nature chemistry

Article

https://doi.org/10.1038/s41557-023-01210-4

Transient water wires mediate selective proton transport in designed channel proteins

Received: 19 May 2022

Accepted: 19 April 2023

Published online: 12 June 2023

Check for updates

Huong T. Kratochvil[1,5] ✉, Laura C. Watkins[2,6], Marco Mravic[1,7], Jessica L. Thomaston[1], John M. Nicoludis[1,8], Noah H. Somberg[3], Lijun Liu[4], Mei Hong[3], Gregory A. Voth[2] ✉ & William F. DeGrado[1] ✉

Selective proton transport through proteins is essential for forming and using proton gradients in cells. Protons are conducted along hydrogen-bonded 'wires' of water molecules and polar side chains, which, somewhat surprisingly, are often interrupted by dry apolar stretches in the conduction pathways, inferred from static protein structures. Here we hypothesize that protons are conducted through such dry spots by forming transient water wires, often highly correlated with the presence of the excess protons in the water wire. To test this hypothesis, we performed molecular dynamics simulations to design transmembrane channels with stable water pockets interspersed by apolar segments capable of forming flickering water wires. The minimalist designed channels conduct protons at rates similar to viral proton channels, and they are at least 10^6-fold more selective for H^+ over Na^+. These studies inform the mechanisms of biological proton conduction and the principles for engineering proton-conductive materials.

The controlled diffusion of protons through transmembrane (TM) proteins is critical for many aspects of physiological function, including substrate transport[1], control of cellular and organelle pH (ref. 2), the creation and use of the pH gradients required for bioenergetics[3,4], and cellular signalling[5]. Proteins conduct protons along precisely defined pathways that prevent the wasteful collapse of Na^+ and K^+ gradients. Because the proton concentration in the cellular cytoplasm is about 10^6-fold lower than that of other ions at neutral pH, proton channels must have selectivities substantially greater than this value. The conduction of protons in water and aqueous pores is facilitated by the formation of water wires consisting of hydrogen-bonded chains of water molecules. In the classical Grotthuss mechanism, protons pass from one water molecule to the next to achieve long-range net transport without the need to move oxygen atoms (Fig. 1a)[6]. In this mechanism, the excess proton first forms a hydrated structure with a hydronium-like core[7], creating a net positive charge defect or hole that propagates 'downstream' of its initial position along the water wire[8].

Proteins achieve high proton selectivity by organizing and gating such water wires in proton conduction pathways often interspersed with ionizable side chains that explicitly participate in Grotthuss shuttling[9,10]. Somewhat surprisingly, however, protons often appear to be conducted through dry stretches of hydrophobic residues that feature in X-ray and cryoelectron microscopic structures of proteins[8,11–13]. Classical molecular dynamics (MD), reactive molecular dynamics (RMD) and quantum mechanical calculations suggest that water can occasionally penetrate through such apolar sectors with the help of polar residues, forming transient water wires not apparent in time-averaged structures (Fig. 1c)[8,11]. Thus, rare equilibrium fluctuations mediated by polar proton loading sites (PLSs) provide one mechanism for transient protonated water wire formation. Additionally, in systems that are energetically activated by chemical reactions, light or a TM potential,

[1]Department of Pharmaceutical Chemistry, University of California-San Francisco, San Francisco, CA, USA. [2]Department of Chemistry, Chicago Center for Theoretical Chemistry, Institute for Biophysical Dynamics and James Franck Institute, The University of Chicago, Chicago, IL, USA. [3]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. [4]DLX Scientific, Lawrence, KS, USA. [5]Present address: Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [6]Present address: Kemper Insurance, Chicago, IL, USA. [7]Present address: Department of Integrative Structural and Computational Biology Scripps Research Institute, La Jolla, CA, USA. [8]Present address: Genentech, San Francisco, CA, USA. ✉e-mail: huong.kratochvil@unc.edu; gavoth@uchicago.edu; william.degrado@ucsf.edu

https://doi.org/10.1038/s41557-023-01210-4

the arrival of an excess proton can induce the formation of transient water wires through confined hydrophobic spaces (Fig. 1d)[8,11,14]. One would expect very high proton selectivity from such a mechanism because a single column of connected water wires in a hydrophobic environment is unable to accommodate or stabilize a hydrated Na^+ or K^+ ion[15–20]. However, experimental evidence for such mechanisms has been indirect or lacking, with the exception of the very extensively studied protein bacteriorhodopsin in which transient spectroscopy and serial crystallography have identified water molecules that form during portions of its photocycle[13,21–23]. However, this system represents only a single example in which large changes are induced by photoisomerization.

We therefore turned to de novo design[24–28] to test and expand the transient water wire hypothesis. The design of a highly proton-selective channel operating by this mechanism not only tests an important concept in biophysics, but it also represents an important challenge in de novo protein design. While it has been possible to design stably packed proteins that bind small molecules and protein interfaces[29–31], more complex functions that rely on chemical dynamics have been difficult to design from scratch. Except for early studies in which transition states were considered as targets in the design of catalytic proteins[32,33], de novo design has generally focused on ground states rather than non-equilibrium high-energy states, as are generated during ion conduction. Computational design algorithms that also favour tight and efficient packing are likely to dampen the essential fluctuations required for catalysis and transport[34–36]. Moreover, de novo protein design does not consider explicit water molecules, and instead relies on approximations of the effect of solvent. In this study, we not only considered water explicitly, but we also accounted for the dynamic formation and breaking of covalent bonds as protons are passed from one molecule of water to the next. Finally, despite a few successes[27,37–40], the design and high-resolution structural characterization of membrane proteins remain a difficult endeavour. Two de novo channel-forming proteins have been structurally characterized, but they were not highly selective, nor were their per-channel conductance rates determined[38]. Other work focused on the conversion of water-soluble nanopores into membrane-spanning channels yielded channels with well-defined single-channel conductance characteristics, but the structure of the ion-conducting form of the channel was not determined[37]. Finally, tetrameric TM bundles were designed to use transition metal ion binding to drive proton translocation and vice versa, but the antiporting efficiency was limited by leakiness to protons[39]. Thus, the purposeful design of highly proton-selective channels that operate by a dynamic wetting/dewetting mechanism represents a notable advance.

Results

MD-guided design of proton channels

To test the transient water wire hypothesis, we designed a series of channels containing a polar PLS adjacent to a hydrophobic pore (Fig. 2). The expected length (l_{exp}) of the longest uninterrupted hydrophobic stretch in the pore, and the number and position of protein loading sites were varied. We chose a cluster of neutral Gln residues as the PLS to help stabilize a proton in the pore without falling into any deep energy wells that might occur with an ionizable residue. Gln and Asn feature in the proton conduction pathways of the S31N mutant of the influenza A matrix protein 2 (M2) and otopetrin proton channels[41,42]. Furthermore, Gln is also flexible so it should be able to stabilize the multiple polarizations of water wires that are created during a conduction cycle.

Our constructs were based on a previously characterized de novo homopentameric TM helical bundle with an interior stabilized by efficient van der Waals packing of alternating layers of apolar Leu

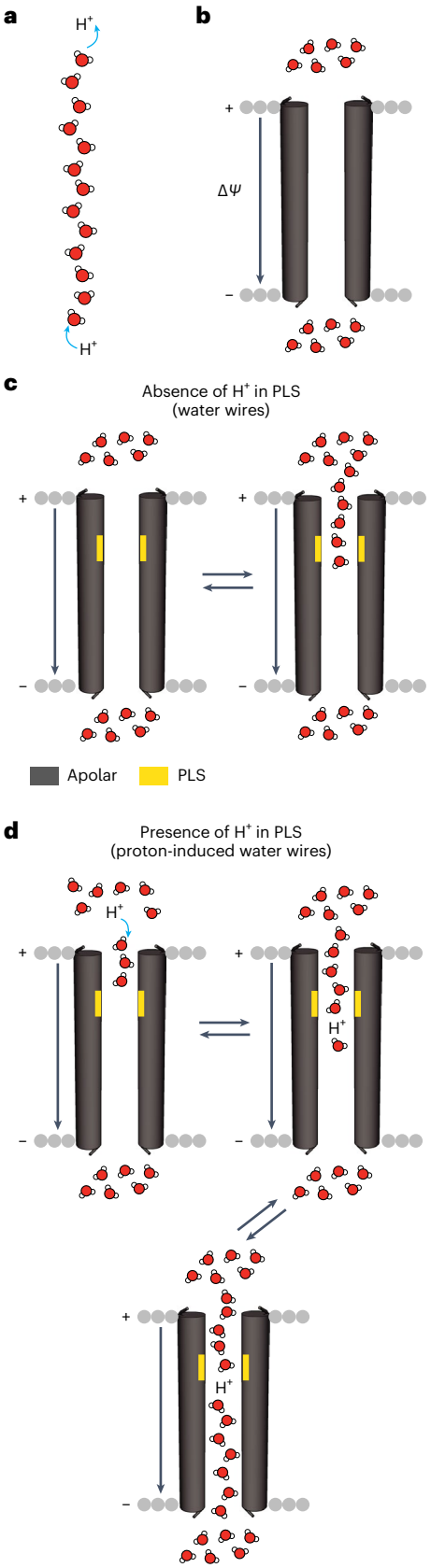
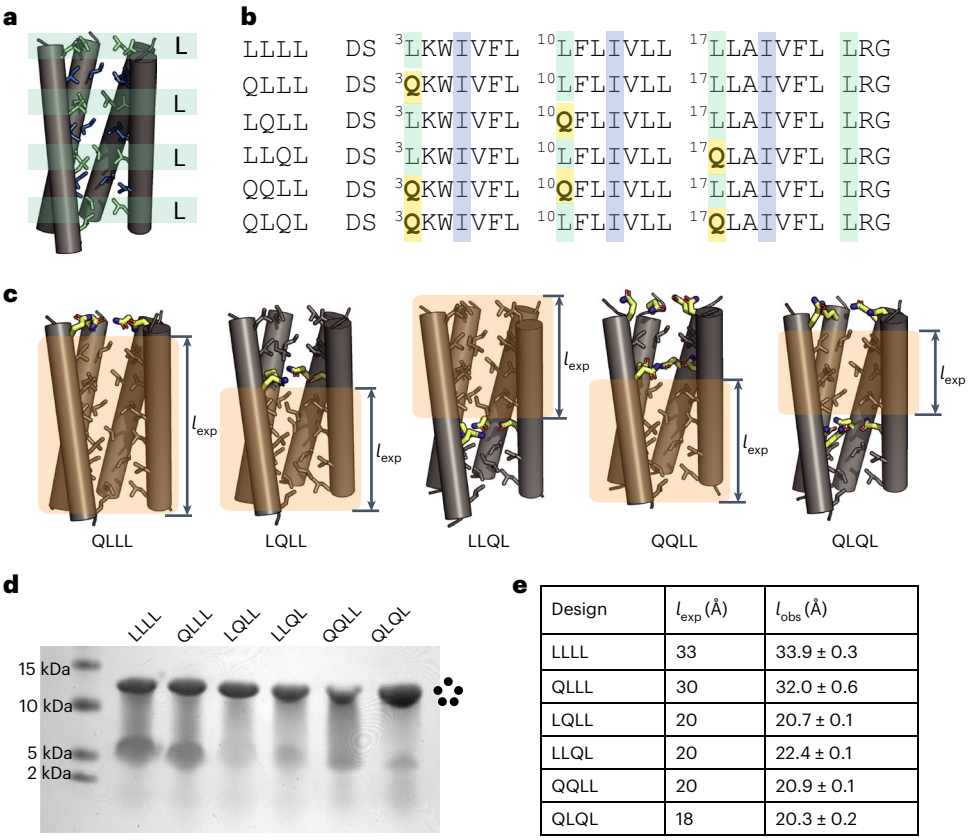


Fig. 1 | Hypothesis of proton-selective transport along transient water wires. **a**, Protons hop across dynamically rearranging hydrogen-bonded wires of water in the Grotthuss mechanism. **b**,**c**, In a wholly apolar channel, the pore remains devoid of water regardless of membrane polarization ($\Delta\Psi$) (**b**), but a polar PLS can mediate the flickering of neutral water in and out of relatively short apolar sectors of the channel (**c**). **d**, The presence of a hydrated excess proton facilitates water wire formation and transport of an excess proton through the hydrophobic sector by Grotthuss shuttling.

https://doi.org/10.1038/s41557-023-01210-4



e

Design	l_{exp} (Å)	l_{obs} (Å)
LLLL	33	33.9 ± 0.3
QLLL	30	32.0 ± 0.6
LQLL	20	20.7 ± 0.1
LLQL	20	22.4 ± 0.1
QQLL	20	20.9 ± 0.1
QLQL	18	20.3 ± 0.2

Fig. 2 | De novo channels incorporate PLSs at key positions to modulate hydrophobic lengths. a, The parent scaffold, LLLL (PDB: 6MCT), contains layers of interdigitating Leu (green) and Ile (blue) residues. **b**, Five Leu-to-Gln variants were designed. Gln residues are highlighted in yellow. **c**, Models of the designed channels illustrate the different expected hydrophobic lengths (l_{exp}) of the channels compared with the parent scaffold. **d**, SDS–PAGE revealed that these designs form stable pentamers (denoted by ring of five circles). **e**, Comparison of the longest hydrophobic lengths expected (l_{exp}) and observed (l_{obs}), determined by classical MD simulations (see Methods and Extended Data Table 1). The fifth helix has been removed from the side views of the models in Fig. 2 for clarity.

and Ile residues (Fig. 2a; Protein Data Bank (PDB): 6MCT)[43]. A narrow, fully hydrophobic pore (~2–3 Å diameter) that is impervious to water in classical MD simulations runs the entire length of the bundle[43]. We introduced single and double Leu-to-Gln substitutions into the pore (Fig. 2b–d), thereby creating a site that filled with water in MD simulations. The Leu residues were targeted for substitution because they project directly towards the interior of the pore without steric clashes. The variants are denoted as LLLL, QLLL, QQLL and so on, according to the nature of the Leu-to-Gln substitutions (Fig. 2). Seven variants with zero, one or two Gln substitutions were synthesized; six formed pentamers, based on gel electrophoresis, and were structurally and functionally characterized. By design, the expected length of the longest apolar path, l_{exp}, was intended to be 33 Å for LLLL and 30 Å for QLLL, and approximately 18–20 Å for the remaining four variants (Fig. 2c). Constructs with both single and double PLSs were evaluated by varying the positions of the introduced Gln residues. A peptide with three Gln residues failed to form a pentameric bundle, so it was not possible to further decrease the value of l_{exp}.

