scientific reports

Check for updates

OPEN

Atomic insights into the signaling landscape of *E. coli* PhoQ histidine kinase from molecular dynamics simulations

Symela Lazaridi[1,2], Jing Yuan[3] & Thomas Lemmin[1]✉

Bacteria rely on two-component systems to sense environmental cues and regulate gene expression for adaptation. The PhoQ/PhoP system exemplifies this crucial role, playing a key part in sensing magnesium (Mg^{2+}) levels, antimicrobial peptides, mild acidic pH, osmotic upshift, and long-chain unsaturated fatty acids, promoting virulence in certain bacterial species. However, the precise details of PhoQ activation remain elusive. To elucidate PhoQ's signaling mechanism at atomic resolution, we combined AlphaFold2 predictions with molecular modeling and carried out extensive Molecular Dynamics (MD) simulations. Our MD simulations revealed three distinct PhoQ conformations that were validated by experimental data. Notably, one conformation was characterized by Mg^{2+} bridging the acidic patch in the sensor domain to the membrane, potentially representing a repressed state. Furthermore, the high hydration observed in a putative intermediate state lends support to the hypothesis of water-mediated conformational changes during PhoQ signaling. Our findings not only revealed specific conformations within the PhoQ signaling pathway, but also hold significant promise for understanding the broader histidine kinase family due to their shared structural features. Our approach paves the way for a more comprehensive understanding of histidine kinase signaling mechanisms across various bacterial species and opens the door for developing novel therapeutics that target PhoQ modulation.

Many microorganisms, encompassing bacteria, archaea, and fungi exhibit the remarkable ability to thrive in diverse and often challenging environments[1,2]. They have evolved intricate regulatory systems to sense and respond to changes in their surroundings[3]. Among these regulatory mechanisms, two-component systems (TCSs) stand as fundamental complexes that orchestrate cellular adaptation[4,5]. TCSs are composed of two distinct units: the histidine kinase (HK) and the response regulator (RR) proteins[6,7]. HK acts as a sensor detecting a large spectrum of environmental stimuli, ranging from shifts in temperature or pH, to availability of nutrients or detection of antibiotics[8–10]. Upon activation, HK undergoes autophosphorylation at a highly conserved histidine residue, marking the onset of a sophisticated phosphorylation relay. This initiates a subsequent interaction between a HK and its cognate RR protein, leading to the transfer of the phosphoryl group typically onto an aspartate residue in the receiver domain of the RR. Finally, the phosphorylated RR dimerizes and binds to specific DNA sequences, regulating gene expression in response to the environmental signal[11–13]. The absence of TCSs from mammals makes them enticing targets for drug development[14].

PhoQ functions as the histidine kinase within the PhoQ/PhoP two-component system, notably abundant among gram-negative bacteria[15,16]. Activation of PhoQ is initiated by a variety of environmental changes[17,18], including a low concentration of Mg^{2+}, acidic pH[19,20], the presence of cationic antimicrobial peptides[21–23], osmotic upshift[24], and long-chain unsaturated fatty acids in bile[25]. PhoP directly controls the transcription of an expansive array of genes, coordinating responses to environmental stress, alterations in the bacterial envelope's composition, and modulating virulence[26,27]. Structurally, PhoQ assembles into a transmembrane dimer with a molecular weight of about 110 kDa. Its configurations closely resemble the architecture of prototypical histidine kinases, comprised of five distinct domains: (i) a periplasmic sensor domain, (ii) a transmembrane (TM) domain linking the sensor domain to the cytosolic domains, which include (iii) the histidine kinases, adenylate cyclases, methyl-accepting

[1]Institute of Biochemistry and Molecular Medicine (IBMM), University of Bern, Bühlstrasse 28, 3012 Bern, Switzerland. [2]Graduate School for Cellular and Biomedical Sciences (GCB), University of Bern, Mittelstrasse 43, 3012 Bern, Switzerland. [3]Max Planck Institute for Terrestrial Microbiology and Center for Synthetic Microbiology, Karl-Von-Frisch-Strasse 14, 35043 Marburg, Germany. ✉email: thomas.lemmin@unibe.ch

chemotaxis proteins, and phosphatases (HAMP) domain[28], (iv) the dimerization and histidine phosphotransfer (DHp) domain, and (v) the catalytic domain, also known as the ATP binding domain (Fig. 1A)[29].

Several domains of PhoQ or homologous HKs have been experimentally characterized, providing valuable structural insights into components of the signaling mechanism. The sensor domain structure has been determined in both its monomeric and various oligomeric states[30–33]. Through a combination of disulfide screen experiments and computational modeling, the dimeric form observed for *E. coli* (PDB id: 3bq8[33]) has been suggested as the most likely representative of PhoQ's physiologically relevant active conformation[34]. The asymmetry observed in this dimeric structure is believed to be essential for signaling, since autophosphorylation selectively takes place in one of the two subunits under physiological conditions[35,36]. A proposed mechanism suggests that upon sensing stimuli, the sensor domain initiates structural rearrangements in the transmembrane domain, which are then propagated to the cytoplasmic domains[35].

In the absence of a high-resolution experimental structure for the TM domain, the phototaxis sensory rhodopsin II-transducer (HtrII, PDB id: 1h2s) complex[37] frequently serves as a model for HK transmembrane four-helical bundles[35,38]. The PhoQ transmembrane domain harbors at its center a critical polar residue (N202) that likely forms a water cavity, similar to that observed in HtrII[39]. This cavity is believed to facilitate the arrangement of side chains around it, enabling the TM helix to bend and change conformation during signaling[35,36]. This hypothesis is supported by ligand-dependent disulfide crosslinking experiments and Bayesian modeling, which identified two distinct structural states differing in the diagonal displacement of the TM helices[40,41].

The HAMP[42] and DHp in the cytoplasm are well-conserved signal-transducing domains that have been extensively characterized experimentally in isolation or in combination with other cytoplasmic domains[43–46]. Various signaling mechanisms have been proposed for the HAMP domain, including the gearbox[47] or the diagonal scissoring models[48]. The signaling initiated by the HAMP domain undergoes a transformation into an asymmetric conformation of the DHp bundle, mediated by a small connecting helix (S-helix). This distortion in the helical bundles of DHp facilitates the release of the catalytic domain, ultimately leading to the autophosphorylation of the conserved histidine[35,43,46].

Although previous studies have yielded crucial insights into the signal mechanism of PhoQ, a holistic understanding remains elusive due to the lack of comprehensive structural information for all the involved constituents. In this study, we therefore leveraged Alphafold2-multimer[49] to produce an atomic model of the complete PhoQ dimer. The predicted model exhibited strong congruence with existing structural data. To delve deeper into the intricate signaling mechanisms governing PhoQ, we combined classical molecular modeling and dynamics simulations and were able to capture two distinct conformations. These models were supported by available experimental data and could be potentially representative of an active and inactive conformation of PhoQ. Collectively, our findings provide new insights into signaling intricacies underpinning PhoQ, and our approach could be applicable to deciphering the signaling mechanisms of other histidine kinases.

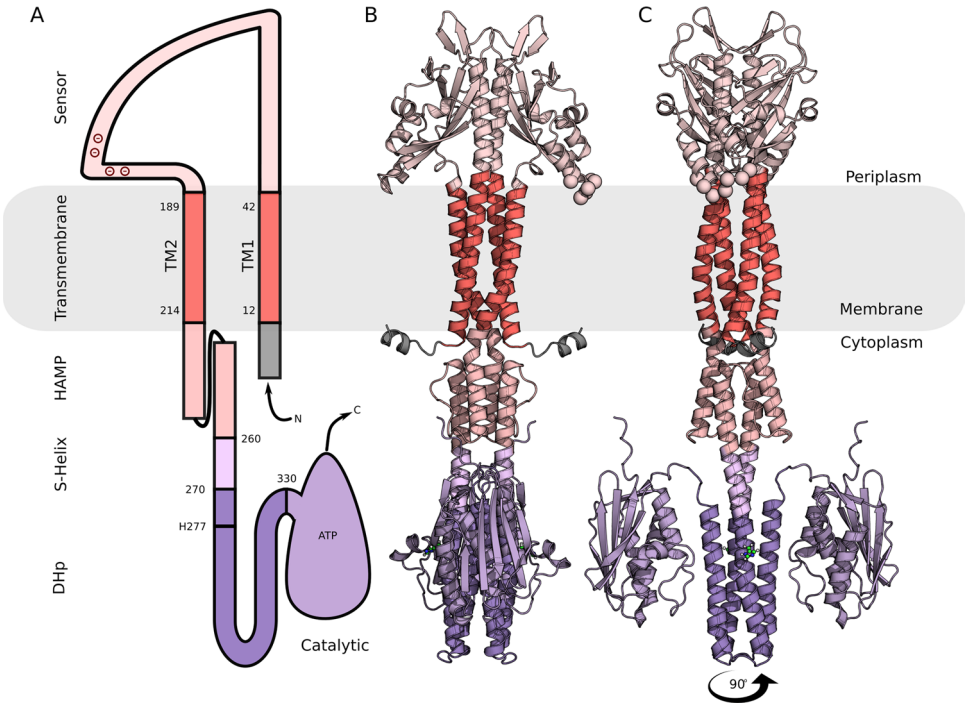

Figure 1. Predicted Structure of E. coli PhoQ. (**A**) Schematic representation of PhoQ topology. (**B**) Cartoon representation of the predicted PhoQ structure using AlphaFold2. Transmembrane domains (red), sensor and HAMP domains (shades of pink), and DHp and catalytic domains (shades of purple) are indicated. The membrane is depicted as a gray rectangle.

