunit of photosystem II. J Phys Chem B 2015;119:12172–81.
[120] Bondar A-N. Mechanisms of long-distance allosteric couplings in proton-binding membrane transporters. Adv Protein Chem Struct Biol 2022;128:199–239.
[121] Shutova T, Klimov VV, Andersson B, Samuelsson G. PsbO protein is involved in proton transfer from the water oxidizing complex of Photosystem II (A cluster of carboxylic groups in) Biochim Biophys Acta 1767;2007:434–40.
[122] Gerland L, Friedrich D, Hopf L, Donovan EJ, Wallmann A, Erdmann N, et al. pH-dependent protonation of surface carboxylates in PsbO enables local buffering and triggers structural changes. ChemBioChem 2020;21:1597–604.
[123] Vo QN, Mahinthichaichan P, Shen J, Ellis CR. How μ-opioid receptor recognizes fentanyl. Nat Comm 2021;12:984.