The hydration of each channel was evaluated in three independent 200 ns classical MD simulations of the design models in phospholipid bilayers (Fig. 3 and Supplementary Figs. 1–6). The pore of the starting LLLL pentamer (Fig. 3a) was devoid of considerable water density throughout the trajectories. The length of the longest apolar path in its pore (l_{obs} = 33.9 ± 0.3 Å; Fig. 2e and Extended Table 1) was in good agreement with the design (l_{exp} = 33 Å). QLLL (Fig. 3b) showed high water density close to its Gln side chain near the entrance to the pore, but the remainder of the pore was dry (l_{obs} = 32.0 ± 0.6 Å). The single-site variant LQLL (Fig. 3c) is strongly hydrated near the Gln carboxamide side chain, which communicates with the bulk water at the top of the channel via infrequent, flickering water wires. The remaining carboxy-terminal pore region is fully dehydrated over a length of ~21 Å (l_{obs} = 20.7 ± 0.1 Å). LLQL (Fig. 3d) shows inverse behaviour with hydrated Gln residues, fluctuating water wires near the bottom of the channel and a 22.4 ± 0.1 Å dehydrated pore at the top. The doubly substituted pentamers show greater hydration associated with the additional interfacial Gln (Fig. 3e,f). However, by design, they still have three consecutive layers of pore-lining Leu and Ile residues, which results in dry regions of similar length to the longest stretch seen in the single-Gln variants (l_{obs} = 20.9 ± 0.1 and 20.3 ± 0.2 Å for QQLL and QLQL, respectively). Thus, by incorporating classical MD simulations into the design process, we can assess the potential of design candidates for experimental validation.

Structure of the designed proton channels

The structures of the five pentameric Gln-containing pentamers were determined by X-ray crystallography. Although they crystallized in different space groups (Extended Data Table 2), their backbone structures were nearly identical, with the root mean squared deviation (RMSD) to the backbone of the starting LLLL pentamer ranging from 0.25 to 0.32 Å (Fig. 4a). The Gln side chains converge in layers near the centre of the bundle, where they form side chain–side chain and side chain–backbone hydrogen bonds (Fig. 4b–f and Supplementary Figs. 7–16). As expected, they also surround puncta of electron density, which were well modelled as water molecules with full occupancy (Extended Data Fig. 1). The Gln side chains are not fully symmetric and different conformations are seen in the individual monomeric units of the

https://doi.org/10.1038/s41557-023-01210-4

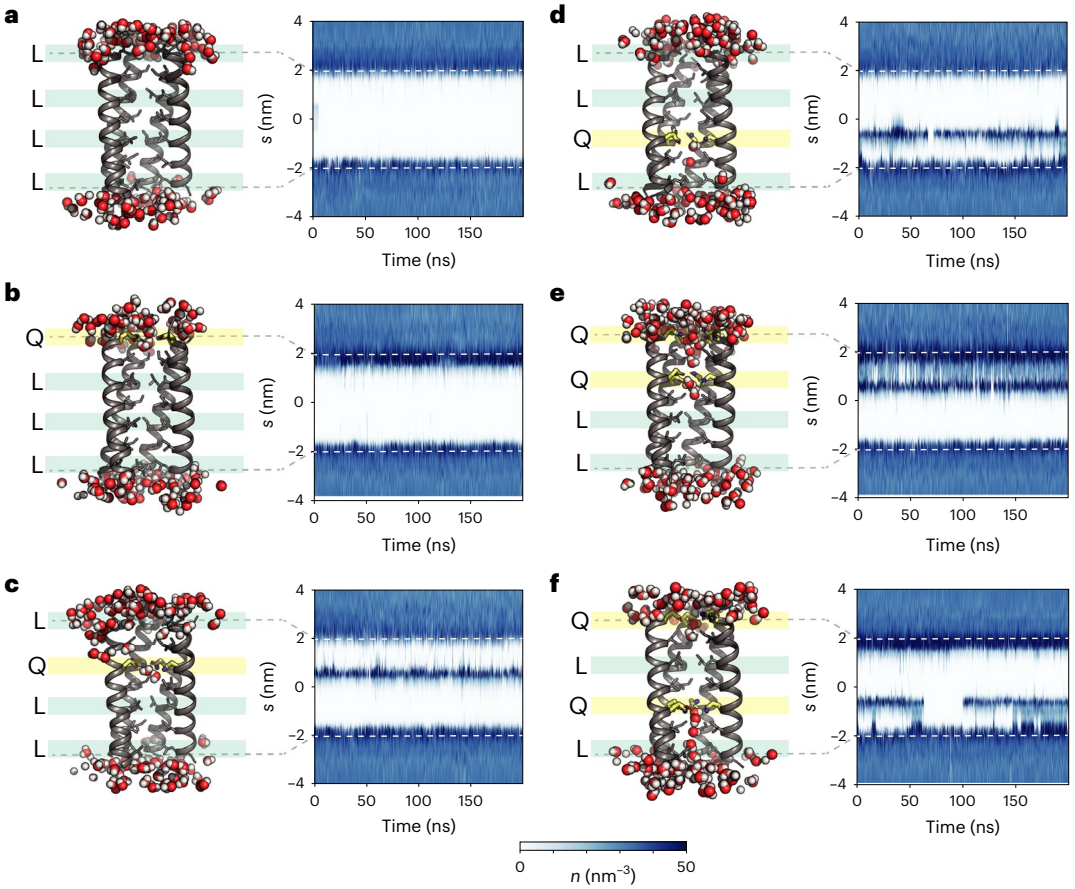


Fig. 3 | Permeation of water into the designed channels correlates with the position of the luminal Gln residues. a–f, The hydration of each channel, as evaluated by classical MD simulations and analysed using the Channel Annotation Package[44]. Snapshots from the MD simulation (left) and plots of pore water density (n, in nm^{-3}) as a function of position along the z-axis to the pore, s, over time (right) reveal no water permeation into the hydrophobic pore in LLLL (**a**) or QLLL (**b**). Pentamers with Gln in the second and third layers of the channel, namely LQLL (**c**), LLQL (**d**), QQLL (**e**) and QLQL (**f**), have high water density around the polar Gln site and flickering water molecules in the shorter apolar segment leading up to the mutation. However, the longer apolar segment remains dehydrated. Dashed lines illustrate location of the TM region of the channel in the water density plots. The fifth helix has been removed from all figures for clarity.

pentamers. They also have higher Debye–Waller factors than the surrounding main-chain and pore-lining side chains. These findings are in good agreement with the hydration of the Gln side chains observed in the MD simulations.

To confirm the hydration observed in the MD simulations and crystal structures, we examined the water-edited solid-state ^{13}C NMR spectra of a series of LLLL and LQLL peptides uniformly labelled with ^{13}C and ^{15}N at the Ile6 or Ile13 positions. As expected from the MD simulations, magnetization transfer from water was more efficient for Ile13 in LQLL than in LLLL for all atoms (Fig. 4g and Extended Data Fig. 2). Water magnetization transfer was observed for both the LLLL and LQLL channels because magnetization was also transferred from water at the membrane surface to the ^{13}C-labelled residues via relayed ^{1}H spin diffusion, in addition to the chemical exchange and nuclear Overhauser mechanisms. Thus, the water magnetization transfer observed for the LLLL channels can be attributed to the surface water, while the additional magnetization transfer for the LQLL channels indicates the presence of water molecules inside the channel pore. Experiments with labelled Gln10 also showed notable saturation transfer from water to the PLS (Fig. 4g). Thus, the dynamics of water observed for the pentamers in hydrated phospholipid bilayers is in good agreement with the MD simulations and crystal structures. However, all three of these methods probe equilibrium hydration in the absence of an excess proton in the pore. The Grotthuss proton hopping process requires consideration of covalent bond making and breaking as the proton moves from water molecule to water molecule. To simulate this process, we turned to RMD simulations.

Multiscale RMD simulations from structures

As noted earlier, while classical simulations are useful for evaluating the hydration of the channel in the absence of an excess proton, they do not account for changes in chemical bonding such as occurs in the Grotthuss shuttling mechanism caused by protons permeating the channel[8,10]. We therefore turned to multiscale RMD (MS-RMD) simulations, which explicitly simulate the entire dynamic trajectory of proton translocation, including the covalent bonding rearrangements and transfer of protons as they pass between water molecules[45–47]. While MS-RMD is substantially faster than explicit quantum mechanical (ab initio MD) calculations, the calculation of proton transport through an entire TM protein by MS-RMD is still computationally very intensive, so we confined our focus to a comparison of LLLL and LQLL based on their crystal structures. We used two collective variables (CVs) to enable enhanced free-energy sampling in the RMD simulations: the position of the 'centre of excess charge' (CEC), which tracks the translocation of a proton charge defect along the pore[45–47], and a water wire connectivity parameter, ϕ, which quantifies the number and connectivity of hydrogen-bonded waters in the pore and is also associated with the excess proton structure[8]. A value of $\phi = 0$ corresponds to a dry pore, and $\phi = 1$ to a pore with a fully connected water wire spanning the length of the pore and containing an excess proton.

https://doi.org/10.1038/s41557-023-01210-4

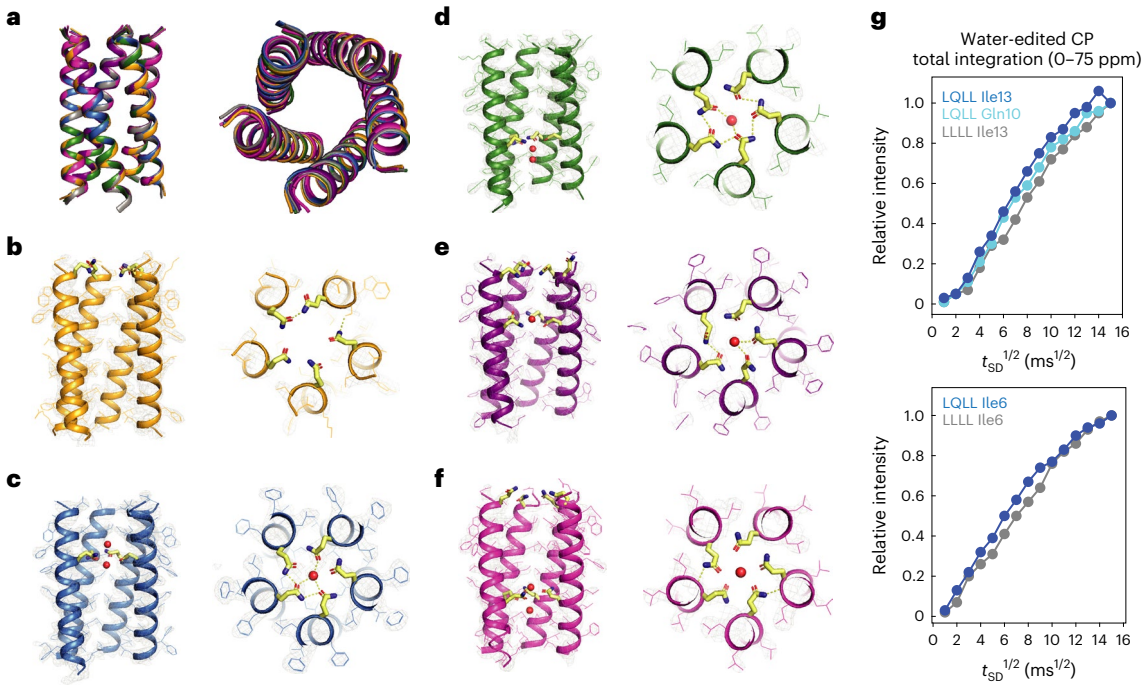


Fig. 4 | Crystal structures of the designed channels demonstrate that the introduction of polar Gln residues mediates stable water pockets. a–**f**, The X-ray structures of the designed channels are within <0.4 Å RMSD of the original design template with LLLL (PDB: 6MCT, grey) (**a**). Representative structures of the designed channels (left) and zoom-in of the Gln layer (right) for QLLL (PDB: 7UDY) (**b**), LQLL (PDB: 7UDZ) (**c**), LLQL (PDB: 7UDV) (**d**), QQLL (PDB: 7UDW) (**e**) and QLQL (PDB: 7UDX) (**f**). The fifth helix has been removed from all figures for clarity. **g**, Water build-up curves for uniformly labelled (^{13}C, ^{15}N) Ile13, Ile6 and Gln10 in the LLLL and LQLL peptides. Both the Ile6 and Ile13 sites in the LQLL sample show faster water build-up than the corresponding sites in the LLLL sample. t_{SD}, spin diffusion period.

The computed potentials of mean force (PMFs) for LLLL and LQLL predict that the saddle in the two-dimensional free-energy 'landscape' as a function of the two CVs for proton translocation is prohibitively high in LLLL, but greatly lowered through the introduction of the single Gln in LQLL (Fig. 5a,b and Supplementary Fig. 17). The landscape and its saddle region for LQLL (rescaled in Fig. 5c to allow easier viewing of its contours) describe the free energy of an excess proton at varying positions along the channel axis (*x* axis; Fig. 5a–c) and varying degrees of protonated water wire formation (*y* axis). The asymmetric location of the Gln-containing PLS in LQLL separates the overall channel into a 'short apolar path' near the amino terminus (right side in Fig. 5a,b and top in Fig. 5d) and a 'long apolar path' near the C terminus (left side in Fig. 5a,b and bottom in Fig. 5d). The lowest energy pathway from the exterior to the PLS through the short apolar path (points (1) to (3) to (4A) to (5) in Fig. 5c,d) features water wires that are fully formed through only this region of the channel and then 'flip' to translocate the proton on through the channel. The lower-free-energy saddle predicts rapid proton transport through the short apolar path. The arrival of a proton in the lower portion of the PLS was also seen to sometimes induce cooperative, fully formed water wires running through the entire long apolar path (points (1) to (3) to (4B) to (5) in Fig. 5c,d and Supplementary Fig. 18). The transmission of protons through this region was thus computed to proceed via two energetically similar pathways (Fig. 5c) that differ in whether water wires are present or absent in the shorter path as the proton moves through the longer path (that is, a fully connected water wire spanning the length of the channel). It is important to note that the broad saddle region along the (vertical) *y* axis (the water wire connectivity) relative to the depth and narrowness of the wells for the proton entry and exit regions (between points (4A) and (4B) in Fig. 5c) indicates a large positive entropy change for the excess proton moving into the saddle region. This is a highly favourable entropy of activation and so one should not interpret the 'barrier' to proton transport as simply arising from the one-dimensional trace along either one of the white curves in Fig. 5c, or from a particular point on the saddle of the two-dimensional free-energy surface in that figure. Instead, the overall proton transport rate must be considered as deriving from an average over many such paths crossing along that vertical saddle, resulting in a quite favourable proton translocation behaviour, as seen in the experiments. In summary, RMD simulations of LLLL and LQLL predict that the introduction of a Gln-rich PLS one-third of the way through the pore should dramatically enhance the transmission of protons via the formation of proton-induced water wires.