Results

Structural characterization of AlphaFold prediction

AlphaFold2-multimer predicted with high confidence the full-length dimeric structure of *E. coli* PhoQ, achieving an average predicted local distance difference test (pLDDT) score of 0.8 [49]. (Supplementary Figure S1A). The generated model formed a symmetrical dimer, with a root mean square deviation (RMSD) of only 0.75 Å between the two monomers. The dimeric interface was characterized by a continuous helix (Fig. 1B,C). Notably, there were marked drops in the pLDDT score in regions aligned well with the expected boundaries between the different domains of PhoQ (Figure S1B). These discernible drops were leveraged to demarcate the start and end points of each domain for subsequent analyses.

The AlphaFold prediction for the sensory Per-Arnt-Sim (PAS) domain[50] shows high agreement with the known crystal structure of *E. coli* (PDB ID: 3bq8, RMSD: 0.6 Å for the monomer[33]). However, significant deviations emerged in the predicted dimer (RMSD: 5.25 Å, Supplementary Figure S2). The interfacial helices in the model formed a shallower angle (10°) compared to the experimental structure (41°), impacting the orientation of the critical acidic patch (residues D136-D150)[51]. To quantify this difference, we embedded the Alphafold model into a lipid bilayer, using the orientations of proteins in membranes (OPM) server[52], and measured angles of 50.21° and 45.25° between the membrane surface and the acidic patch helices for chains A and B, respectively. To compare the orientations of the acidic patches in the experimental structure, we aligned the sensor interfaces (residues T48-K64) with the oriented Alphafold model. The angle for chain A remained comparable (47.02°), but was more parallel for chain B (25.85°), due to the inherent asymmetry observed in the experimental structure. Additionally, the crucial salt bridge between D179 and R50'[33] present in the experimental structure was absent in the prediction, despite a comparable Cα-Cα distance (approximately 11.4 Å for both).

The transmembrane domain of PhoQ formed a four-helical bundle with antiparallel helices per monomer, defined as TM1 (S12—V42) and TM2 (Y189—A214). This structural arrangement is strongly corroborated by prior research[33,34,53]. Both TM1-TM1' and TM2-TM2' assembled in coiled coils. To assess potential disruptions in the coiled-coil packing, we calculated the accommodation index[54]. The obtained score of 1.0 suggests the presence of single-residue insertions within the coiled-coils, at positions V25 (TM1) and S200 (TM2). To assess the novelty of this fold, we employed the structural search tool FoldSeek to query the Protein Data Bank (PDB) (TableS1)[55]. The four-helix bundle aligned with various structures; however few were transmembrane proteins (best PDB ID: 6sss_C, TM-score: 0.46, RMSD: 3.90 Å). This probably reflects the scarcity of such structures in the database. In contrast, comparison with the commonly used HtrII transmembrane bundle[38] using the TM-align mode[56] (RCSB.org[57]), yielded a significantly higher TM-score of 0.55 and a reduced RMSD of 2.56 Å. Additionally, we compared the Alphafold model to a previously published atomic model of the transmembrane domain[41], which has been validated experimentally[58,59] (Supplementary Figure S3A, B). Despite some variations (increased TM1 separation and closer TM2 helices in the published model, RMSD: 3.8 Å, TM-score: 0.56), the overall folds aligned well, enabling a seamless connection between the sensor interface helices and the HAMP domain (Supplementary Figure S3C).

The predicted HAMP domain adopted the expected two-antiparallel helix configuration, connected by a 10-residue loop. The FoldSeek search against the PDB database identified eight HAMP chains from sensor proteins among the top hits but with low structural similarity (TM-Score: 0.6). The key difference resided in the interhelical angle. We further quantified this variability by comparing the modelled HAMP to experimentally solved HAMP domains[60]. This revealed a wide distribution of angles, suggesting a lack of a single preferred conformation for the HAMP domain in these related sensor proteins (Supplementary Figure S4).

The predicted DHp domain consisted of two alpha helices (a1 and a2) connected by a hairpin loop, forming a left-handed four-helix bundle in the dimer. This structure closely resembles other experimentally determined DHp domains (best TM-score: 0.84). Notably, the conserved histidine residue (H277), critical for the phosphorylation cascade, resides within the a1 helix with its side chain exposed to the cytoplasm. Moreover, the handedness of the bundle places the intra-protomeric catalytic domain in close proximity to this conserved histidine, strongly suggesting a cis-autophosphorylation mechanism[61].

The catalytic domain adopted the expected Bergerat fold and aligns, closely matching the experimentally solved structure of *E. coli* (PDB ID: 1id0, RMSD: 1.7 Å)[62]. However, a key structural difference resided in the loop forming the ATP-lid, which connects helix α3 and beta-strand β3. In the lid0 structure, the ATP-lid extended further, exposing the ligand to the solvent (Supplementary Figure S5). Although Alphafold was not trained with ligands, it has been shown that the configuration of binding pockets can often accommodate them[63]. We successfully positioned the ATP molecule and Mg^{2+} cation from the experimental structure (1id0) into the predicted binding pocket by aligning the main helices of the catalytic domains. In both structures, the Mg^{2+} cation was chelated by three oxygens of the triphosphate head and an additional interaction with N388 or N389, N385, and Q442 in the predicted and experimental structures, respectively. Conserved interactions were observed between K392 and N389 with the triphosphate head in both structures. However, the predicted structure lacks interactions with the ATP-lid residues R434 and R439. Furthermore, the adenosine moiety of the ATP engaged in pi-pi interactions with Y393 and formed a hydrogen bond with D415 in both structures.

Structural refinement with all-atom MD simulations

In order to further refine the predicted structure of the full-length PhoQ dimer, we carried out a set of all-atom molecular dynamics simulations. The AlphaFold model ($PhoQ_{AF}$) was inserted into a bacterial mimic membrane bilayer and three replicas of all-atom MD simulations were conducted. These simulations were performed under near-physiological conditions (310 K and 1 atm pressure), with each replica extending up to approximately 1.4 microseconds (Supplementary Table S2). The models remained stable throughout the simulations (average RMSD: 8.76 ± 2.57 Å, Supplementary Figure S6). The major difference was associated with the rearrangement

of the catalytic domains, moving slightly away from the DHp. While the protein's core exhibited only marginal deviation from the predicted model (Fig. 2A), one replica exhibited a notable deviation. In this replica, a subtle change in the transmembrane (TM) helix bundle of this replica, bringing TM1' and TM2 helices closer and reducing the kink at P208, caused the DHp to move away from the core axis leading to an increase in the protein core RMSD (Supplementary Figure S7B). Interestingly, the catalytic domain was also released in this replica. The cross angle between the helices of the sensor interface remained consistently stable, averaging 7.38° ± 3.76° (Supplementary Figure S9). During the simulations, we also observed the dynamic formation of a salt bridge between D179 and R50' (42% of simulation time with a cut-off distance of 3.6 Å) and R50 and D179' (22% of simulation time).

We attempted to refine the AlphaFold model's sensor interface using targeted molecular dynamics (TMD) simulations ($PhoQ_{AF}^{TMD}$) to match the conformation observed in the experimental dimer structure (PDB ID: 3bq8). However, this introduced significant structural distortions (RMSD of full protein: 10.32 ± 3.76 Å and core: 5.19 ± 2.03 Å). In particular, the helix connecting the sensor to TM1 partially unfolded (residues D45 to T47), potentially leading to the decoupling of the sensor domain from the transmembrane region (Supplementary Figure S8B). We concluded that the TMD simulation employed was insufficient to induce the desired conformational alteration in the AlphaFold model, suggesting a more intricate structural coupling between the sensor and transmembrane domains that may hinder the desired conformational change in the model.

The high degree of structural similarity between the sensor-TM1 interface (I30-S43) and the TM2-HAMP interface (R219-A226) in the previously modeled transmembrane domain[41] and $PhoQ_{AF}$ allowed the replacement of the predicted TM structure with the modeled one (Supplementary Figure S3). We thus constructed a hybrid model ($PhoQ_H$) by combining the modeled transmembrane domain with the AlphaFold-predicted sensor and cytoplasmic domains (Supplementary Figure S8C). We then performed all-atom MD simulations in triplicate to refine the structure and assess its stability. The core of PhoQ remained stable throughout the simulations (4.16 ± 0.47 Å, Supplementary Figure S6, Supplementary Table S2) and the sensor interfacial helices maintained a parallel alignment (cross angle: 13.10 ± 4.00°, Supplementary Figure S9). The simulations also captured the release of one of the catalytic domains.

Encouraged by the stability of the $PhoQ_H$, we conducted a TMD simulation ($PhoQ_H^{TMD}$) to induce an interface angle resembling the experimental dimer (PDB ID: 3bq8). Even though the structure remained stable during the subsequent 1.5 µs unrestrained MD simulations, conducted in three replicas, (RMSD of 7.46 ± 3.24 Å, Supplementary Figure S6 and Supplementary Table S2), the resulting conformation deviated significantly from $PhoQ_{AF}$ (Fig. 2B and Supplementary Figure S8D). The equilibrated conformation of $PhoQ_H^{TMD}$ was characterized by a highly asymmetric and curved structure, with several kinks along the core. Notably, the sensor domain tilted significantly (49.12° ± 5.11°) relative to the membrane compared to its perpendicular orientation in $PhoQ_{AF}$. Furthermore, a kink at position F44 in chain B allowed the sensor interface to retain a cross-angle more closely aligned with the experimental structure, approximately 27.22° ± 4.63°. This conformation was similar to the experimental structure (PDB ID: 3bq8, RMSD: 4.07 ± 0.44 Å). The interface helices pivoted around E55, while the E55 (chain A) and R53 (chain B) formed a distance of 3.7 Å for 33% of the simulation time, agreeing with the experimental structure. The increased cross-angle led to a near vertical alignment of the acidic patch helix on chain B (~ 71.2°), bringing the acidic cluster close to the membrane surface. In contrast, the acidic patch helix of the other protomer aligned almost parallel to the membrane surface, forming an angle of approximately 7°. Although the conformations of the HAMP and DHp domains remained very similar to $PhoQ_{AF}$ (1.82 ± 0.30 Å RMSD), the catalytic domain of chain A was released and pivoted towards the DHp region.