Function through proton flux measurements

To demonstrate the ability of these designed channels to transport protons, we experimentally measured the flux of protons driven by an electrical gradient. We employed a vesicle assay (adapted from refs. 48,49) that uses a chemi-osmotic-induced electrical potential to drive the carrier-mediated translocation of protons into phospholipid vesicles. In this assay, vesicles containing K^+ buffer are rapidly diluted into Na^+ buffer, creating a chemical potential across the bilayer. Valinomycin, a K^+ carrier that is highly selective for K^+ over Na^+, is then added to allow K^+ to diffuse down its chemical potential out of the vesicle, thereby creating a TM electrical potential (Fig. 6a and Extended Data Fig. 3). If a proton carrier or channel is present in the bilayer, protons will then follow the induced electrical potential and diffuse against a growing concentration gradient of protons into the vesicle, leading to a drop in the interior pH (pH_{in}), which is detected by a pH-sensitive fluorescent dye (Fig. 6a). It is noteworthy that this system requires that the vesicles be non-leaky and highly impermeable to Na^+, otherwise Na^+ will diffuse into the vesicle dissipating the electrical potential (Extended Data Fig. 3). Figure 6c shows proton conduction data for the well-characterized proton channel M2, a proton-selective viroporin from the influenza A virus[50]. Induction of an electrical gradient by

https://doi.org/10.1038/s41557-023-01210-4

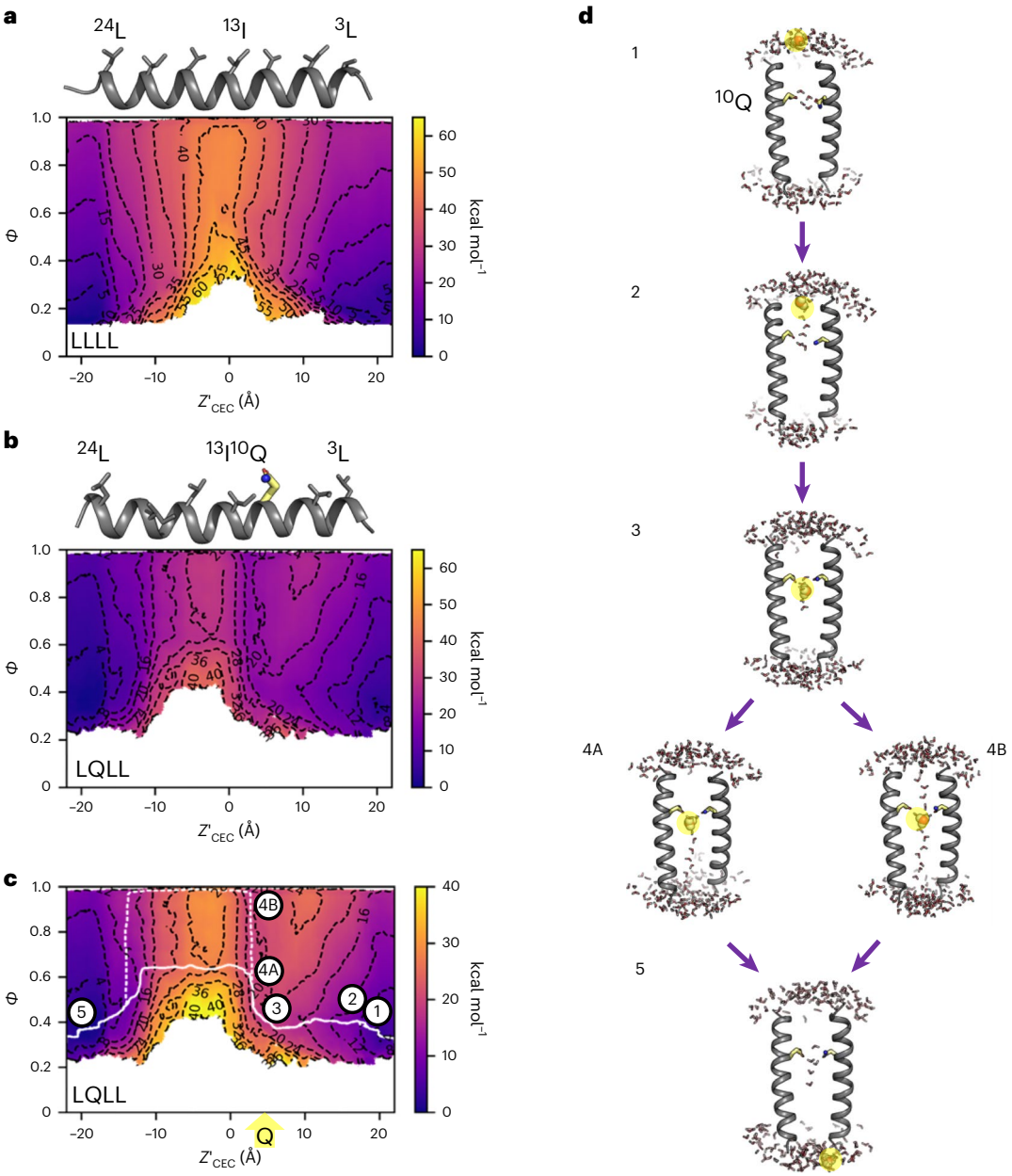


Fig. 5 | MS-RMD predicts that introduction of the PLS enables the formation of proton-conductive transient water wires. a, Two-dimensional PMF of LLLL shows a high barrier when the proton is at $Z'_{CEC} = 0$ Å, or the centre of mass of the channel at the Ile13 α-carbon atoms. **b**, Addition of the Gln residue at +4 Å in LQLL shifts the barrier to the C-terminal side of the channel and decreases the barrier height by ~20 kcal mol^{-1}. **c**, The two lowest mean free-energy paths (white solid and dashed lines), derived from string theory (see Methods) through the LQLL channel. Note the scale change on the colour bar in **b** compared with **c**. **d**, Snapshots along the two pathways for LQLL (shown in **c**) reveal the mechanism of proton-induced water wires mediating proton translocation. The most hydronium-like structures are highlighted in yellow. Only two helices are represented for clarity.

addition of valinomycin (at time $t = 120$ s) leads to a change in pH_{in}, as expected from the proton selectivity of M2, but no change in pH_{in} was observed for empty vesicles containing no protein (Fig. 6b). Following the addition of valinomycin, the protonophore, carbonyl cyanide *m*-chlorophenyl hydrazone (CCCP), was added (Extended Data Fig. 3); the observed additional bolus of proton flux assures that electrical and pH gradients have been maintained throughout the experiment, even in the order of hours (Extended Data Figs. 4–7 and Supplementary Figs. 19–23).

Using this assay, we determined that LLLL ($I_{obs} = 33.9 \pm 0.3$ Å) had no significant conductance over the background, and that QLLL ($I_{obs} = 32.0 \pm 0.6$ Å) had a very low conduction that was not significantly different from that of LLLL (Fig. 6g). In contrast, the remaining pentamers with apolar tracks of approximately 20 Å (LQLL, LLQL, QQLL and QLQL) showed significant proton conduction well above background, reaching rates on a par with M2 at equivalent symmetrical pH_{in} and pH_{out} and peptide/lipid ratios (Fig. 6g). These four functional proton channels showed very similar conductance rates (within a factor of two), consistent with their similar hydrophobic lengths. This finding was confirmed by performing six to nine independent replicates for each channel ($p < 0.0001$; Supplementary Table 3). Given that the pH is 7.5 ($[H^+] = 10^{-7.5}$ M) and $[Na^+]_{out}$ is 0.16 M at the beginning of the experiment, the selectivity of the channels for protons over Na^+ must be at least 10^6-fold.

https://doi.org/10.1038/s41557-023-01210-4

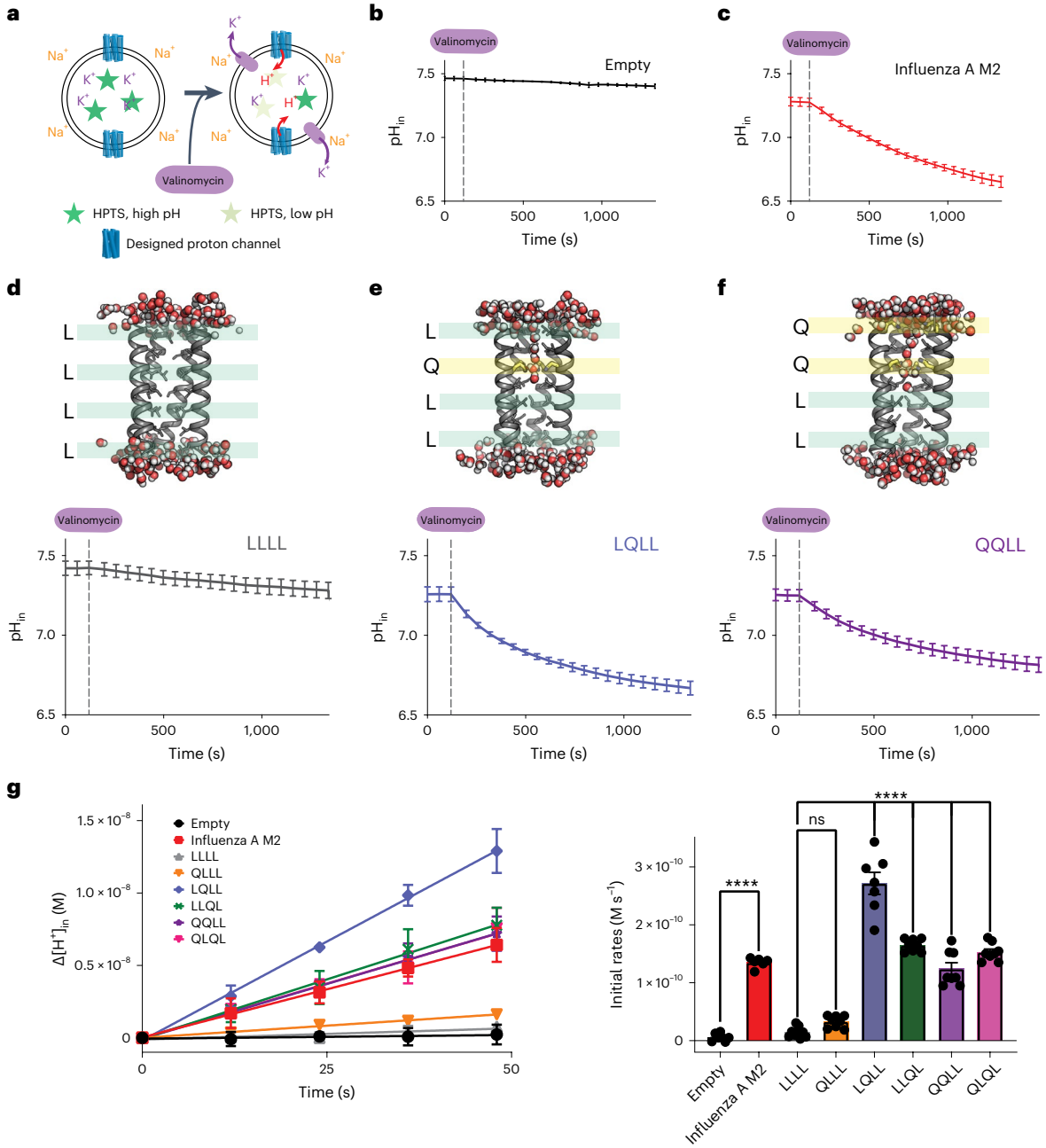


Fig. 6 | Designed channels selectively move protons across the membrane. **a**, Schematic for proton flux assays using a vesicle-entrapped pH-sensitive fluorescent dye, 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS). At $t = 0$ the pH is the same inside and outside of the vesicles. **b**, Following addition of valinomycin, the pH_{in} of empty vesicles does not change, because there is no proton channel or carrier included. **c**, When a proton channel is present, such as the influenza A M2 channel shown in this panel, addition of valinomycin enables the transport of protons down the electrochemical gradient created by the efflux of potassium. This results in a notable decrease in the pH_{in} over time as protons move into the vesicle up a concentration gradient. **d**–**f**, Snapshots from MD simulations (top) and proton flux assays (bottom) for LLLL (**d**), LQLL (**e**) and QQLL (**f**) indicate that addition of polar Gln near the middle of the channel enables water permeation events into the pore, which facilitates proton-selective transport. The fifth helix has been removed from **d**–**f** for clarity. **g**, Change in the H^+ concentration of representative samples upon addition of valinomycin at $t = 0$. For the control samples, influenza A M2 showed significant (****$P < 0.0001$) initial rates relative to Empty. QLLL initial rates are not significant (ns, $P = 0.3947$) relative to LLLL. The fitting of the initial rates shows that LQLL, LLQL, QQLL and QLQL have significant proton transport activity (****$P < 0.0001$) relative to LLLL. All data in **b**–**f** are presented as the mean ± SD for $n > 6$ independent experiments. The data in **g** are shown as the mean ± SD (left) and the mean ± SEM (right) of the independent experiments (black circles). The data were analysed using the unpaired t-test for empty versus influenza A M2, and ordinary one-way analysis of variance for multiple comparisons (Dunn's test) of LLLL versus the designed channels (see Supplementary Tables 4 and 5 for the complete analysis).

Discussion

We have successfully engineered minimalist proton channels that have enabled us to explore the roles that transient water wires through apolar regions play in proton selectivity and conduction. In keeping with our minimalist strategy, a ring of Gln side chains was chosen as the PLS. In future work, ionizable residues in the pore would allow for pH and metal ion[51,52] gating and modulation of the proton conductance. Moreover, although the peptide pentamers were randomly distributed in the

membrane in our preparations (Extended Data Fig. 8), methods to allow unidirectional insertion of the channels would provide information on the symmetry of conduction. Previous theoretical work has predicted that the arrival of a proton at a loading site can trigger transient water wires that span surprisingly long distances[11], which we have shown here can be as long as ~20 Å. While this was predicted by MS-RMD, our classical MD simulations, which did not explicitly consider a mobile Grotthuss shuttling excess proton, predicted that water wires could only span a length of around 10 Å from the exterior bulk water to a position within the Gln pentad. Thus, while classical simulations are quite useful for identifying pre-existing water wires that occasionally flicker on in the neutral state (Fig. 1c), they are intrinsically less suited to identifying water wires that are induced as a proton enters the channel (Fig. 1d). The M2 proton channel from the influenza A virus[53,54] and the voltage-gated proton channel Hv1 (refs. 55–57) use electrical or chemical gradients to drive vectoral Grothuss proton conduction across apolar constrictions within channels. In other cases, such as bacteriorhodopsin[13,21–23], light absorption leads to the release of protons that induce the formation of water wires. Similarly, proton transport to the PLS of cytochrome *c* oxidase traverses an apolar region[58–62]; classical and MS-RMD simulations revealed that transient water structures facilitate this important step of this proton pump[63,64]. In each case, proton conduction occurs along water wires that are impervious to larger ions. By contrast, many channels that conduct larger ions have apolar pores that are dry in the 'off' state but become substantially larger and hydrated in response to a larger conformational gating transition[16–19,65].

During proton conduction, each molecule of water in a water wire changes the direction of its dipole as protons transfer from one molecule to the next. The orientation of the water molecules must be reset to regenerate the initial polarization before a new proton can be transported. Water wires in restricted hydrophobic environments are well suited to stabilizing both polarizations because they do not form strong dipolar interactions with the pore that would bias their orientation. Also, the Gln side chains are relatively mobile in our crystal structures, as indicated by the Debye–Waller factors and MD and MS-RMD simulations (Supplementary Fig. 24), which show rapid conformational fluctuations. This behaviour contrasts with the requirements for water channels such as aquaporins, which feature stable water-binding sites with strong polarization that undermine the orientations of the hydrogen-bonded networks and dipolar switching that would otherwise enable proton-selective transport[66,67].

In this work it was necessary to move beyond static structures to include the dynamic processes required for proton migration, including both protein and water dynamics, which de novo design generally ignores in favour of structural stability and computational efficiency. MS-RMD allowed even deeper consideration of the bond-making and -breaking steps required for Grotthuss proton migration through water wires. While computational speed is currently too low to incorporate RMD into the early stages of protein design, it can clearly provide an important filter to assess potential designs. Indeed, Mondal and co-workers have used an empirical valence bond approach to screen combinatorial libraries of enzyme variants[68].

Our work also highlights enabling design principles for the development of new proton-conductive materials. We have shown that the close positioning of a PLS proximal to a dry pore results in channels that are highly selective for protons. Our minimalist designs show that the PLS need not be elaborate in design. Indeed, this work was inspired in part by experimental and computational studies of hard materials composed of carbon nanotubes[69,70]. Our design principles also have implications for the design of soft-material proton-selective membranes. For example, Jiang and co-workers have designed proton-selective copolymers consisting of apolar segments interspersed with occasional polar ethylene glycol units[71], showing similarities to the two-component design of our channels. Our current work extends a computational approach to mechanistically interrogate and design materials with even greater efficiency and selectivity. Indeed, because our designs are based on fundamental physical chemical principles and molecular rather than bioinformatic algorithms, they are not limited to the production of natural proteins or synthetic peptides. Instead, they should be able to translate to the design of novel non-proteinaceous molecular assemblies and polymers for applications ranging from water purification to energy storage and use.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41557-023-01210-4.

References

1. Moriyama, Y. & Futai, M. H^+-ATPase, a primary pump for accumulation of neurotransmitters, is a major constituent of brain synaptic vesicles. *Biochem. Biophys. Res. Commun.* **173**, 443–448 (1990).
2. Nishi, T. & Forgac, M. The vacuolar (H^+)-ATPases—nature's most versatile proton pumps. *Nat. Rev. Mol. Cell Biol.* **3**, 94–103 (2002).
3. Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144–148 (1961).
4. Nicholls, D. G. Mitochondrial ion circuits. *Essays Biochem.* **47**, 25–35 (2010).
5. Diering, G. H. & Numata, M. Endosomal pH in neuronal signaling and synaptic transmission: role of Na^+/H^+ exchanger NHE5. *Front. Physiol.* **4**, 412 (2014).
6. Agmon, N. The Grotthuss mechanism. *Chem. Phys. Lett.* **244**, 456–462 (1995).
7. Calio, P. B., Li, C. & Voth, G. A. Resolving the structural debate for the hydrated excess proton in water. *J. Am. Chem. Soc.* **143**, 18672–18683 (2021).
8. Li, C. & Voth, G. A. A quantitative paradigm for water-assisted proton transport through proteins and other confined spaces. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.2113141118 (2021).
9. Wraight, C. A. Chance and design—proton transfer in water, channels and bioenergetic proteins. *Biochim. Biophys. Acta* **1757**, 886–912 (2006).
10. Decoursey, T. E. Voltage-gated proton channels and other proton transfer pathways. *Physiol. Rev.* **83**, 475–579 (2003).
11. Peng, Y., Swanson, J. M., Kang, S. G., Zhou, R. & Voth, G. A. Hydrated excess protons can create their own water wires. *J. Phys. Chem. B* **119**, 9212–9218 (2015).
12. Banh, R. et al. Hydrophobic gasket mutation produces gating pore currents in closed human voltage-gated proton channels. *Proc. Natl Acad. Sci. USA* **116**, 18951–18961 (2019).
13. Garczarek, F. & Gerwert, K. Functional waters in intraprotein proton transfer monitored by FTIR difference spectroscopy. *Nature* **439**, 109–112 (2006).
14. Kaur, D., Khaniya, U., Zhang, Y. & Gunner, M. R. Protein motifs for proton transfers that build the transmembrane proton gradient. *Front. Chem.* **9**, 660954 (2021).
15. Kalra, A., Garde, S. & Hummer, G. Osmotic water transport through carbon nanotube membranes. *Proc. Natl Acad. Sci. USA* **100**, 10175–10180 (2003).
16. Ben-Abu, Y., Zhou, Y., Zilberberg, N. & Yifrach, O. Inverse coupling in leak and voltage-activated K^+ channel gates underlies distinct roles in electrical signaling. *Nat. Struct. Mol. Biol.* **16**, 71–79 (2009).
17. Jensen, M. O. et al. Principles of conduction and hydrophobic gating in K^+ channels. *Proc. Natl Acad. Sci. USA* **107**, 5833–5838 (2010).

18. Aryal, P., Sansom, M. S. & Tucker, S. J. Hydrophobic gating in ion channels. *J. Mol. Biol.* **427**, 121–130 (2015).
19. Zhu, F. & Hummer, G. Drying transition in the hydrophobic gate of the GLIC channel blocks ion conduction. *Biophys. J.* **103**, 219–227 (2012).
20. Rasaiah, J. C., Garde, S. & Hummer, G. Water in nonpolar confinement: from nanotubes to proteins and beyond. *Annu. Rev. Phys. Chem.* **59**, 713–740 (2008).
21. Wang, T. et al. Deprotonation of D96 in bacteriorhodopsin opens the proton uptake pathway. *Structure* **21**, 290–297 (2013).
22. Weinert, T. et al. Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography. *Science* **365**, 61–65 (2019).
23. Freier, E., Wolf, S. & Gerwert, K. Proton transfer via a transient linear water-molecule chain in a membrane protein. *Proc. Natl Acad. Sci. USA* **108**, 11435–11439 (2011).
24. Regan, L. & DeGrado, W. F. Characterization of a helical protein designed from first principles. *Science* **241**, 976–978 (1988).
25. Walsh, S. T., Cheng, H., Bryson, J. W., Roder, H. & DeGrado, W. F. Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proc. Natl Acad. Sci. USA* **96**, 5486–5491 (1999).
26. Kuhlman, B. et al. Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364–1368 (2003).
27. Vorobieva, A. A. et al. De novo design of transmembrane β barrels. *Science* https://doi.org/10.1126/science.abc8182 (2021).
28. Yang, C. et al. Bottom-up de novo design of functional proteins with complex structural features. *Nat. Chem. Biol.* **17**, 492–500 (2021).
29. Polizzi, N. F. & DeGrado, W. F. A defined structural unit enables de novo design of small-molecule-binding proteins. *Science* **369**, 1227–1233 (2020).
30. Cao, L. et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
31. Fleishman, S. J. et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816–821 (2011).
32. Jiang, L. et al. De novo computational design of retro-aldol enzymes. *Science* **319**, 1387–1391 (2008).
33. Lassila, J. K., Privett, H. K., Allen, B. D. & Mayo, S. L. Combinatorial methods for small-molecule placement in computational enzyme design. *Proc. Natl Acad. Sci. USA* **103**, 16710–16715 (2006).
34. Polizzi, N. F. et al. De novo design of a hyperstable non-natural protein–ligand complex with sub-A accuracy. *Nat. Chem.* **9**, 1157–1164 (2017).
35. Leaver-Fay, A. et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* **487**, 545–574 (2011).
36. Koga, N. et al. Principles for designing ideal protein structures. *Nature* **491**, 222–227 (2012).
37. Scott, A. J. et al. Constructing ion channels from water-soluble α-helical barrels. *Nat. Chem.* **13**, 643–650 (2021).
38. Xu, C. et al. Computational design of transmembrane pores. *Nature* **585**, 129–134 (2020).
39. Joh, N. H. et al. De novo design of a transmembrane Zn^{2+}-transporting four-helix bundle. *Science* **346**, 1520–1524 (2014).
40. Lu, P. et al. Accurate computational design of multipass transmembrane proteins. *Science* **359**, 1042–1046 (2018).
41. Thomaston, J. L. et al. X-ray crystal structure of the influenza A M2 proton channel S31N mutant in two conformational states: an open and shut case. *J. Am. Chem. Soc.* **141**, 11481–11488 (2019).
42. Saotome, K. et al. Structures of the otopetrin proton channels Otop1 and Otop3. *Nat. Struct. Mol. Biol.* **26**, 518–525 (2019).
43. Mravic, M. et al. Packing of apolar side chains enables accurate design of highly stable membrane proteins. *Science* **363**, 1418–1423 (2019).
44. Klesse, G., Rao, S., Sansom, M. S. P. & Tucker, S. J. CHAP: a versatile tool for the structural and functional annotation of ion channel pores. *J. Mol. Biol.* **431**, 3353–3365 (2019).
45. Lee, S., Liang, R., Voth, G. A. & Swanson, J. M. Computationally efficient multiscale reactive molecular dynamics to describe amino acid deprotonation in proteins. *J. Chem. Theory Comput.* **12**, 879–891 (2016).
46. Knight, C., Lindberg, G. E. & Voth, G. A. Multiscale reactive molecular dynamics. *J. Chem. Phys.* **137**, 22A525 (2012).
47. Yamashita, T., Peng, Y., Knight, C. & Voth, G. A. Computationally efficient multiconfigurational reactive molecular dynamics. *J. Chem. Theory Comput.* **8**, 4863–4875 (2012).
48. Moffat, J. C. et al. Proton transport through influenza A virus M2 protein reconstituted in vesicles. *Biophys. J.* **94**, 434–445 (2008).
49. Ma, C. et al. Identification of the functional core of the influenza A virus A/M2 proton-selective ion channel. *Proc. Natl Acad. Sci. USA* **106**, 12283–12288 (2009).
50. Leiding, T., Wang, J., Martinsson, J., DeGrado, W. F. & Arskold, S. P. Proton and cation transport activity of the M2 proton channel from influenza A virus. *Proc. Natl Acad. Sci. USA* **107**, 15409–15414 (2010).
51. Slope, L. N. & Peacock, A. F. De novo design of xeno-metallo coiled coils. *Chem. Asian J.* **11**, 660–666 (2016).
52. Pinter, T. B. J., Koebke, K. J. & Pecoraro, V. L. Catalysis and electron transfer in de novo designed helical scaffolds. *Angew. Chem. Int. Ed.* **59**, 7678–7699 (2020).
53. Khurana, E. et al. Molecular dynamics calculations suggest a conduction mechanism for the M2 proton channel from influenza A virus. *Proc. Natl Acad. Sci. USA* **106**, 1069–1074 (2009).
54. Yi, M., Cross, T. A. & Zhou, H. X. A secondary gate as a mechanism for inhibition of the M2 proton channel by amantadine. *J. Phys. Chem. B* **112**, 7977–7979 (2008).
55. Ramsey, I. S. et al. An aqueous H^+ permeation pathway in the voltage-gated proton channel Hv1. *Nat. Struct. Mol. Biol.* **17**, 869–875 (2010).
56. Chamberlin, A. et al. Hydrophobic plug functions as a gate in voltage-gated proton channels. *Proc. Natl Acad. Sci. USA* **111**, E273–E282 (2014).
57. Takeshita, K. et al. X-ray crystal structure of voltage-gated proton channel. *Nat. Struct. Mol. Biol.* **21**, 352–357 (2014).
58. Wikstrom, M., Krab, K. & Sharma, V. Oxygen activation and energy conservation by cytochrome c oxidase. *Chem. Rev.* **118**, 2469–2490 (2018).
59. Hofacker, I. & Schulten, K. Oxygen and proton pathways in cytochrome c oxidase. *Proteins* **30**, 100–107 (1998).
60. Wikström, M., Verkhovsky, M. I. & Hummer, G. Water-gated mechanism of proton translocation by cytochrome c oxidase. *Biochim. Biophys. Acta Bioenerg.* **1604**, 61–65 (2003).
61. Tashiro, M. & Stuchebrukhov, A. A. Thermodynamic properties of internal water molecules in the hydrophobic cavity around the catalytic center of cytochrome c oxidase. *J. Phys. Chem. B* **109**, 1015–1022 (2005).
62. Goyal, P., Lu, J., Yang, S., Gunner, M. R. & Cui, Q. Changing hydration level in an internal cavity modulates the proton affinity of a key glutamate in cytochrome c oxidase. *Proc. Natl Acad. Sci. USA* **110**, 18886–18891 (2013).
63. Liang, R., Swanson, J. M. J., Wikstrom, M. & Voth, G. A. Understanding the essential proton-pumping kinetic gates and decoupling mutations in cytochrome c oxidase. *Proc. Natl Acad. Sci. USA* **114**, 5924–5929 (2017).
64. Liang, R., Swanson, J. M., Peng, Y., Wikstrom, M. & Voth, G. A. Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase. *Proc. Natl Acad. Sci. USA* **113**, 7420–7425 (2016).

65. Lynch, C. I., Rao, S. & Sansom, M. S. P. Water in nanopores and biological channels: a molecular simulation perspective. *Chem. Rev.* https://doi.org/10.1021/acs.chemrev.9b00830 (2020).
66. Chen, H. et al. Charge delocalization in proton channels, I: the aquaporin channels and proton blockage. *Biophys. J.* **92**, 46–60 (2007).
67. Murata, K. et al. Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599–605 (2000).
68. Mondal, D., Kolev, V. & Warshel, A. Combinatorial approach for exploring conformational space and activation barriers in computer-aided enzyme design. *ACS Catal.* **10**, 6002–6012 (2020).
69. Tunuguntla, R. H., Allen, F. I., Kim, K., Belliveau, A. & Noy, A. Ultrafast proton transport in sub-1-nm diameter carbon nanotube porins. *Nat. Nanotechnol.* **11**, 639–644 (2016).
70. Geng, J. et al. Stochastic transport through carbon nanotubes in lipid bilayers and live cell membranes. *Nature* **514**, 612–615 (2014).
71. Jiang, T. et al. Single-chain heteropolymers transport protons selectively and rapidly. *Nature* **577**, 216–220 (2020).

Methods

Synthesis of proton channel peptides

The proton channel peptides (Fig. 2 and Supplementary Table 1) were synthesized using 9-fluorenylmethoxycarbonyl-based solid-phase peptide synthesis on either a Biotage Initiator Alstra microwave synthesizer or a Syro II parallel peptide synthesizer. All peptides were synthesized with a free amine N terminus and a C-terminal carboxamide using Tentagel S-RAM resin (Chem-Impex) with 0.22–0.24 mmol loading per g resin. Following their synthesis, the peptides were cleaved from the resin with a solution of trifluoroacetic acid (TFA)–triisopropylsilane–water (95:2.5:2.5), precipitated with cold diethyl ether, redissolved in 50% 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) and water, and purified using reversed-phase HPLC on a C4 preparation column (Vydac) with a gradient of solvents A (water containing 0.1% TFA) and B (isopropanol–acetonitrile–water–TFA, 60:30:9.9:0.1) at a flow rate of 10 ml min^{-1}. The peptides were then lyophilized until dry. Following successful purification, the peptides were confirmed to be >90% pure by analytical HPLC over a C4 column (Phenomenex) and confirmed to have the correct mass by matrix-assisted laser desorption/ionization–time of flight mass spectrometry (Shimadzu AXIMA Performance) using α-cyano-4-hydroxycinnamic acid (Sigma) as the matrix (Supplementary Table 1). The peptides were then dissolved in ethanol to the appropriate stock concentrations and subsequently used in all experiments as described.

Expression of full-length influenza A M2 polypeptide

The influenza A M2 polypeptide from previous work[49] was expressed and purified from *Escherichia coli* cells. Briefly, the gene containing the A/Udorn/72 Cys-free W15F variant (W15F, C17S, C19S and C50S) and a C-terminal 6× His tag was cloned into a pEXP-5-NT plasmid and transformed into chemically competent *E. coli* BL21(DE3) cells (Invitrogen) using heat shock methods. The cells were grown in terrific broth medium (Invitrogen) with ampicillin at 37 °C. When an optical density value of 0.6–0.8 was achieved at a wavelength of 600 nm, the expression culture was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside. The cells were then collected no more than 2.5 h after induction and centrifuged at 4 °C for 10 min. The pellets were then frozen before purification. For purification, the cells were resuspended in lysis buffer (50 mM Tris pH 8, 300 mM NaCl, 5% glycerol, 25 mM imidazole, 0.1% decyl maltose neopentyl glycol and 8 M urea) and sonicated with a microtip sonicator for 5 min at 20% amplitude (2 s on/off). The lysed cells were then spun for 15 min at 18,000 r.p.m. (~41,000*g*) and purified using nickel (II) nitriloacetate beads using batch purification methods. The wash buffer was the same as the lysis buffer. The protein was eluted in elution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5% glycerol, 250 mM imidazole, 0.1% decyl maltose neopentyl glycol and 8 M urea) and mixed with HFIP before loading onto the C4 column for secondary purification by HPLC. The polypeptide was lyophilized until dry and reconstituted in ethanol to the appropriate stock concentrations. SDS–PAGE and matrix-assisted laser desorption/ionization–time of flight mass spectrometry analyses confirmed the expression of the polypeptide.