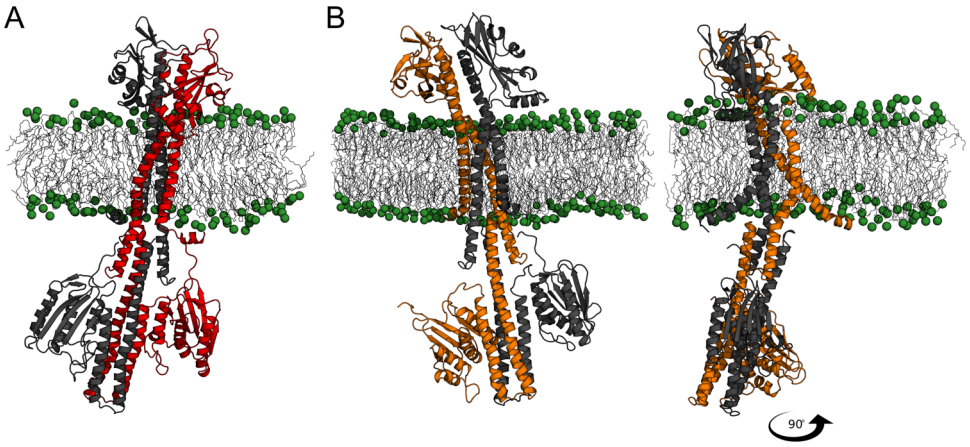

Figure 2. Representative structural models of PhoQ after MD simulation refinement. Cartoon representations of (**A**) $PhoQ_{AF}$ and (**B**) $PhoQ_H^{TMD}$ at the end of the MD simulation. The different chains are colored in gray and red or orange, respectively. The hydrophobic core of the membrane is depicted with gray sticks, and phosphate groups are shown as green spheres.

Conformational change of the transmembrane bundle

To understand how the altered transmembrane domain between $PhoQ_{AF}$ and $PhoQ_H^{TMD}$ affected the protein structure, we investigated structural deviations propagated along this domain. We divided the transmembrane domain into three slices, each defined by two alpha carbons per protomer and oriented parallel to the membrane (Fig. 3 and Supplementary Figure S10). The top slice (closest to the sensor domain) comprised residues A36 and W194, the middle slice (center of the membrane) residues L28 and N202, and the bottom slice (closest to the HAMP domain) residues T21 and L210. Compared to $PhoQ_{AF}$, $PhoQ_H^{TMD}$ exhibited consistently larger TM1-TM1' distances and smaller TM2-TM2' distances. This resulted in TM1-TM1' helices forming a cone shape, narrowing towards the sensor domain, whereas TM2-TM2' helices adopted an inverted cone shape. Interestingly, despite the overall asymmetry of $PhoQ_H^{TMD}$, the arrangement of residues within each slice displayed a more symmetrical, diamond-like shape compared to the distorted appearance in $PhoQ_{AF}$ (Figs. 3 and Supplementary Figure S10). The configuration of $PhoQ_H^{TMD}$ was accompanied by a bent at I207 of approximately ~ 3.5 degrees proximal to P208 in chain A (compared to ~ 6.8 degrees in $PhoQ_{AF}$), and a pistonning movement between transmembrane helices (Supplementary Figure S11). A partial unfolding on the connecting region between TM2 and the sensor domain (I181 – Y189) was also observed in $PhoQ_H^{TMD}$. In both models, the amphipathic N-terminal helices of TM1 (M1 to P8) were folded back and bound to the membrane surface.

Comparison to experimental cross-linking data

When compared to available experimental cross-linking data for the sensor interface, transmembrane (TM), and HAMP domains[40], all structural models exhibited generally good agreement between the cross-linking fractions and average inter-protomeric Cα-Cα distances (Fig. 4). However, some key discrepancies emerged. Notably, the absence of cross-links for residues A20-L30 and the decreased periodicity for residues S43-D45 align with the larger interatomic distances and kink in the TM1 helix as it exits the membrane observed within the $PhoQ_H^{TMD}$ model.

Solvation of the transmembrane bundle

Previous experiments have shown that the polarity of N202 is essential for the function of PhoQ[39,41]. This residue, located at the center of the transmembrane bundle, is thought to be crucial for solvating the bundle and

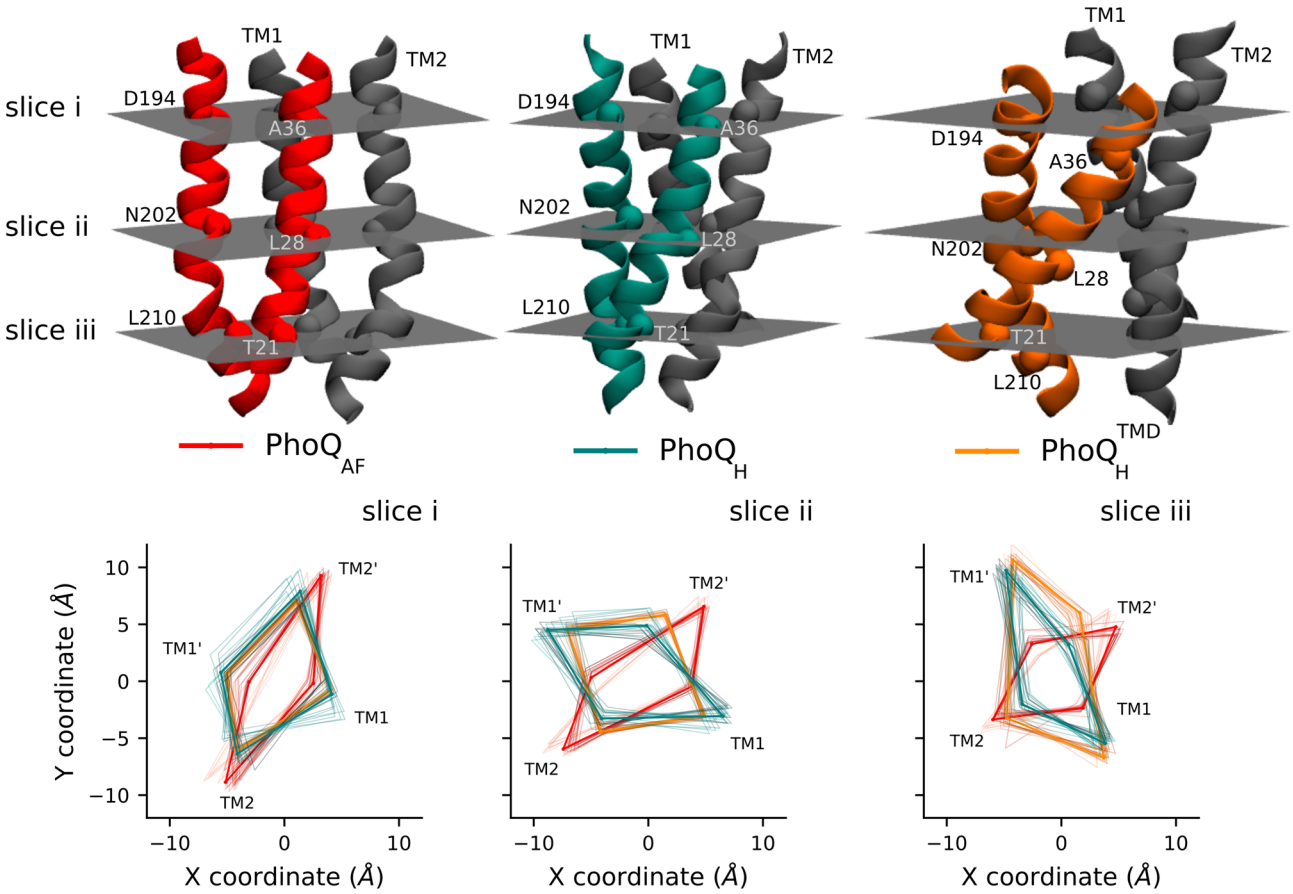

Figure 3. Interprotomer Cα distances within TM helices. Top panel: Cartoon representation of the PhoQ transmembrane (TM) helices divided into three slices (**i**–**iii**) defined by Cα atoms: (**i**) D194 and A36, (**ii**) N202 and L28, (**iii**) L210 and T21. Bottom panel: Distances measured during the MD simulations between Cα atoms within each slice for the $PhoQ_{AF}$ (red), $PhoQ_H$ (teal), and $PhoQ_H^{TMD}$ (orange) models, sampled every 100 ns.