SDS–PAGE gels of proton channels

First, 25–50 μg peptide was lyophilized from the peptide stocks and dissolved in a solution of 25 mM Tris pH 7.5 (Thermo Fisher) with 50 mM octyl-β-glucopyranoside (OG; Carbosynth) for a final concentration of 2.5 μg $μl^{-1}$. An equal volume of 2× lithium dodecyl sulfate loading buffer (Invitrogen) was added to each sample and the samples were subsequently boiled for 15 min at 95 °C. Then, 5 μg peptide was loaded into each well of a 12% Bis-Tris NuPAGE gel (Thermo Fisher) with Precision Plus Protein Dual Xtra pre-stained protein standards (Bio-Rad). The gel was subsequently allowed to run for 30 min at 200 V.

Proton channel liposomal flux assays and analysis

The proton channel liposomal flux assays were adapted from refs. 48,49; the exact methodologies for the preparation and performance of these assays are described below.

Preparation of proteoliposomes. All proteoliposome samples were prepared as stock solutions of 1:500 peptide/lipids with final concentrations of 10 μM and 5 mM, respectively. To prepare the proteoliposome stock, 10 μl of 100 mM of a ratiometric pH-sensitive dye, the HPTS trisodium salt, was added to a total of 2 ml K^+ buffer (50 mM K_2SO_4 and 30 mM K_2HPO_4, pH 7.5). The solution was then added to a dried film of 3:1 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC)/1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (Avanti Lipids) to yield a 10 mM lipid solution, which was subsequently vortexed for 1 min before sonicating with a microtip sonicator for 5 min at 20% amplitude for 2 s on/off. The sonicated solution was then divided into 200 μl aliquots to which 1 M OG was added to a final concentration of 26 mM OG. After 1 h on the rotisserie at room temperature, the peptide in 26 mM OG in K^+ buffer (or, in the case of empty, 15 μl of 26 mM OG in K^+ buffer) was added to the solution of detergent-solubilized lipids for a final concentration of 20 μM peptide. The peptide–detergent–lipid (PDL) solution was equilibrated for 1 h at room temperature on the rotisserie. Following incubation of the PDL solution, 50 μl of an XAD biobead solution (Sigma-Aldrich) in K^+ buffer was added to the PDL mixture every 5–10 min five times for a final volume of around 500 μl and final peptide and lipid concentrations as described above. After the final addition of biobeads, the solution was left to rotate on the rotisserie overnight at 4 °C. The following morning, the samples were spun in an ultracentrifuge at 96,000 r.p.m. (~328,000*g*) for 10 min to pellet the liposomes. The liposome pellets were resuspended in the same amount of dye-free K^+ buffer to afford the stock solution used for the liposomal assays. Dynamic light scattering (Malvern Panalytical) of the vesicle samples enabled their size to be determined before proton flux assays to ensure liposome formation.

Preparation of working solutions for proton flux assays. A working 3 μM valinomycin solution was prepared by dissolving 1 mg ml^{-1} (~900 mM) valinomycin stock in dimethylsulfoxide (Sigma-Aldrich) in Na^+ buffer (50 mM Na_2SO_4 and 30 mM Na_2HPO_4, pH 7.5). Similarly, a 3 μM CCCP working solution was generated by dissolving 600 μM CCCP stock in dimethylsulfoxide in Na^+ buffer.

For the assay, 17.3 μl of the proteoliposome stock solution was diluted into 632.7 μl of Na^+ buffer mixed with 10 mM *p*-xylene-bis-pyridinium bromide (Invitrogen) for a total 650 μl of sample at working concentrations. The presence of membrane-impermeable *p*-xylene-bis-pyridinium bromid quenches any extraliposomal HPTS fluorescence. Therefore, the fluorescence measured in this experiment comes only from intraliposomal HPTS. The sample was allowed to equilibrate for 20 min before 190 μl was aliquoted into three wells on a black, U-shaped-bottom 96-well plate (Greiner) for fluorescence measurements (collected every 55 s).

Instrumentation. All data reported in Fig. 6, Extended Data Figs. 5–7 and Supplementary Figs. 19–23 were collected using a Biotek Synergy 2 microplate reader equipped with 405 nm/20 nm and 460 nm/40 nm bandwidth excitation filters and a 528 nm/20 nm bandwidth emission filter (Biotek). Valinomycin and CCCP solutions (10 μl each) were added to the wells using an injector system at 120 and 1,350 s into the assay, respectively, with 3 s of vigorous shaking following each addition. The fluorescence was measured at the two excitation wavelengths every 12 s.

For the long-time kinetics experiments (Extended Data Fig. 4), the data were collected on a Biotek Synergy H1 microplate reader using monochromators. The excitation wavelengths were set at 417 and 460 nm, and the emission wavelength was set to 515 nm.

Calibration and analysis of flux data. For all the data reported in Fig. 6, Extended Data Figs. 5–7 and Supplementary Figs. 19–23, the fluorescence signals were calibrated using the ratio of the fluorescence signals of the deprotonated dye, F^- ($\lambda_{excitation}$ = 460 nm, $\lambda_{emission}$ = 528 nm), and the isosbestic point, F_{iso} ($\lambda_{excitation}$ = 405 nm, $\lambda_{emission}$ = 528 nm), as a function of pH from pH 8 to pH 4 (Extended Data Fig. 3). Similarly, for the long-time kinetics experiments, a calibration curve was prepared from the ratio of the fluorescence signals of F^- ($\lambda_{excitation}$ = 460 nm, $\lambda_{emission}$ = 515 nm) and F_{iso} ($\lambda_{excitation}$ = 417 nm, $\lambda_{emission}$ = 515 nm) as a function of pH (Extended Data Fig. 1). Fluorescence measurements were performed in triplicate for each reported sample and then converted into pH_{in} values using the defined calibration curves.

For the data reported in Fig. 6, Extended Data Figs. 5–7 and Supplementary Figs. 19–23, the initial rates of proton conduction were determined for each sample using the data collected in the first 60 s after the addition of valinomycin. Similarly, for the long-time kinetics experiments (Extended Data Fig. 4), the initial rates were determined by fitting the data collected in the first 220 s after the addition of valinomycin. All data were fitted using linear regression methods and the initial rates were extracted from the fitted slope.

Determination of peptide orientation in vesicles

Liposomes containing 108 μM peptide (peptide/lipid ratio of 1:50) were used for the following experiment. Calibration curves were prepared as follows: 75 μl of liposome solution was mixed with 37.5 μl of 1 M OG and 37.5 μl of water (for non-reacted samples) or 37.5 μl of 20 mM methyltetrazine-sulfo-N-hydroxysuccinimide ester (Click Chemistry Tools) in water (for reacted samples) for 30 min. The samples were then quenched with 16.67 μl of 1 M Tris pH 8 for 5 min. The non-reacted and reacted samples were mixed in differing ratios and diluted in 50% HFIP–water before running on reversed-phase HPLC on an analytical C4 column with a gradient of solvents A (water containing 0.1% TFA) and B (isopropanol–acetonitrile–water–TFA, 60:30:9.9:0.1) at a flow rate of 1 ml min^{-1}. The areas under the curves of the non-reacted and reacted peaks were calculated for each sample to generate the calibration curve (Extended Data Fig. 8). The traces in the figures are background-subtracted and offset for ease of viewing.

To determine the orientation of samples, 50 μl of each liposome sample was mixed with 50 μl of 10 mM methyltetrazine sulfo-N-hydroxysuccinimide ester in water for 30 min and subsequently quenched with 5.56 μl of 1 M Tris pH 8 for 5 min. The samples were then diluted with 50:50 1 M OG and HFIP before running on HPLC. The areas under the curves of the non-reacted and reacted peaks were calculated and compared with the calibration curves to estimate the orientation of the peptide (Extended Data Fig. 8d,e). All samples were measured in duplicate.

Lipidic cubic phase crystallography of proton channels

The methods used for lipidic cubic phase (LCP) crystallization were originally described in detail by Caffrey and Cherezov and have been slightly modified as follows[72,73]. The peptide from the ethanol stock solution was mixed with 60 mg monoolein (Sigma-Aldrich) until a clear solution was obtained. The solution containing peptide and monoolein was then dried under a gentle stream of N_2 and lyophilized overnight. To prepare the LCP, the monoolein–peptide mixture was heated at 42 °C until it became liquid, and subsequently extruded ~200 times with two-thirds times the volume of 50 mM OG in coupled gas-tight Hamilton syringes at room temperature. Successful LCP formation was confirmed as the solution became clear and was not birefringent in the cross-polarizer. The final concentration of peptide for each sample is listed in Supplementary Table 2.

For crystallization, 50 nl of the LCP mixture was dispensed onto 96-well Laminex plastic sandwich plates (Molecular Dimensions) with 0.5–1 μl of precipitant solution using the TTP Labtech LCP Mosquito robot at room temperature. The plates were sealed with plastic coverslips (Molecular Dimensions) and monitored using a Formulatrix RockImager at 20 °C. Crystals of each construct were harvested under the conditions noted in Supplementary Table 2, cryoprotected with 30% (v/v) polyethylene glycol 400 if necessary, and flash-frozen in liquid nitrogen.

X-ray diffraction data collection and analysis

Crystals were mounted under a cryostream at 100 K and data were collected at both the 8.3.1 beamline at the Advanced Light Source at Lawrence Berkeley National Laboratory at a wavelength of 1.115832 Å or at the 23-ID-B and 23-ID-D beamlines at Argonne National Laboratory at a wavelength of 1.03318 Å. The data were processed using the XDS package[74] and reduced with AIMLESS within the CCP4 suite[75]. The structures were determined by molecular replacement with Phaser[76] using a previously designed de novo protein (PDB: 6MCT) as a model. The models were rebuilt in Coot[77] and the structures were then refined using PHENIX[78]. The data processing and structural refinement statistics are reported in Extended Data Table 2.

Solid-state NMR experiments

The LQLL and LLLL peptides with the appropriate site-specific ^{13}C and ^{15}N labels at the Ile6, Ile13 or Gln10 position were prepared by solid-phase peptide synthesis methods as described previously. The purified peptides were reconstituted into d54-DMPC lipids at a protein/lipid molar ratio of 1:12 with a final peptide mass of ~5 mg. The samples were concentrated to a hydration level of ~40% (wt/wt) and pelleted into 3.2 mm rotors for NMR experiments.

Magic-angle-spinning (MAS) solid-state NMR spectra were measured on a Bruker Avance III HD 600 MHz (14.1 T) spectrometer with a Bruker 3.2 mm HXY MAS probe operating in double-resonance $^1H/^{13}C$ mode. All ^{13}C NMR spectra were externally referenced to the adamantane CH_2 resonance at 38.48 ppm on the trimethylsilane scale. All spectra were recorded under 10.5 kHz MAS at a sample temperature of 277 K. The sample temperature was estimated from the ^1H chemical shift of bulk water at 4.97 ppm using the equation T (K) = 96.9 × (7.83 − δ_{H_2O}) (ref. 79). Typical radiofrequency field strengths were 50–80 kHz for ^1H and 50–60 kHz for ^{13}C. All spectra were recorded with a recycle delay of 3 s.

Water-edited ^{13}C cross-polarization (CP) experiments were performed to probe the water accessibility and hydration of the channels[80–84]. These experiments involved a ^1H excitation pulse at 71.4 kHz, followed by a 52 rotor period (5.0 ms) Hahn echo with a selective Gaussian 180° pulse of 4.8 ms placed on resonance with water. This water-selective echo period was followed by a ^1H spin diffusion period, whose duration, t_{SD}, was varied from 1 to 225 ms. The ^1H spin diffusion period was followed by 500 μs ^1H–^{13}C CP for ^{13}C detection, during which ^1H two-phase modulation decoupling was applied at 71.4 kHz.

Water-edited ^{13}C spectral intensities were analysed as integrated intensities from 0–75 ppm for all samples and as peak heights for the resolved signals of Ile13 in the LQLL and LLLL samples. These intensities were corrected for water ^1H spin lattice T_1 relaxation by dividing each value by $\exp(-t_{SD}/T_1)$. The T_1-corrected intensities were then normalized to the value at 225 ms. The water ^1H T_1 values were measured in an inversion recovery experiment and ranged from 1.0–1.5 s.

Classical MD simulations

The model for the engineered pentameric membrane protein LLLL was based on its crystal structure (PDB: 6MCT). Molecular models of the Leu-to-Gln (LQ) variant proteins were built and side chains repacked to the global energetic minima using SCRWL4 (ref. 85), with the crystal structure of LLLL as the input template. The simulations were performed before the LCP X-ray structures were solved, and thus they were modelled a priori. The initial TM orientation of LLLL in a lipid bilayer was predicted using the Orientations of Proteins in Membranes database

and Positioning of Proteins in Membranes server[86], and the LQ variant proteins were modelled with an identical geometry to LLLL relative to the membrane by structural superposition.

The MD system was built through an automatic script merging protein and membrane components using Visual Molecular Dynamics (VMD)[87] and the GROMACS engine[88]. First, a pre-assembled 7.5 × 7.5 nm POPC bilayer was modelled using VMD's membrane builder application and aligned with the implicit bilayer used to predict the membrane proteins' insertion geometry. The oriented membrane protein was merged with the lipid bilayer and lipid molecules that clashed with protein atoms (<1.2 Å overlap) were removed. The system was treated as a periodic box, 7.0 nm in the Z direction, and hydrated with TIP3P water (water regions of ~20 Å above and below the bilayer). KCl was added to the system to neutralize protein charge and to yield a final 0.15 M ion concentration. The system used CHARMM36 parameters[89], and the GROMACS 2018 engine was used for minimization and dynamics simulations. The recommended CHARMM36 cut-offs ($r_{Coulomb}$, r_{vdW} = 1.2 nm), switching (1.0 nm) and particle mesh Ewald distances were used. A 2 fs time step was used for Langevin dynamics.

The system was minimized with the steepest descent algorithm (5,000 steps maximum, tolerance of <1,000.0 kJ mol^{-1} nm^{-1}) without atomic position restraints. A 50 ps NVT dynamics simulation was initiated from the minimized model using harmonic position restraints on all non-hydrogen protein atoms (1 kcal mol^{-1} Å^{-2}) using the velocity-rescale thermostat fixed at 298.15 K with a 0.1 ps coupling constant. Next, a 15 ns restraint NPT equilibration simulation was run using a semi-isotropic Berendsen barostat (P = 1 bar, pressure coupling time constant = 5 ps, compressibility = 4.5 × 10^{-5} bar^{-1}) and a Berendsen thermostat (T = 298.15 K and 1.0 ps time constant) while maintaining 1 kcal mol^{-1} Å^{-2} harmonic restraints on protein α-carbon atoms relative to the input structure. Unrestrained production dynamics simulations were then run for 200 ns with a Nosé–Hoover thermostat at 298.15 K with a 1.0 ps time constant and a semi-isotropic Parrinello–Rahman barostat fixed at 1 bar with a pressure coupling constant of 5 ps. Coordinate frames were extracted at 20 ps intervals in these production simulation trajectories. Three independent simulation trajectories were launched for each unique protein sequence with different initial atomic velocities.

The backbone atoms from every frame in the trajectory had an atomic RMSD in the range of 0.6–1.2 Å relative to the initial frame for all variants and each triplicate trajectory. Likewise, for all variants, the average RMSD between any two frames throughout the MD trajectory was <1 Å. From these backbone RMSD metrics, we determined that the protein fold and tertiary structure are very stable for all variants, and deviations from the LLLL model are of the same order as thermal fluctuations. The backbone RMSDs of the medoid frame of each triplicate trajectory relative to the LCP X-ray structures of the LQ variants were all <1.2 Å.

Analysis of classical MD simulations

The channels and their water content were analysed using the Channel Annotation Package (CHAP; https://www.channotation.org) available from the Sansom lab[44]. The trajectories from the production runs were centred on the protein using the GROMACS software and sampled every 200 ps to create a compressed xtc trajectory file for CHAP analysis. Water density plots (Figs. 3 and 6 and Supplementary Figs. 1–6) were generated for each of the three independent 200 ns trajectories performed for each of the pentameric bundle designs. Similarly, the time-averaged water density profiles (Supplementary Figs. 1c–6c) were generated by calculating the average water density at every given position along the z-axis to the pore, s, over the course of the 200 ns simulations. These time-averaged water density profiles were then used to calculate the observed hydrophobic lengths, l_{obs}, reported in Fig. 2c and Extended Data Table 1 using a second-derivative method.

MS-RMD simulations

The X-ray crystal structures for LQLL (PDB: 7UDZ) and LLLL (PDB: 6MCT) were used as the starting structures for simulations. Classical simulations were first performed to equilibrate the protein structures and the simulation systems. Each protein was embedded in a POPC bilayer and solvated with water using CHARMM-GUI (refs. 90–93), and the membrane and water were equilibrated using a standard equilibration protocol. Classical equilibration with no restraints was performed for 500 ns. The CHARMM36 forcefield[94] was used to model all interactions, and simulations were run at 298 K in the NPT ensemble using GROMACS[95].

MS-RMD[45–47,96] was subsequently used to model the water and excess proton in all simulations performed in our analyses. The MS-RMD method captures proton delocalization in water by allowing H–O bonds to break and form. This is achieved by taking a linear combination of possible bonding topology states at every time step. See our previous work for the theory and a detailed description[7,9,11]. The Multi-State Empirical Valence Bond Version 3.2 parameters were used to describe the hydrated excess proton[97]. The excess proton CEC ($\mathbf{r}_{CEC}$) is defined as[98]:

$$\mathbf{r}_{CEC} = \sum_i^N c_i^2 \, \mathbf{r}_{COC}^i \quad (1)$$

where $\mathbf{r}_{COC}^i$ is the coordinate of the CEC of the ith diabatic state, c_i^2 is the squared amplitude of that state, and N is the total number of diabatic states. The CEC defines the position of the delocalized excess proton. The CHARMM36 forcefield was used to model the remaining interactions. Simulations were run at 298 K in the NVT ensemble using LAMMPS[99] with the MS-RMD package.

Umbrella sampling simulations were performed in two dimensions to model the proton transport process. The first CV used was the CEC position along the channel axis, Z'_{CEC}. The channel axis for each system is defined as the average principal component of the protein from a 750 ps MS-RMD simulation after equilibration. The position along this axis is calculated in reference to the centre of mass of the Ile13 α-carbon atoms, such that Z'_{CEC} = 0 Å at that point. The second CV used, ϕ, is a recently developed CV that measures the water connectivity within a channel using graph theory[8]. This CV is a notable improvement over water density, which does not directly bias the formation of a continuous water wire and can result in unphysical water 'clumps' as the bias increases. Instead, the new water connectivity CV measures the length of transient water wire formations on a scale from 0 to 1, where 0 is no water and 1 is water fully connected throughout the channel, agnostic to the number of water molecules. We refer the reader to ref. 8 for further theoretical details.

The open-source, community-developed PLUMED library[100,101] was used to define the umbrella bias potentials. Umbrella sampling windows were set every 0.5 Å in the range of −22 to 22 Å for Z'_{CEC} and every 0.035 in the range 0.140–0.980 and 0.245–0.980 for ϕ for LLLL and LQLL, respectively, for a total of ~2,000 windows for each system. The initial windows were generated by using steered MD to pull water into the channel, and the excess proton was placed at each point along the channel. Subsequent windows were pulled from nearby windows. Windows were equilibrated for 100 ps and then run for 0.5–3.5 ns. A few additional windows were added to ensure sampling overlap.

The two-dimensional PMFs were calculated using WHAM-2D (ref. 102). Error bars were calculated using the block method with four blocks. The minimum free-energy path was calculated using the string method.

The two-dimensional plots and graphs in Figure 5 and Supplementary Figs. 17, 18 and 24 were generated with Matplotlib[103]. The simulation snapshots in Figure 5 were generated using VMD[87].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Coordinates and data files have been deposited with the PDB with accession codes 7UDY (QLLL), 7UDZ (LQLL), 7UDV (LLQL), 7UDW (QQLL) and 7UDX (QLQL). Source data are provided with this paper. All other data and materials are available from the corresponding authors on request.

References

72. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* **4**, 706–731 (2009).
73. Caffrey, M. Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu. Rev. Biophys.* **38**, 29–51 (2009).
74. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
75. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
76. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
77. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
78. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
79. Böckmann, A. et al. Characterization of different water pools in solid-state NMR protein samples. *J. Biomol. NMR* **45**, 319–327 (2009).
80. Luo, W. & Hong, M. Conformational changes of an ion channel detected through water–protein interactions using solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **132**, 2378–2384 (2010).
81. Williams, J. K. & Hong, M. Probing membrane protein structure using water polarization transfer solid-state NMR. *J. Magn. Reson.* **247**, 118–127 (2014).
82. Mandala, V. S. et al. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat. Struct. Mol. Biol.* **27**, 1202–1208 (2020).
83. Gelenter, M. D. et al. Water orientation and dynamics in the closed and open influenza B virus M2 proton channels. *Commun. Biol.* **4**, 338 (2021).
84. Hong, M. et al. Coupling amplification in 2D MAS NMR and its application to torsion angle determination in peptides. *J. Magn. Reson.* **129**, 85–92 (1997).
85. Krivov, G. G., Shapovalov, M. V. & Dunbrack, R. L. Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **77**, 778–795 (2009).
86. 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 (2012).
87. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
88. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
89. Huang, J. & MacKerell, A. D. Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
90. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
91. Jo, S., Kim, T. & Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS ONE* **2**, e880 (2007).
92. Wu, E. L. et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
93. Lee, J. et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).
94. Best, R. B. et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
95. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
96. Nelson, J. G., Peng, Y., Silverstein, D. W. & Swanson, J. M. Multiscale reactive molecular dynamics for absolute pK_a predictions and amino acid deprotonation. *J. Chem. Theory Comput.* **10**, 2729–2737 (2014).
97. Biswas, R., Tse, Y. L., Tokmakoff, A. & Voth, G. A. Role of presolvation and anharmonicity in aqueous phase hydrated proton solvation and transport. *J. Phys. Chem. B* **120**, 1793–1804 (2016).
98. Day, T. J. F., Soudackov, A. V., Čuma, M., Schmitt, U. W. & Voth, G. A. A second generation multistate empirical valence bond model for proton transport in aqueous systems. *J. Chem. Phys.* **117**, 5839–5849 (2002).
99. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
100. Bonomi, M. B. et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **16**, 670–673 (2019).
101. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
102. Grossfield, A. WHAM: the weighted histogram analysis method, v 2.0.9 (University of Rochester, 2002); http://membrane.urmc.rochester.edu/wordpress/?page_id=126
103. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
104. Hodel, A., Kim, S.-H. & Brünger, A. T. *Acta Crystallogr. A* **48**, 851–858 (1992).

Acknowledgements

Diffraction data were collected at the GM/CA@APS and ALS BL 8.3.1. GM/CA@APS is supported by the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006 and P30GM138396), and the Eiger 16M detector by the National Institutes of Health (NIH; S10 OD012289). We also acknowledge the Advanced Photon Source, supported by the US Department of Energy (DE; contract DE-AC02-06CH11357), and beamline 8.3.1 at the Advanced Light Source operated by the University of California at San Francisco with support from the NIH (R01 GM124149 and P30 GM124169), Plexxikon and the Integrated Diffraction Analysis Technologies programme (US Department of Energy Office of Biological and Environmental Research). The Advanced Light Source at Lawrence Berkeley National Laboratory is supported by DE (DE-AC02-05CH11231). H.T.K. was supported by the NIH (K99GM138753). W.F.D. was supported by the NIH (R35 GM122603), NSF (CHE 1709506) and the Air Force Office of Scientific Research (FA9550-19-1-0331). L.C.W. and G.A.V. were supported by the NIH (R01 GM053148). J.M.N. was supported by the NIH (5T32HL007731 and F32GM133085). J.L.T. was supported by the NIH (R35GM122603).

Author contributions

H.T.K. designed the channels, ran flux measurements, crystallized and collected the X-ray diffraction data, and ran and analysed the classical MD simulations. L.C.W. ran the MS-RMD simulations. L.C.W. and G.A.V. analysed the data. M.M. ran classical MD simulations. J.L.T, J.M.N. and L.L. processed and refined the crystal structures. H.T.K. and W.F.D. analysed the experimental data. All authors contributed to the data analysis and writing of the paper.

Competing interests
The authors declare no competing interests.

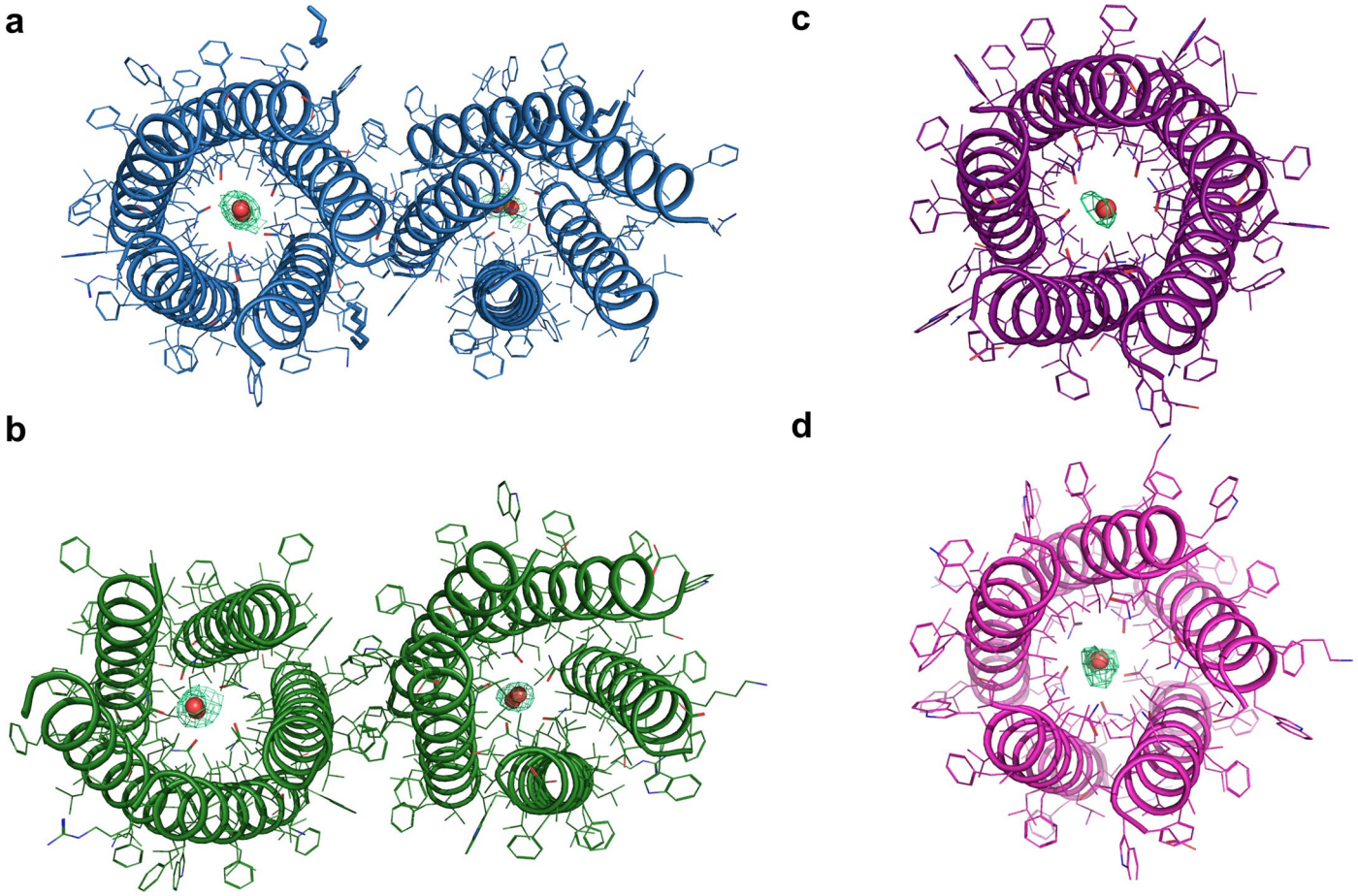
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41557-023-01210-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41557-023-01210-4.

Correspondence and requests for materials should be addressed to Huong T. Kratochvil, Gregory A. Voth or William F. DeGrado.