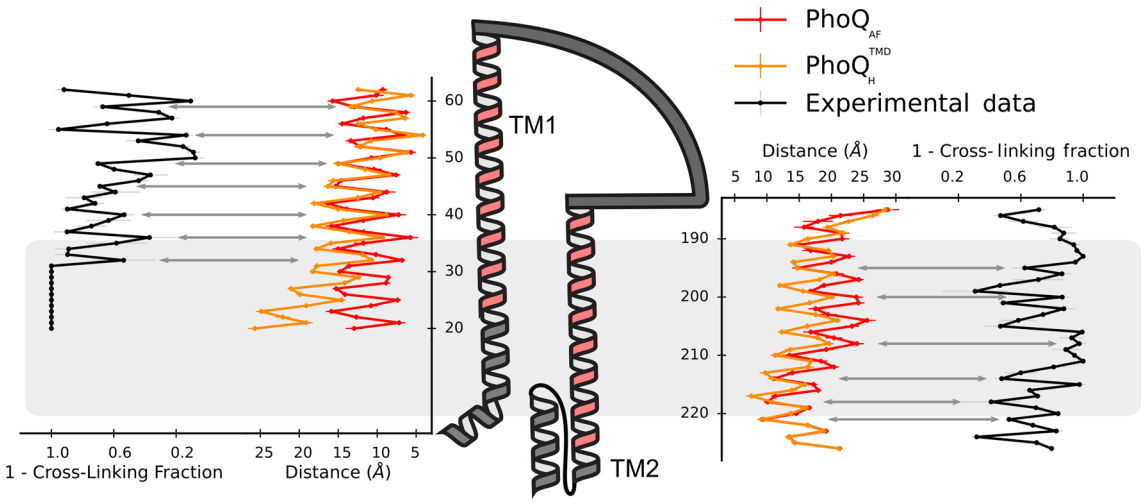

Figure 4. Comparison of cross-linking data with Cα distances in PhoQ models. Cross-linking data is represented in black, and the corresponding Cα inter-residue distances for $PhoQ_{AF}$ and $PhoQ_H^{TMD}$ are shown in red and orange, respectively. The membrane is depicted with a grey rectangle.

enabling transitions between PhoQ signaling states. Consistent with this hypothesis, the $PhoQ_H^{TMD}$ and $PhoQ_H$ models exhibited extensive transmembrane solvation throughout the simulations, contrasting with the complete absence of water molecules in the $PhoQ_{AF}$ model (Fig. 5). Detailed analysis revealed a dynamic pattern of water occupancy within the transmembrane bundle. On average, 1.7 water molecules were present within the TM bundle of $PhoQ_H^{TMD}$. The most frequent configuration involved two water molecules occupying the bundle (45 ± 11%). Occasionally, only one water molecule remained (23 ± 26%), and rarely, the bundle became dehydrated (14 ± 12%). The average residence time for an individual water molecule was around 287 ± 156 ns. Interestingly,

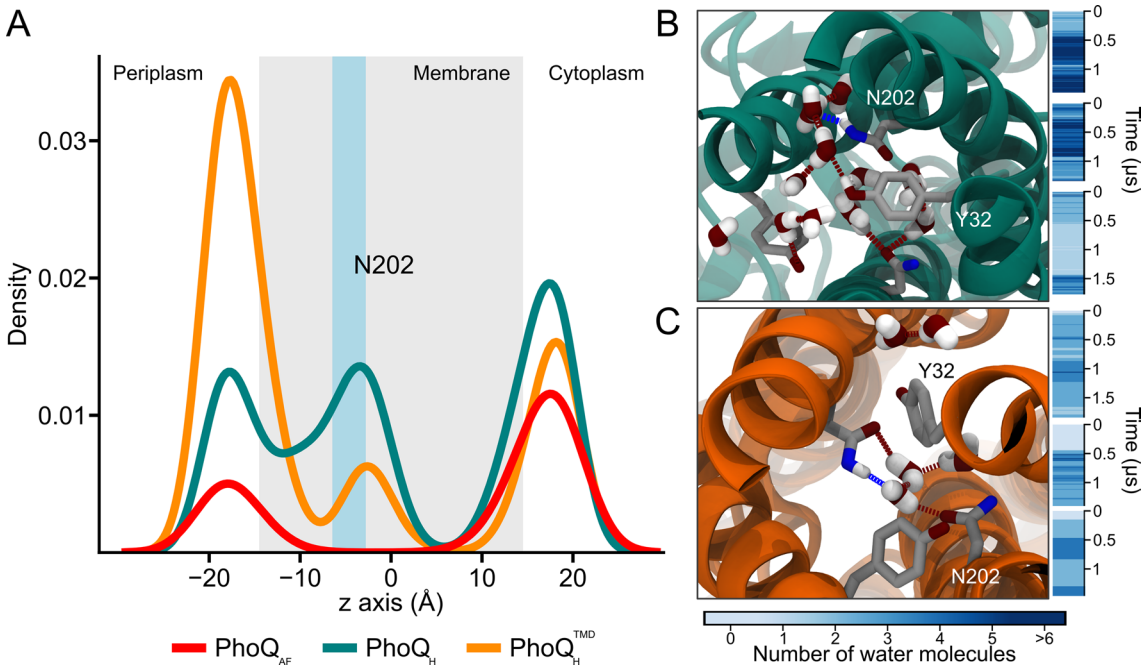

Figure 5. Solvation of the transmembrane domain. (**A**) Probability density profile of water molecules along the axis perpendicular to the membrane (z-axis) of the $PhoQ_{AF}$ (red), $PhoQ_H$ (teal), and $PhoQ_H^{TMD}$ (orange) models. (**B**,**C**) Molecular representations of frames with the highest water occupancy for $PhoQ_H$ and $PhoQ_H^{TMD}$ respectively. The protein is shown with a cartoon representation in teal ($PhoQ_H$) and orange ($PhoQ_H^{TMD}$) and residues N202 and Y32 are highlighted in licorice. Water molecules are shown in red (oxygen) and white (hydrogen). The bottom blue intensity plots represent water molecule count throughout the simulations (three replicas), with higher intensity indicating more water molecules.

some water molecules resided within the bundle for significantly longer durations, even up to 1 microsecond (μs) in one replica.

The transmembrane domain of $PhoQ_H$ displayed a greater degree of water occupancy, with an average of 3.5 water molecules per frame. The primary configurations involved either one (20 ± 30%) or two water molecules (19 ± 8%) or a more dynamic distribution of three to seven water molecules (27 ± 24%). The average water residence time was about 70 ± 50 ns, with one instance remaining in the bundle for the entire simulation (1.4 μs). To further investigate the functional role of N202, we conducted two additional 0.5 μs simulations of the N202A mutant. These simulations revealed a significant decrease in solvation within the transmembrane bundle, with only residual solvation remaining due to interactions with Y32.

Interactions with Mg^{2+} cations

Considering the crucial role of PhoQ in Mg^{2+} sensing, we investigated the presence of Mg^{2+} ions in the vicinity of the acidic patch (residues: L131—D150), known to be important for sensing cations. Minimal interactions with Mg^{2+} were observed for the $PhoQ_{AF}$ and $PhoQ_H$ models (less than 1% of the total simulation time for both chains). In striking contrast, the acidic patch of the $PhoQ_H^{TMD}$ model exhibited extensive Mg^{2+} interactions with the acidic patch, particularly residues D149 and D152 of chain A (100% simulation time). Interestingly, each bound Mg^{2+} cation also interacted with a lipid phosphate group of the membrane, either directly or indirectly through water molecules (Fig. 6). Although positioned further from the membrane surface, residue D125 in the acidic patch of chain B for $PhoQ_H^{TMD}$ still bound to a Mg^{2+} ion for 95% of the simulation time.

Dynamics of the cytoplasmic region

Despite overall similar appearances across simulations, principal component analysis (PCA) identified subtle but distinct conformational changes within the cytoplasmic domain of PhoQ (Supplementary Figure S12). The first principal component (PC1) captured a dominant bending motion in the B-chain of the S249-E296 helix (Supplementary Figure S12B, D), whereas the second principal component (PC2) described a scissoring motion of the N-terminal helices of the HAMP domain (R219-H234) (Supplementary Figure S12C, E). The residual motions captured by PC3 and PC4 describe complex bending movements (Supplementary Figure S13 and Movie S3, S4). These small conformational shifts within the HAMP and DHp domains resulted in the dissociation of the catalytic domain from the core region in several simulations. To characterize these diverse conformations, we employed two key metrics: (i) the distance between a catalytic residue (H277) and the center of mass of the catalytic domain (excluding the lid region), and (ii) the angle formed between the DHp domain core and the catalytic domain's gripper helix (G445-E457). This analysis identified seven distinct poses adopted by the catalytic domain (Fig. 7). The most prevalent conformation (red in Fig. 7) closely resembled the Alphafold prediction, with the gripper helix bound to the DHp domain. Notably, at least one cytoplasmic domain adopted this conformation in almost all simulations (Supplementary Figure S14). Two additional conformations (light orange and green in Fig. 7) exhibited an orientation that resembled a trans-phosphorylation state. These conformations positioned the catalytic domain near the core, with the gripper helix perpendicular to the DHp domain. However, they crucially lacked interaction with the key residue H277. The light orange conformation, observed primarily in chain A of the $PhoQ_H^{TMD}$ model, positions the domain slightly closer to the core compared to the green conformation. The

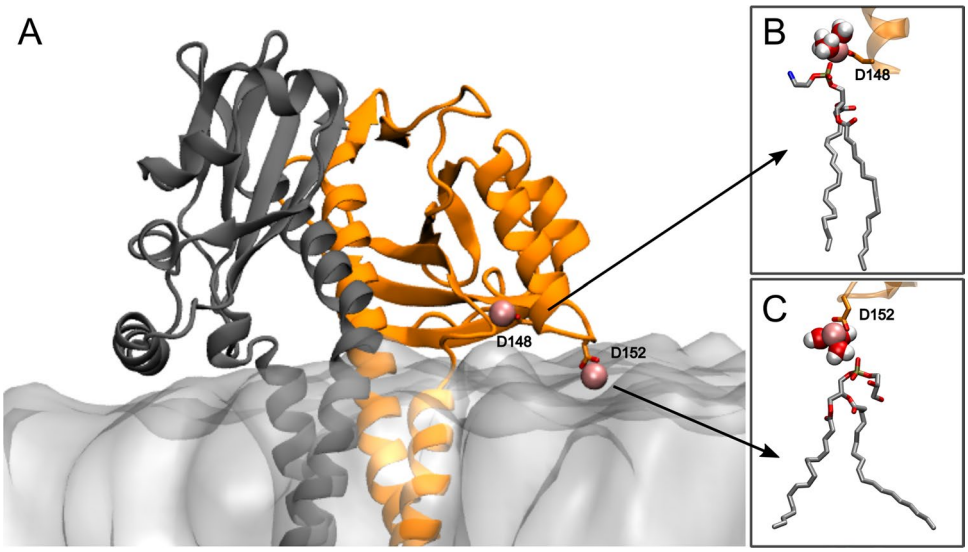

Figure 6. Mg^{2+}-bridged interactions of the acidic patch and the membrane. (**A**) Cartoon representation of the sensor domain (gray/orange) interacting with the membrane (transparent gray surface). Mg^{2+} ions (pink spheres) bridge the acidic patch residues (D148, D152, sticks) and the membrane. Representative snapshots of interactions occurring directly (**B**) and via a water molecule (**C**).