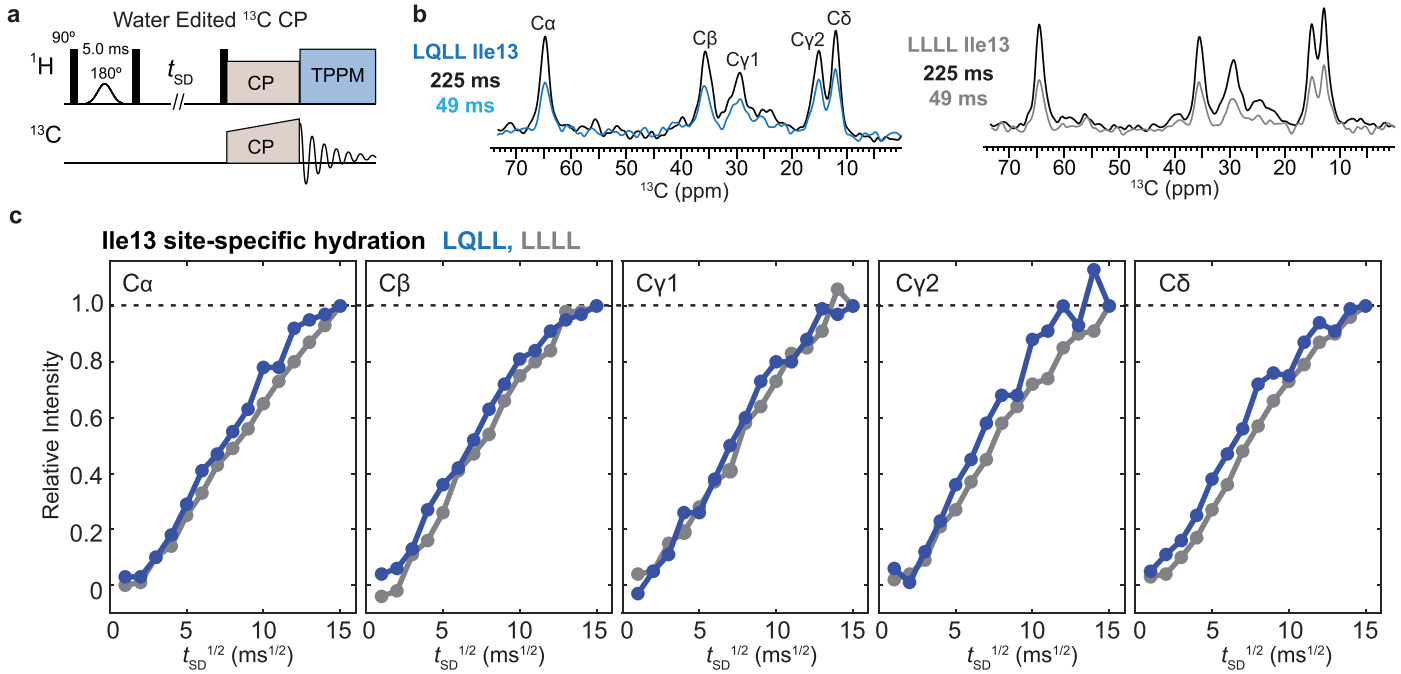
Peer review information *Nature Chemistry* thanks Mihail Barboiu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41557-023-01210-4



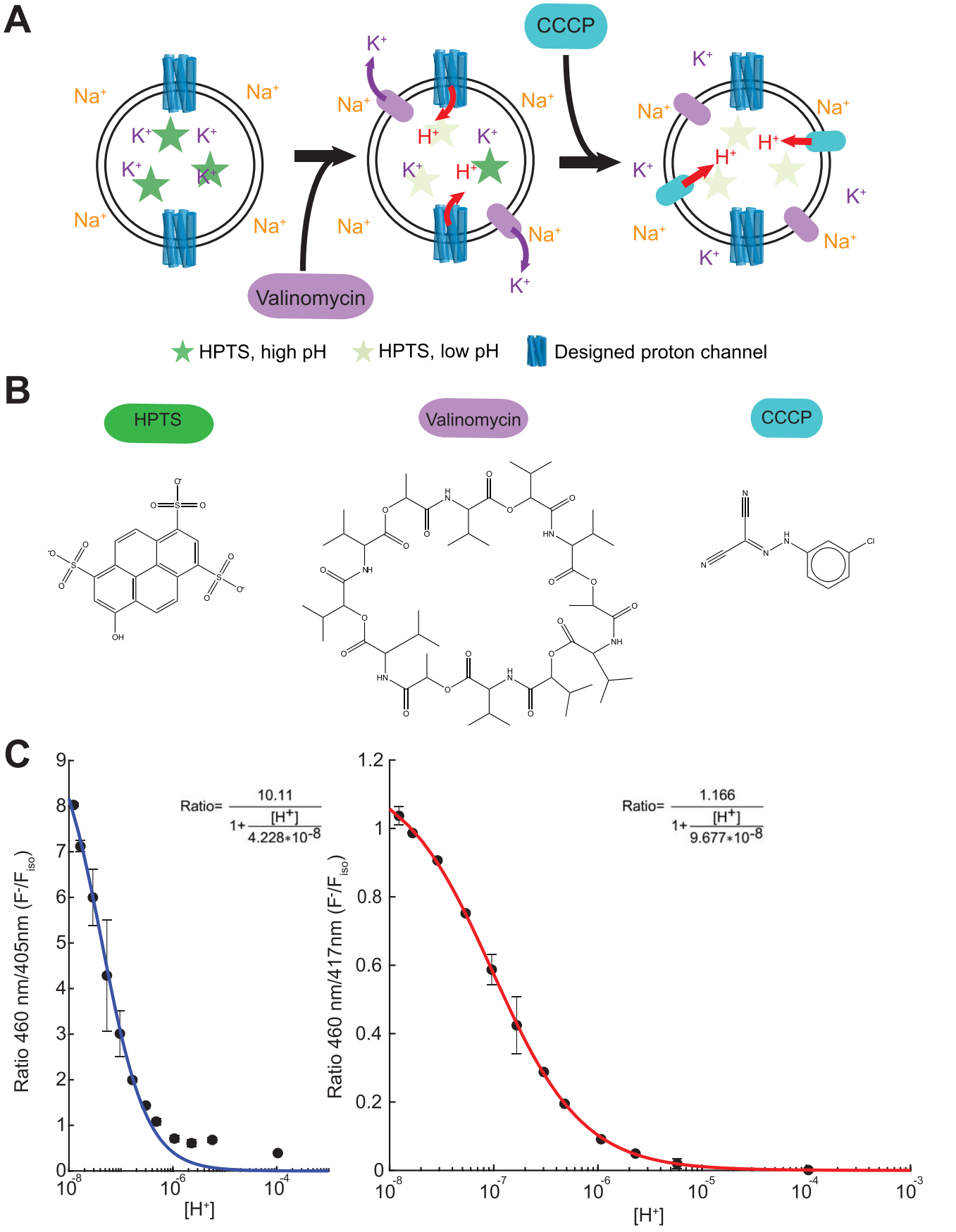
Extended Data Fig. 1 | Composite omit maps (2mFo-DFc) of designed proton channels. Composite omit maps of the asymmetric unit for **a**, LQLL, **b**, LLQL, and one pentamer from the asymmetric unit for **c**, QQLL, and **d**, QLQL (shown only for the waters for clarity). All contours at σ = 1.0. Omit maps with simulated cartesian annealing were generated using Phenix, using methodology described in Hodel, et al.[104].

https://doi.org/10.1038/s41557-023-01210-4



Extended Data Fig. 2 | Pulse diagram and water-edited 13 C spectra, water buildup curves of membrane-bound LQLL and LLLL peptides. a, Pulse diagram of water-edited 13 C CP experiment. **b**, Representative water-edited 13 C spectra of I13 in LQLL and LLLL, measured with 225 ms and 49 ms ^{1}H mixing. The relative intensities of the 49 ms spectrum to the 225 ms spectrum are higher for LQLL Ile13 than LLLL Ile13, especially for the sidechain Cγ2 and Cδ carbons. **c**, Site-resolved water buildup curves for Ile13 in LQLL and LLLL. For all ^{13}C sites, LQLL shows a faster water buildup than LLLL, consistent with water molecules in the pore lumen due to the Gln10 PLS.

https://doi.org/10.1038/s41557-023-01210-4

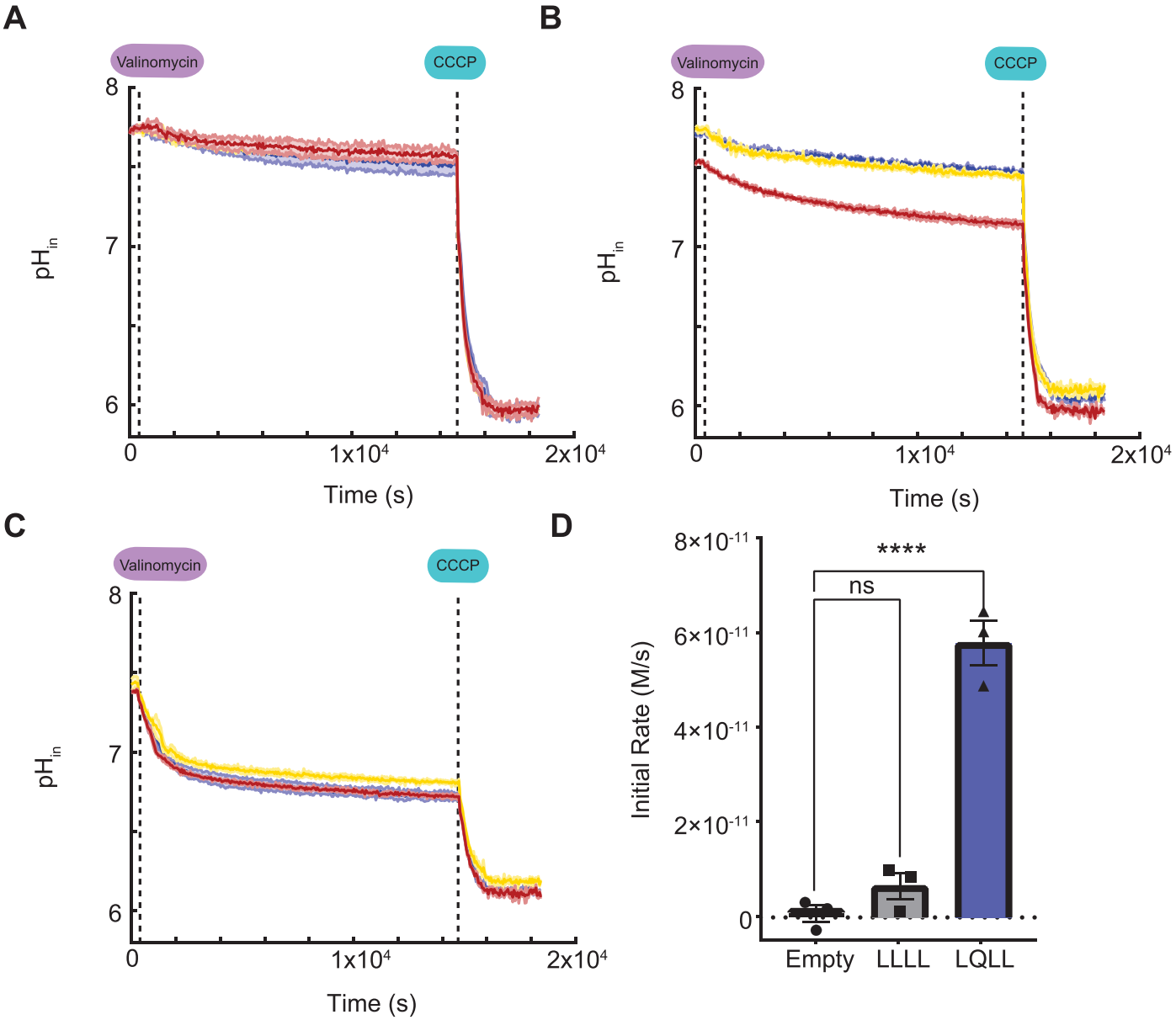


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41557-023-01210-4

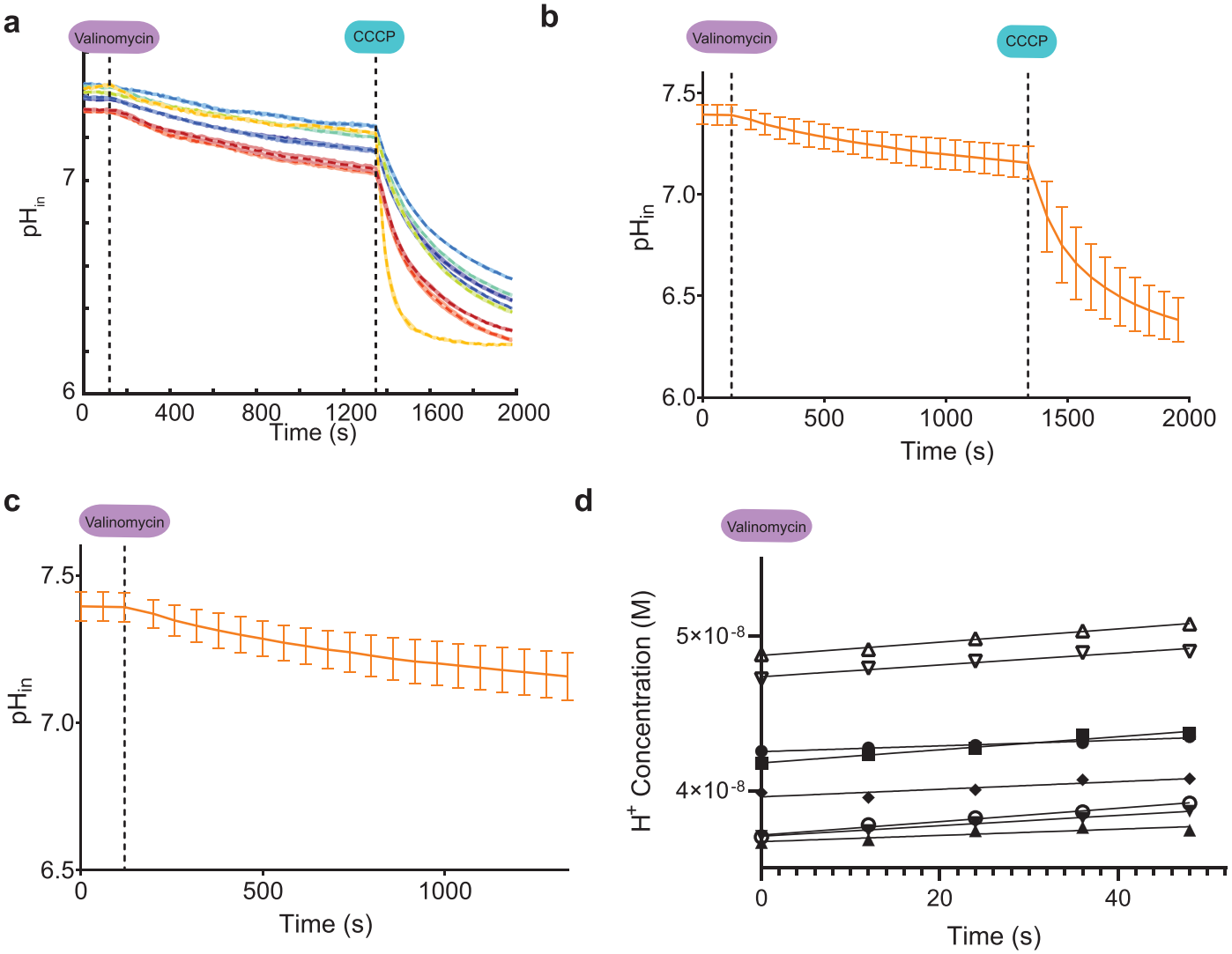
Extended Data Fig. 3 | Overview of proton flux measurements. a, Full schematic for proton flux measurement including CCCP step, which is included to check vesicle leakiness and confirm proton selectivity. **b**, Chemical structures of key components of vesicle assay. **c**, Calibration curves for HPTS at ~5 μM in 12 solutions of 50 mM K_2SO_4, 30 mM K_2HPO_4 at different pH values for two plate readers used in data collection process. Unless stated, all data collected with instrument that generated the blue calibration curve. Fits used for downstream data processing shown for each of the two instruments with adjusted R-squared values of 0.9866 and 0.9998 for the left and right curves, respectively. Data for n = 3 independent samples shown as mean values +/− SD.