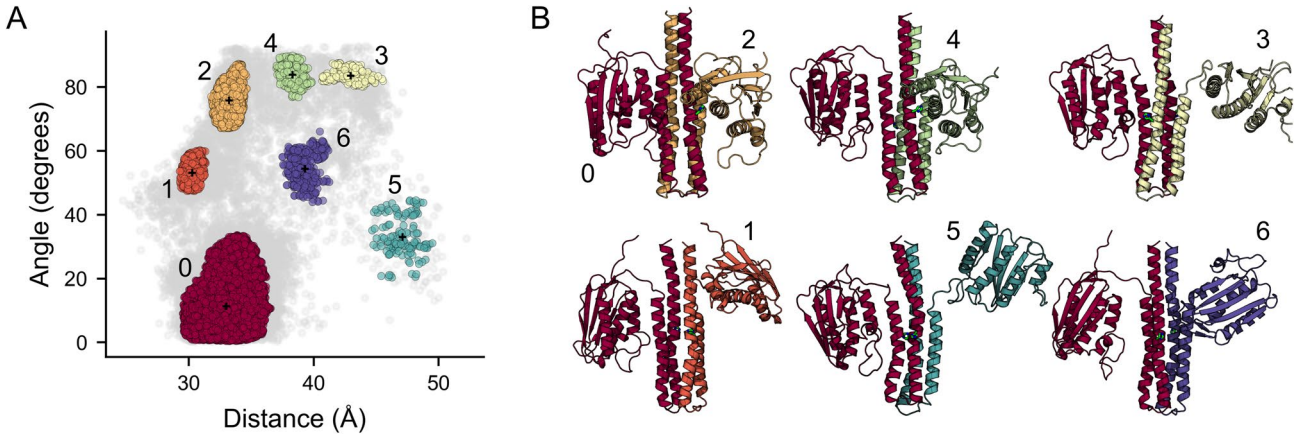

Figure 7. Conformational dynamics of the catalytic domain. (**A**) Scatter plot depicting the distance between H277 (DHp bundle) and the catalytic domain's center of mass (x-axis) versus the angle between the gripper helix and the DHp core axis (y-axis). Each chain was analyzed separately and the data was merged in the plot. A multi-step clustering identified seven clusters (colored dots) and gray dots represent transition conformations between clusters. The black cross indicates the center of mass of each cluster. (**B**) Cartoon representations of the major identified conformations. Each chain is colored according to its cluster assignment in panel (**A**).

remaining clusters (yellow, purple, cyan, and orange) all represented detached conformations of the catalytic domain lacking interaction with the DHp core. Notably, the cyan cluster corresponded to a conformation where the domain moved towards and directly interacted with the membrane surface.

Dynamic network analysis

To further explore the signal transduction mechanism of PhoQ within each simulation, we employed dynamic network analysis. We focused on identifying the shortest paths connecting the Mg^{2+}-sensing acidic patch (D152) to the phosphorylation site (H277) for each acidic patch. We then selected the shortest path for each histidine residue on both chains. The shortest path lengths ranged from 22 amino acids in $PhoQ_H^{TMD}$ to 27 in $PhoQ_H$. In all simulations, the identified minimum shortest paths terminated always on the H277 of the same chain (chain A for $PhoQ_H$, and chain B for $PhoQ_{AF}$ and $PhoQ_H^{TMD}$), supporting the proposed asymmetric model of PhoQ activation. It should be noted that the difference in shortest path lengths between protomers was minimal, with an average variation of only 0.11 + /- 0.02% between inter and intra-path lengths (average path length: 22.84 + /- 1.57). The specific residues involved in the shortest paths differed between protomers. In $PhoQ_{AF}$, the path switched protomers from R14 to E246', whereas in $PhoQ_H^{TMD}$, it switched from I221 to L250'. The shortest path for chain A of $PhoQ_H$ switched protomers twice (Y32 to L205' and V228' to L257), while the path for chain B switched only once (V228' to L257) (Supplementary Figure S15 A).

To identify the most crucial amino acids involved in PhoQ's signaling mechanisms, we analyzed the network of each state using the betweenness centrality. This metric identifies residues that frequently act as bridges between different communication pathways within a network. The analysis revealed distinct patterns for the different PhoQ states (Supplementary Figure S15 B). In the $PhoQ_{AF}$ model, the analysis identified key residues located at the interface between the two protein chains (L13-N57, E246-T273) and within the structured part of the HAMP domain. Additionally, the interconnecting loop bridging the DHp and catalytic domains (S326-S349) exhibited high centrality, suggesting its potential importance in signal transmission. For the hybrid models ($PhoQ_H$ and $PhoQ_H^{TMD}$,), the analysis highlighted residues with high centrality at the interfaces between various domains. These included the TM2-HAMP (L210-E222), sensor-TM1 (R50-A31), and HAMP-DHp interfaces (N255-L279) in both protein chains. Notably, similar to the $PhoQ_{AF}$ model, the interconnecting loop between the DHp and catalytic domains (T331-A341) emerged as important in these models as well.

To investigate further how the internal dynamics of PhoQ change across different functional states, we performed a community analysis. Communities are defined as clusters of amino acid residues that interact strongly (high cross-correlation), potentially representing functional units within the protein. The analysis revealed distinct clustering patterns for each PhoQ state, particularly in the periplasmic and transmembrane regions (Supplementary Figure S15 C). In the $PhoQ_H^{TMD}$ model, the transmembrane (TM) domains and sensor interface of both protomers, formed a tightly-knit community, while the remaining parts of the sensor domain constituted distinct clusters. In contrast, in $PhoQ_H$ the entire sensor and TM1 formed community for each protomer, while the TM2 and TM2' domains represented a separate community. Finally, in $PhoQ_{AF}$, the sensor and transmembrane domains of chain B formed a community, while an additional community encompassing the acidic patch and part of TM2 was also identified in chain A. Interestingly, the cytoplasmic region displayed a consistent pattern across all models, where the HAMP-DHp domains formed a single community and each catalytic domain belonged to two separate communities.

Discussion

This study leveraged AlphaFold structural predictions and MD simulations to explore the conformational dynamics of PhoQ, a sensor protein critical for bacterial signal transduction. Although the initial AlphaFold2 prediction displayed high confidence and domain arrangements consistent with expectations, it depicted a near-parallel interface in the sensor domain, unlike the ~ 40° angle observed in the known X-ray structure (PDB ID: 3bq8). By integrating a previously published model of the transmembrane domain[39] with targeted MD simulations, we successfully captured two additional conformations of PhoQ (hybrid models: $PhoQ_H^{TMD}$ and $PhoQ_H$). Even though all models remained stable and compatible with available cross-linking data[40], small structural deviations within the sensor and transmembrane domains resulted in substantial overall conformational changes between the models, providing insights into PhoQ signaling at atomic resolution (Fig. 8 and Supplementary Movie S5).

In particular, the hybrid structures exhibited a high degree of asymmetry, aligning with previous evidence suggesting an asymmetric nature of PhoQ signaling[13,33], and generally of histidine kinases[35]. The opening of the sensor domain interface resulted in the reorientation of one of the acidic patches of $PhoQ_H^{TMD}$, thus allowing the bridging of Mg^{2+} cation with the membrane bilayer. This observation corresponds perfectly with previous hypotheses based on X-ray structures[33,51]. These studies proposed that orienting the acidic patch towards the membrane through Mg^{2+} bridges triggers PhoQ signaling. This bridging interaction between Mg^{2+} and both the acidic patch and the membrane is crucial since the negatively charged acidic patch and the negatively charged phosphate groups in the membrane would otherwise repel each other. The Mg^{2+} interaction occurred either directly with membrane phosphate groups or mediated through a water molecule. Based on this, we propose that $PhoQ_H^{TMD}$ represents the repressed state of PhoQ, where Mg^{2+} binding stabilizes the opening of the sensor interface helices.

In contrast to the complete absence of water molecules observed in $PhoQ_{AF}$ and the partial hydration of $PhoQ_H^{TMD}$, $PhoQ_H$ displayed a significantly higher degree of water molecules occupying the core of its transmembrane domain. This finding agrees with prior research indicating the importance of the polar residue N202 for signaling and the proposed presence of a water-filled cavity[39,41], also apparent in the HtrII structure[38,64]. It was hypothesized that this hydration pocket facilitates transitions between the different states of PhoQ, by relaxing the tertiary packing interactions in the bundle[39]. The high level of hydration in $PhoQ_H$ TM could indicate that this conformation is an intermediate state within the signaling landscape of PhoQ. The higher intrinsic preference of the sensor domain for the kinase on-state (not bound to Mg^{2+}) could explain the residual hydration observed in $PhoQ_H^{TMD}$, which would allow an easier switching away from the repressed state[39,41].

Despite significant conformational changes in the sensor and transmembrane domains, the HAMP and DHp domains remained remarkably similar across all conformations. This observation lends further support to the previously proposed allosteric coupling model[65] between these domains, where communication occurs through subtle structural changes rather than a single, concerted conformational shift. Notably, a slight bend in the cytoplasmic S-helix, aligning with solved structures of similar domains[46], and a scissoring motion observed in the HAMP domain, consistent with proposed signaling of HAMP domain[60,66], appear sufficient to transmit these signals towards the catalytic domain. Furthermore, the catalytic domain exhibited higher mobility in the $PhoQ_H^{TMD}$ (repressed state) simulations. Although none of the observed catalytic domain conformations

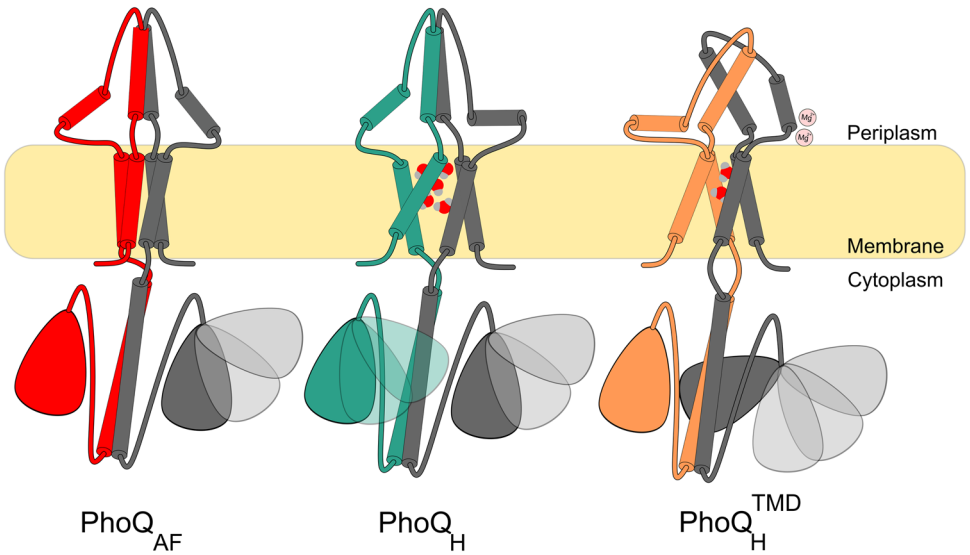

Figure 8. Schematic representation of identified conformations of PhoQ within its signaling landscape. Three potential conformations of the PhoQ signaling protein were identified through computational modeling. AlphaFold predicted a stable and symmetrical structure ($PhoQ_{AF}$, red), which might represent a pre-activated state before autophosphorylation. A hydrated conformation ($PhoQ_H$, teal) could correspond to a transition state during signaling. Finally, a highly asymmetric structure with minimal hydration and two bound Mg^{2+} ions ($PhoQ_H^{TMD}$, orange) indicates the repressed state of PhoQ.

facilitated autophosphorylation of the histidine residue, this enhanced mobility could represent another regulatory mechanism for repressing PhoP phosphorylation.

Beyond elucidating PhoQ's specific activation pathway, these insights hold significant promise for guiding future investigations into the broader histidine kinase family. The shared structural features observed in PhoQ suggest a potential for these findings to be extrapolated to related proteins, paving the way for a more comprehensive understanding of histidine kinase signaling mechanisms across various biological systems. Furthermore, the validity of our generated models opens exciting avenues for further studies on PhoQ itself. These models can be instrumental in exploring PhoQ's modulation by small peptides[67] and antimicrobial agents, potentially leading to the development of novel therapeutic strategies.

Methods

Molecular modeling

The initial model was obtained using the sequence of *E. coli* K12 (P23837) and the colab implementation of AlphaFold2-multimer[49], with default parameters. The hybrid model was constructed by replacing the transmembrane domain of the predicted model with the transmembrane domain from a previously published atomic model[41]. Specifically, residues M1 to S43 in TM1 and S193 to L224 in TM2 were exchanged, by aligning sensor-TM1 interface (I30-S43) and the TM2-HAMP interface (R219-A226). To ensure smooth junctions between the different models, the Rosetta kinematic closure protocol (KIC) was applied for relaxation[68].

Molecular dynamics

The models of PhoQ dimers were embedded in a bacterial mimetic bilayer composed of a 3:1 ratio of phosphatidylethanolamine (POPE) and phosphatidylglycerol (POPG). The system was solvated with a water box with a padding of 20 Å and neutralized using NaCl (150 mM) and $MgCl_2$ (5 mM).

The simulations were performed with the CHARMM36 force field[69], including CMAP corrections for the protein. The water molecules were described with the TIP3P water parameterization[70].

Targeted molecular dynamics simulations

Targeted molecular dynamics simulations using NAMD2.14[71] were used to enforce the conformation of the interface helices in the sensor domain observed in the X-ray structure (PDB: 3bq8). The cutoff for non-bonded interactions was set to 12 Å with a switching distance at 10 Å. Periodic electrostatic interactions were computed using Particle-Mesh Ewald (PME) summation with a grid spacing smaller than 1 Å. To maintain a constant temperature of 310 K, Langevin dynamics with a damping coefficient of 1.0 ps was employed. Additionally, a constant pressure of 1 atm was maintained using the Langevin barostat[72].

To preserve the stability of the secondary structure during the simulation, extra bonds were applied to the torsional angles of the protein's backbone. The targeted dynamics simulation was carried out over 50 ns. The interface helices were then kept fixed for an additional 50 ns before finally releasing all constraints, except for the secondary structure, and allowing the system to evolve for an extra 50 ns.

Unrestrained molecular dynamics

Each system underwent extended Molecular Dynamics simulations using the OpenMM7.7.0 molecular engine[73]. The systems were first minimized for 5000 steps followed by an equilibration (1.75 ns), progressively releasing positional restraints on the backbone atoms. The cutoff for non-bonded interactions was set to 12 Å with a switching distance at 10 Å. The periodic electrostatic interactions were computed using particle-mesh Ewald (PME) summation. Constant temperature of 310 K was imposed by Langevin dynamics with a damping coefficient of 1.0 ps. Constant pressure of 1 atm was maintained with Monte Carlo barostat[74]. The hydrogen mass repartitioning scheme was used to achieve a 4 fs time-step[75]. Snapshots from each simulation were extracted at 1 ns time intervals for structural analysis. Each simulation was carried out up to at least 1.4 µs (Supplementary Table S2).

Analyses

We employed a combination of custom scripting and established software packages to analyze the data. Custom Tcl scripts were written to automate the extraction of features within VMD 1.9.3 software[76–78]. Structural alignments were carried out on the backbone atoms using the VMD 1.9.3 tool (RMSDTT). The extracted data was analyzed with pandas[79,80] and NumPy 1.23.4[81], and graphics were generated using Matplotlib[82]. The Principal Component Analysis (PCA) was carried out on the alpha carbon using the implementation provided by the ProDy 2.0 library[83]. The curvature of the helices was measured with the Bendix 1.1 tool[84].

Foldseek was employed for a fast structure-based search against the PDB databank, using the 3diaa mode. The retrieved structures were structurally aligned using the TM-align[56] method via the RCSB.org[57].

The HAMP domain principle component analysis was conducted using the HAMP structures from[60] and Foldseek alignment tool locally with minimum TM-score of 0.5, sensitivity of 9.5 minimum sequence identity of 20%, and coverage of 80% in both query and target.

The accommodation index was employed to assess the perturbation of the coiled-coil core of PhoQ and specifically between the regions S12-L62 and S187-N235 [https://grigoryanlab.org/cccp/index.ai.php].

To identify groups of similar conformations of PhoQ cytoplasmic domain, we employed a multi-step clustering approach. Initially, we utilized Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for exploratory analysis[85,86]. Due to uneven data point density across the conformational space, we implemented an additional step. This involved manually dividing the space into sub-regions and running localized DBSCAN analyses with adjusted bandwidths for each sub-region.

The network analysis was conducted on a dynamical network constructed using the correlation between residues computed using carma[87]. The network was further analyzed using the networkx 2.8.4 library[88] implemented in python. The shortest path was computed using the Dijkstra algorithm, the communities were computed via the Clauset-Newman-Moore greedy modularity maximization[89] and the key residues were computed via the shortest-path betweenness centrality[90].

The visualization and molecular renderings were produced with VMD 1.9.3 and Pymol 2.5.0[91,92]. Morphing was performed using ChimeraX 1.4[93–95].

Data availability

All simulations are available at 10.5281/zenodo.10988292 Script for analyses are available at https://github.com/symelaz/PhoQ.

Received: 18 April 2024; Accepted: 22 July 2024
Published online: 26 July 2024

References

1. Gadd, G. M. Metals, minerals and microbes: Geomicrobiology and bioremediation. *Microbiology* **156**(3), 609–643. https://doi.org/10.1099/mic.0.037143-0 (2010).
2. Fierer, N. Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* **15**(10), 579–590. https://doi.org/10.1038/nrmicro.2017.87 (2017).
3. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: Many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* **6**(8), 613–624. https://doi.org/10.1038/nrmicro1932 (2008).
4. Groisman, E. A. Feedback control of two-component regulatory systems. *Annu. Rev. Microbiol.* **70**(1), 103–124. https://doi.org/10.1146/annurev-micro-102215-095331 (2016).
5. Cheung, J. & Hendrickson, W. A. Sensor domains of two-component regulatory systems. *Curr. Opin. Microbiol.* **13**(2), 116–123. https://doi.org/10.1016/j.mib.2010.01.016 (2010).
6. Stock, A. M., Robinson, V. L., & Goudreau, P. N. Two-component signal transduction (2000).
7. Groisman, E. A. & Mouslim, C. Sensing by bacterial regulatory systems in host and non-host environments. *Nat. Rev. Microbiol.* **4**(9), 705–709. https://doi.org/10.1038/nrmicro1478 (2006).
8. Hirakawa, H., Nishino, K., Hirata, T. & Yamaguchi, A. Comprehensive studies of drug resistance mediated by overexpression of response regulators of two-component signal transduction systems in *Escherichia coli*. *J. Bacteriol.* **185**(6), 1851–1856. https://doi.org/10.1128/JB.185.6.1851-1856.2003 (2003).
9. Nishino, K., Honda, T. & Yamaguchi, A. Genome-wide analyses of *Escherichia coli* gene expression responsive to the BaeSR two-component regulatory system. *J. Bacteriol.* **187**(5), 1763–1772. https://doi.org/10.1128/JB.187.5.1763-1772.2005 (2005).
10. Nizet, V. Antimicrobial peptide resistance mechanisms of human bacterial pathogens. *Curr. Issues Mol. Biol.* https://doi.org/10.21775/cimb.008.011 (2006).
11. Montagne, M., Martel, A. & Le Moual, H. Characterization of the catalytic activities of the PhoQ histidine protein kinase of *Salmonella enterica* Serovar Typhimurium. *J. Bacteriol.* **183**(5), 1787–1791. https://doi.org/10.1128/JB.183.5.1787-1791.2001 (2001).
12. Sanowar, S. & Le Moual, H. Functional reconstitution of the *Salmonella typhimurium* PhoQ histidine kinase sensor in proteoliposomes. *Biochem. J.* **390**(3), 769–776. https://doi.org/10.1042/BJ20050060 (2005).
13. Yeo, W.-S. *et al.* Intrinsic negative feedback governs activation surge in two-component regulatory systems. *Mol. Cell* **45**(3), 409–421. https://doi.org/10.1016/j.molcel.2011.12.027 (2012).
14. Wuichet, K., Cantwell, B. J. & Zhulin, I. B. Evolution and phyletic distribution of two-component signal transduction systems. *Curr. Opin. Microbiol.* **13**(2), 219–225. https://doi.org/10.1016/j.mib.2009.12.011 (2010).
15. Miller, S. I., Kukral, A. M. & Mekalanos, J. J. A two-component regulatory system (phoP phoQ) controls Salmonella typhimurium virulence. *Proc. Natl. Acad. Sci.* **86**(13), 5054–5058. https://doi.org/10.1073/pnas.86.13.5054 (1989).
16. Hancock, R. E. W. & McPhee, J. B. Salmonella's sensor for host defense molecules. *Cell* **122**(3), 320–322. https://doi.org/10.1016/j.cell.2005.07.023 (2005).
17. Velkinburgh, J. C. V. & Gunn, J. S. PhoP-PhoQ-regulated loci are required for enhanced bile resistance in Salmonella spp. *Infect. Immun.* **67**, 1614–1622 (1999).
18. Baker, S. J., Gunn, J. S., & Morona, R. The Sahonella @phi rnelittin resistance gene pgaB affects intracellular growth in PMA= differentiated U937 cells, polymyxin B resistance and Iipopolysaccharide.
19. Véscovi, E. G., Soncini, F. C. & Groisman, E. A. Mg2+ as an extracellular signal: Environmental regulation of salmonella virulence. *Cell* **84**(1), 165–174. https://doi.org/10.1016/S0092-8674(00)81003-X (1996).
20. Prost, L. R. *et al.* Activation of the bacterial sensor kinase PhoQ by acidic pH. *Mol. Cell* **26**(2), 165–174. https://doi.org/10.1016/j.molcel.2007.03.008 (2007).
21. Brodsky, I. E., Ernst, R. K., Miller, S. I. & Falkow, S. *mig - 14* Is a *Salmonella* gene that plays a role in bacterial resistance to antimicrobial peptides. *J. Bacteriol.* **184**(12), 3203–3213. https://doi.org/10.1128/JB.184.12.3203-3213.2002 (2002).
22. Groisman, E. A., Parra-Lopez, C., Salcedo, M., Lipps, C. J. & Heffron, F. Resistance to host antimicrobial peptides is necessary for Salmonella virulence. *Proc. Natl. Acad. Sci.* **89**(24), 11939–11943. https://doi.org/10.1073/pnas.89.24.11939 (1992).
23. Shprung, T., Peleg, A., Rosenfeld, Y., Trieu-Cuot, P. & Shai, Y. Effect of PhoP-PhoQ activation by broad repertoire of antimicrobial peptides on bacterial resistance. *J. Biol. Chem.* **287**(7), 4544–4551. https://doi.org/10.1074/jbc.M111.278523 (2012).
24. Wang, B. *et al.* Affinity-based capture and identification of protein effectors of the growth regulator ppGpp. *Nat. Chem. Biol.* **15**(2), 141–150. https://doi.org/10.1038/s41589-018-0183-4 (2019).
25. Viarengo, G. *et al.* Unsaturated long chain free fatty acids are input signals of the Salmonella enterica PhoP/PhoQ regulatory system. *J. Biol. Chem.* **288**(31), 22346–22358. https://doi.org/10.1074/jbc.M113.472829 (2013).
26. Lejona, S., Aguirre, A., Cabeza, M. L., Véscovi, E. G. & Soncini, F. C. Molecular characterization of the Mg^{2+} -responsive PhoP-PhoQ regulon in *Salmonella enterica*. *J. Bacteriol.* **185**(21), 6287–6294. https://doi.org/10.1128/JB.185.21.6287-6294.2003 (2003).
27. Shin, D., Lee, E.-J., Huang, H. & Groisman, E. A. A positive feedback loop promotes transcription surge that jump-starts *Salmonella* virulence circuit. *Science* **314**(5805), 1607–1609. https://doi.org/10.1126/science.1134930 (2006).
28. Aravind, L. & Ponting, C. P. The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. *FEMS Microbiol. Lett.* **176**(1), 111–116. https://doi.org/10.1111/j.1574-6968.1999.tb13650.x (1999).
29. Mascher, T., Helmann, J. D. & Unden, G. Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol. Mol. Biol. Rev.* **70**(4), 910–938. https://doi.org/10.1128/MMBR.00020-06 (2006).
30. Cheung, J. & Hendrickson, W. A. Structural analysis of ligand stimulation of the histidine kinase NarX. *Structure* **17**(2), 190–201. https://doi.org/10.1016/j.str.2008.12.013 (2009).

31. Sevvana, M. *et al.* A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA. *J. Mol. Biol.* **377**(2), 512–523. https://doi.org/10.1016/j.jmb.2008.01.024 (2008).
32. Moore, J. O. & Hendrickson, W. A. An asymmetry-to-symmetry switch in signal transmission by the histidine kinase receptor for TMAO. *Structure* **20**(4), 729–741. https://doi.org/10.1016/j.str.2012.02.021 (2012).
33. Cheung, J., Bingman, C. A., Reyngold, M., Hendrickson, W. A. & Waldburger, C. D. Crystal structure of a functional dimer of the PhoQ sensor domain. *J. Biol. Chem.* **283**(20), 13762–13770. https://doi.org/10.1074/jbc.M710592200 (2008).
34. Goldberg, S. D., Soto, C. S., Waldburger, C. D. & DeGrado, W. F. Determination of the physiological dimer interface of the PhoQ sensor domain. *J. Mol. Biol.* **379**(4), 656–665. https://doi.org/10.1016/j.jmb.2008.04.023 (2008).
35. Bhate, M. P., Molnar, K. S., Goulian, M. & DeGrado, W. F. Signal transduction in histidine kinases: Insights from new structures. *Structure* **23**(6), 981–994. https://doi.org/10.1016/j.str.2015.04.002 (2015).
36. Jiang, P., Peliska, J. A. & Ninfa, A. J. Asymmetry in the autophosphorylation of the two-component regulatory system transmitter protein nitrogen regulator II of Escherichia coli. *Biochemistry* **39**(17), 5057–5065. https://doi.org/10.1021/bi992921w (2000).
37. Gordeliy, V. I. *et al.* Molecular basis of transmembrane signalling by sensory rhodopsin II–transducer complex. *Nature* https://doi.org/10.1038/nature01109 (2002).
38. Moukhametzianov, R. *et al.* Development of the signal in sensory rhodopsin and its transfer to the cognate transducer. *Nature* https://doi.org/10.1038/nature04520 (2006).
39. Goldberg, S. D., Clinthorne, G. D., Goulian, M. & DeGrado, W. F. Transmembrane polar interactions are required for signaling in the Escherichia coli sensor kinase PhoQ. *Proc. Natl. Acad. Sci. USA* **107**(18), 8141–8146. https://doi.org/10.1073/pnas.1003166107 (2010).
40. Molnar, K. S. *et al.* Cys-scanning disulfide crosslinking and Bayesian modeling probe the transmembrane signaling mechanism of the histidine kinase, PhoQ. *Structure* **22**(9), 1239–1251. https://doi.org/10.1016/j.str.2014.04.019 (2014).
41. Lemmin, T., Soto, C. S., Clinthorne, G., DeGrado, W. F. & Peraro, M. D. Assembly of the transmembrane domain of *E. coli* PhoQ histidine kinase: Implications for signal transduction from molecular simulations. *PLoS Comput. Biol.* **9**(1), e1002878. https://doi.org/10.1371/journal.pcbi.1002878 (2013).
42. Dunin-Horkawicz, S. & Lupas, A. N. Comprehensive analysis of HAMP domains: Implications for transmembrane signal transduction. *J. Mol. Biol.* **397**(5), 1156–1174. https://doi.org/10.1016/j.jmb.2010.02.031 (2010).
43. Wang, C. *et al.* Mechanistic insights revealed by the crystal structure of a histidine kinase with signal transducer and sensor domains. *PLOS Biol.* **11**(2), e1001493. https://doi.org/10.1371/journal.pbio.1001493 (2013).
44. Ferris, H. U. *et al.* Mechanism of regulation of receptor histidine kinases. *Structure* **20**(1), 56–66. https://doi.org/10.1016/j.str.2011.11.014 (2012).
45. Albanesi, D. *et al.* Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proc. Natl. Acad. Sci.* **106**(38), 16185–16190. https://doi.org/10.1073/pnas.0906699106 (2009).
46. Mechaly, A. E., Sassoon, N., Betton, J.-M. & Alzari, P. M. Segmental helical motions and dynamical asymmetry modulate histidine kinase autophosphorylation. *PLoS Biol.* **12**(1), e1001776. https://doi.org/10.1371/journal.pbio.1001776 (2014).
47. Hulko, M. *et al.* The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell* **126**(5), 929–940. https://doi.org/10.1016/j.cell.2006.06.058 (2006).
48. Swain, K. E., Gonzalez, M. A. & Falke, J. J. Engineered socket study of signaling through a four-helix bundle: Evidence for a Yin–Yang mechanism in the kinase control module of the aspartate receptor. *Biochemistry* **48**(39), 9266–9277. https://doi.org/10.1021/bi901020d (2009).
49. Evans *et al.*, R. Protein complex prediction with AlphaFold-Multimer. *Bioinformatics*, preprint (2021). https://doi.org/10.1101/2021.10.04.463034.
50. Ponting, C. P. & Aravind, L. PAS: A multifunctional domain family comes to light. *Curr. Biol.* **7**(11), R674–R677. https://doi.org/10.1016/S0960-9822(06)00352-6 (1997).
51. Cho, U. S. *et al.* Metal bridges between the PhoQ Sensor domain and the membrane regulate transmembrane signaling. *J. Mol. Biol.* **356**(5), 1193–1206. https://doi.org/10.1016/j.jmb.2005.12.032 (2006).
52. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376. https://doi.org/10.1093/nar/gkr703 (2012).
53. Milburn, M. V. *et al.* Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* **254**(5036), 1342–1347. https://doi.org/10.1126/science.1660187 (1991).
54. Schmidt, N. W., Grigoryan, G. & DeGrado, W. F. The accommodation index measures the perturbation associated with insertions and deletions in coiled-coils: Application to understand signaling in histidine kinases. *Protein Sci.* **26**(3), 414–435. https://doi.org/10.1002/pro.3095 (2017).
55. Van Kempen, M. *et al.* Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-023-01773-0 (2023).
56. Zhang, Y. & Skolnick, J. TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33**(7), 2302–2309. https://doi.org/10.1093/nar/gki524 (2005).
57. Berman, H. M. *et al.* The protein data bank. *Nucleic Acids Res.* **28**(1), 235–242. https://doi.org/10.1093/nar/28.1.235 (2000).
58. Yuan, J., Jin, F., Glatter, T. & Sourjik, V. Osmosensing by the bacterial PhoQ/PhoP two-component system. *Proc. Natl. Acad. Sci. USA* **114**(50), E10792–E10798. https://doi.org/10.1073/pnas.1717272114 (2017).
59. Yadavalli, S. S. *et al.* Functional determinants of a small protein controlling a broadly conserved bacterial sensor kinase. *J. Bacteriol.* https://doi.org/10.1128/JB.00305-20 (2020).
60. Winski, A. *et al.* AlphaFold2 captures the conformational landscape of the HAMP signaling domain. *Protein Sci.* **33**(1), e4846. https://doi.org/10.1002/pro.4846 (2024).
61. Ashenberg, O., Keating, A. E. & Laub, M. T. Helix bundle loops determine whether histidine kinases autophosphorylate in cis or in trans. *J. Mol. Biol.* **425**(7), 1198–1209. https://doi.org/10.1016/j.jmb.2013.01.011 (2013).
62. Marina, A., Mott, C., Auyzenberg, A., Hendrickson, W. A. & Waldburger, C. D. Structural and mutational analysis of the PhoQ histidine kinase catalytic domain. *J. Biol. Chem.* **276**(44), 41182–41190. https://doi.org/10.1074/jbc.M106080200 (2001).
63. Hekkelman, M. L., de Vries, I., Joosten, R. P. & Perrakis, A. AlphaFill: Enriching AlphaFold models with ligands and cofactors. *Nat. Methods* https://doi.org/10.1038/s41592-022-01685-y (2023).
64. Bergo, V. B., Spudich, E. N., Spudich, J. L. & Rothschild, K. J. Active Water in protein–protein communication within the membrane: The case of SRII–HtrII signal relay. *Biochemistry* **48**(5), 811–813. https://doi.org/10.1021/bi802180a (2009).
65. Mensa, B. *et al.* Allosteric mechanism of signal transduction in the two-component system histidine kinase PhoQ. *eLife* **10**, e73336. https://doi.org/10.7554/eLife.73336 (2021).
66. Gushchin, I. *et al.* Mechanism of transmembrane signaling by sensor histidine kinases. *Science* **356**(6342), eaah6345. https://doi.org/10.1126/science.aah6345 (2017).
67. Jiang, S. *et al.* The inhibitory mechanism of a small protein reveals its role in antimicrobial peptide sensing. *Proc. Natl. Acad. Sci.* **120**(41), e2309607120. https://doi.org/10.1073/pnas.2309607120 (2023).
68. Mandell, D. J., Coutsias, E. A. & Kortemme, T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat. Methods* **6**(8), 551–552. https://doi.org/10.1038/nmeth0809-551 (2009).
69. Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**(25), 2135–2145. https://doi.org/10.1002/jcc.23354 (2013).

70. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**(2), 926–935. https://doi.org/10.1063/1.445869 (1983).
71. Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**(4), 044130. https://doi.org/10.1063/5.0014475 (2020).
72. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103**(11), 4613–4621. https://doi.org/10.1063/1.470648 (1995).
73. Eastman, P. *et al.* OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Comput. Biol.* **13**(7), e1005659. https://doi.org/10.1371/journal.pcbi.1005659 (2017).
74. Åqvist, J., Wennerström, P., Nervall, M., Bjelic, S. & Brandsdal, B. O. Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chem. Phys. Lett.* **384**(4), 288–294. https://doi.org/10.1016/j.cplett.2003.12.039 (2004).
75. Feenstra, K. A., Hess, B. & Berendsen, H. J. C. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **20**(8), 786–798. https://doi.org/10.1002/(SICI)1096-987X(199906)20:8%3c786::AID-JCC5%3e3.0.CO;2-B (1999).
76. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**(1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5 (1996).
77. Eargle, J., Wright, D. & Luthey-Schulten, Z. Multiple alignment of protein structures and sequences for VMD. *Bioinformatics* **22**(4), 504–506. https://doi.org/10.1093/bioinformatics/bti825 (2006).
78. Stone, J. E., Gullingsrud, J., & Schulten, K. A system for interactive molecular dynamics simulation," in *Proceedings of the 2001 symposium on Interactive 3D graphics*, ACM, Mar. 2001, pp. 191–194. https://doi.org/10.1145/364338.364398.
79. The pandas development team, "pandas-dev/pandas: Pandas." Feb. 2020. [Online]. https://doi.org/10.5281/zenodo.3509134
80. McKinney, W. *Data Structures for Statistical Computing in Python*. 2010, p. 61. https://doi.org/10.25080/Majora-92bf1922-00a.
81. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2 (2020).
82. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**(3), 90–95. https://doi.org/10.1109/MCSE.2007.55 (2007).
83. Bakan, A., Meireles, L. M. & Bahar, I. ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics* **27**(11), 1575–1577. https://doi.org/10.1093/bioinformatics/btr168 (2011).
84. Dahl, A. C. E., Chavent, M. & Sansom, M. S. P. Bendix: Intuitive helix geometry analysis and abstraction. *Bioinformatics* **28**(16), 2193–2194. https://doi.org/10.1093/bioinformatics/bts357 (2012).
85. Schubert, E., Sander, J., Ester, M., Kriegel, H. P. & Xu, X. DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Trans. Database Syst.* **42**(3), 1–21. https://doi.org/10.1145/3068335 (2017).
86. Ester, M., Kriegel, H.-P., & Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.
87. Glykos, N. M. Software news and updates carma: A molecular dynamics analysis program. *J. Comput. Chem.* **27**(14), 1765–1768. https://doi.org/10.1002/jcc.20482 (2006).
88. Hagberg, A. A., Schult, D. A., & Swart, P. J. Exploring Network Structure, Dynamics, and Function using NetworkX (2008).
89. Clauset, A., Newman, M. E. J. & Moore, C. Finding community structure in very large networks. *Phys. Rev. E* **70**(6), 066111. https://doi.org/10.1103/PhysRevE.70.066111 (2004).
90. Brandes, U. On variants of shortest-path betweenness centrality and their generic computation. *Soc. Netw.* **30**(2), 136–145. https://doi.org/10.1016/j.socnet.2007.11.001 (2008).
91. Stone, J. E. An efficient library for parallel ray tracing and animation.
92. Schrödinger, L., & Warren, D. PyMOL. May 20, 2020. [Online]. Available: http://www.pymol.org/pymol
93. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci. Publ. Protein Soc.* **27**(1), 14–25. https://doi.org/10.1002/pro.3235 (2018).
94. Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **32**(11), e4792. https://doi.org/10.1002/pro.4792 (2023).
95. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **30**(1), 70–82. https://doi.org/10.1002/pro.3943 (2021).

Acknowledgements

T.L. and S.L. acknowledges funding from Swiss National Science Foundation (SNSF: PCEFP3_194606) and J.Y. from the Max Planck Society.

Author contributions

S.L. study design, data acquisition, data analysis, article drafting, critical revision, and final approval of the manuscript. J.Y. conception of the idea, study design, article drafting, critical revision, and final approval of the manuscript. T.L. conception of the idea, study design, article drafting, critical revision, and final approval of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-68206-z.

Correspondence and requests for materials should be addressed to T.L.