https://doi.org/10.1038/s41557-023-01210-4



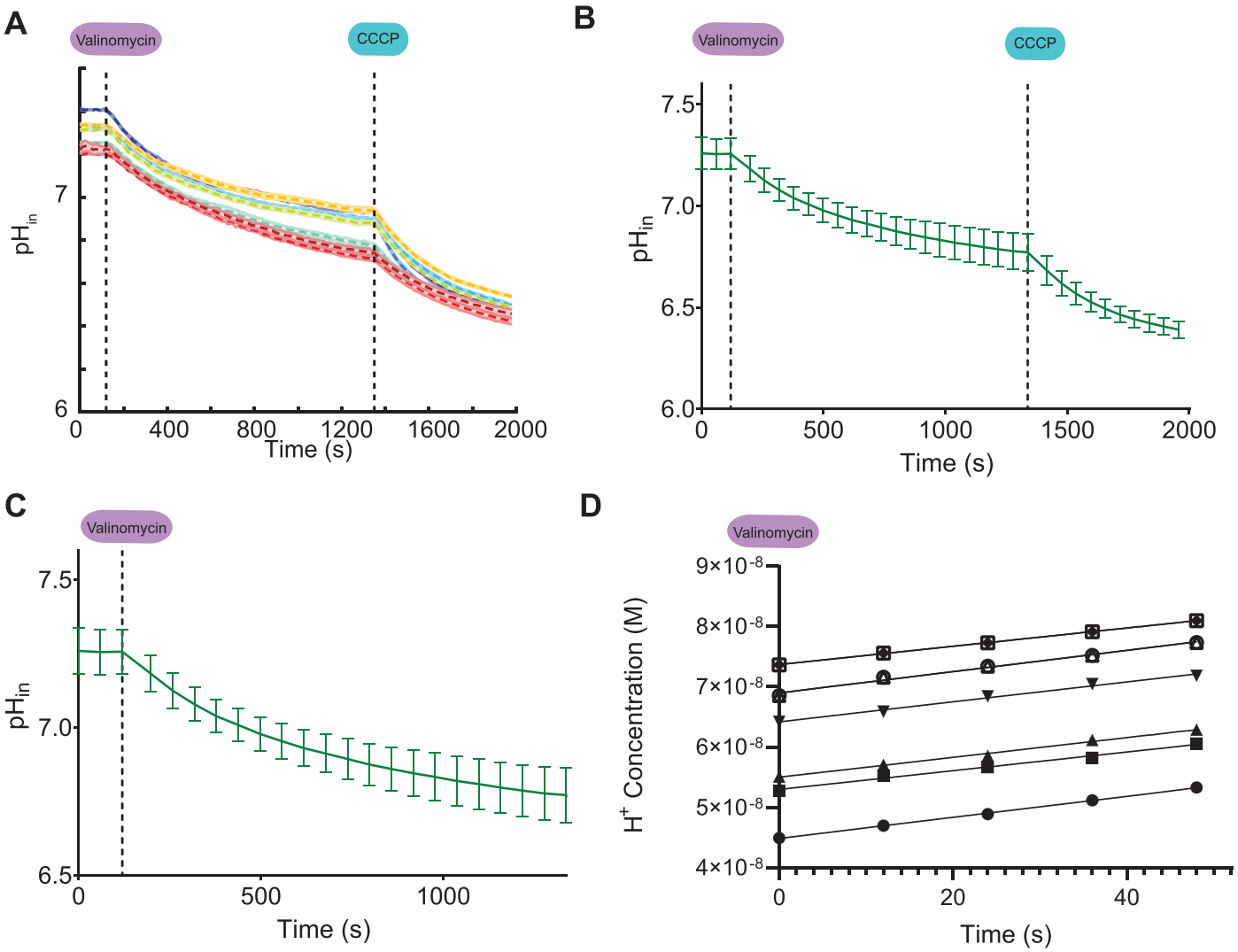
Extended Data Fig. 4 | All proton flux assay data for long kinetics runs. Long kinetics runs of about 5 hours total for **a**, empty, **b**, LLLL, and **c**, LQLL vesicles. Dotted lines denote times in experiment when valinomycin and CCCP were added to the samples. Three samples of the different conditions were measured in triplicate. These long-time measurements reveal that the vesicles are not leaky to Na^+, K^+, or H^+ and maintain their cargo and assembly over the entire course of the measurement. **d**, From the linear regression fits of the first 220 s following addition of valinomycin, all slopes (which give the initial rates (in M/s)) were used to calculate the mean and standard error. The one-way ANOVA analysis (with Dunn's test) reveals that LLLL rates are not significantly different ($p > 0.05$, adjusted $p = 0.4021$) when compared to the control empty vesicles. LQLL rates, however, are statistically significant ($p < 0.0001$, $p = 3.23E\text{-}5$) when compared to the control empty vesicles using one-way ANOVA analysis with Dunn's test. All data from $n = 3$ independent samples are shown as mean values +/− SD.

https://doi.org/10.1038/s41557-023-01210-4



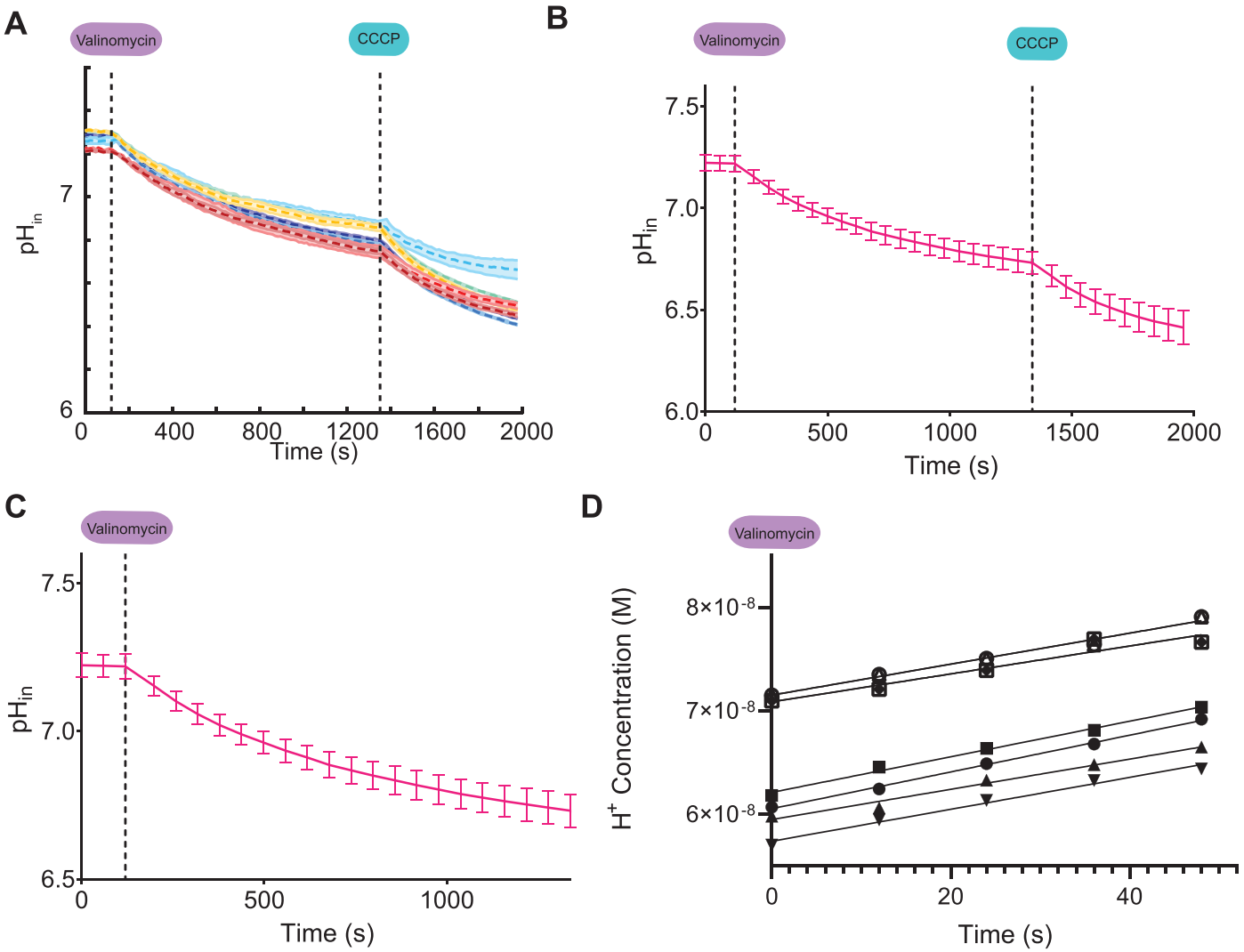
Extended Data Fig. 5 | All proton flux assay data for QLLL vesicle samples. Nine samples (each run in triplicate with shaded error bars shown) containing 1:500 peptide:lipid ratio; samples were run independently in the assay. **a**, pH_{in} as a function of time throughout the measurement for each independent sample. **b**, Mean and standard deviation for data collection. **c**, Data prior to CCCP addition shows little change in pH_{in} after addition of valinomycin. **d**, Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.

https://doi.org/10.1038/s41557-023-01210-4



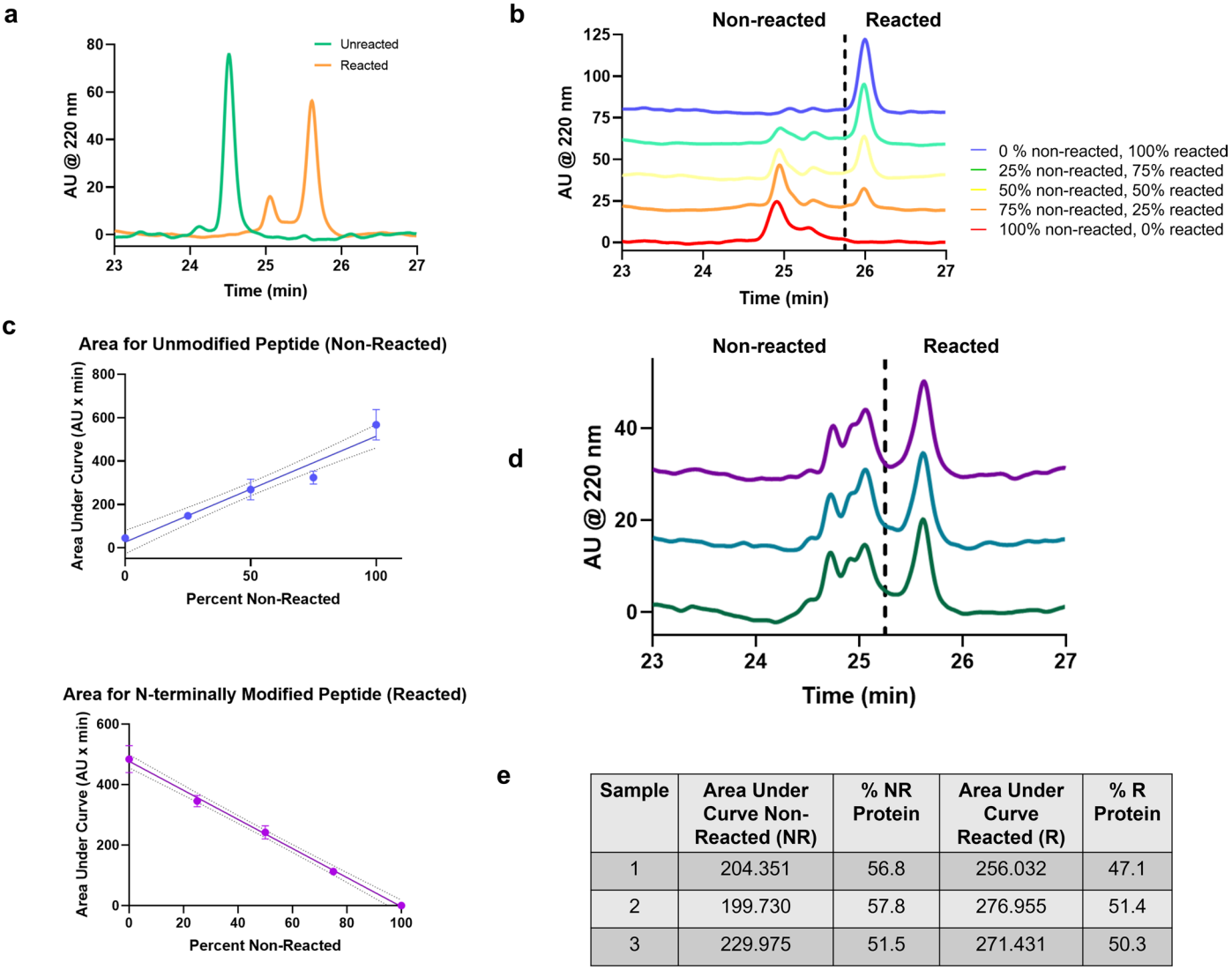
Extended Data Fig. 6 | All proton flux assay data for LLQL vesicle samples. Eight samples (each run in triplicate with shaded error bars shown) containing 1:500 peptide:lipid ratio; samples were run independently in the assay. **a**, pH_{in} as a function of time throughout the measurement for each independent sample. **b**, Mean and standard deviation for data collection. **c**, Data prior to CCCP addition shows notable change in pH_{in} after addition of valinomycin. **d**, Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.

https://doi.org/10.1038/s41557-023-01210-4



Extended Data Fig. 7 | All proton flux assay data for QLQL vesicle samples. Seven samples (each run in triplicate with shaded error bars shown) containing 1:500 peptide:lipid ratio; samples were run independently in the assay. **a**, pH_{in} as a function of time throughout the measurement for each independent sample. **b**, Mean and standard deviation for data collection. **c**, Data prior to CCCP addition shows notable change in pH_{in} after addition of valinomycin. **d**, Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.

https://doi.org/10.1038/s41557-023-01210-4



e

Sample	Area Under Curve Non-Reacted (NR)	% NR Protein	Area Under Curve Reacted (R)	% R Protein
1	204.351	56.8	256.032	47.1
2	199.730	57.8	276.955	51.4
3	229.975	51.5	271.431	50.3

Extended Data Fig. 8 | Determining orientation of pentamers in vesicles. a, HPLC trace of unreacted and reacted peptides following reaction with the highly polar, amine-reactive methyltetrazine 3-sulfo-N-hydroxysuccinimide ester (methyltetrazine sulfo-NHS, see Methods). The only amine-reactive groups are the N-terminus or the N-terminal lysine sidechain. Thus, only peptides in which N-terminus is exposed on the outside of the vesicle should react to the dye. **b**, HPLC traces of mixtures of different ratios of non-reacted and reacted peptides. **c**, Calibration curves for area under the curve for nonreacted and reacted peaks in the HPLC traces corresponding to the different mixtures in **b**. Data shown are for n = 2 independent experiments and shown as mean values +/− SD. **d**, Traces of three independent samples of LQLL pentamers from vesicles after reaction with methyltetrazine sulfo-NHS. **e**, Using the calibration curves in **c**, the area under the curve was determined for each sample. The data indicate that half the amines react, as expected from a random orientation of pentamers in the lipid vesicle.

https://doi.org/10.1038/s41557-023-01210-4

Extended Data Table 1 | Length (l_{obs}) of longer and shorter ($l_{obs,sh}$) hydrophobic stretches in designed channels

Design	l_{obs}	$l_{obs,sh}$
LLLL	33.9 ± 0.3 Å	-
QLLL	32.0 ± 0.6 Å	-
LQLL	20.7 ± 0.1 Å	11.6 ± 0.4 Å
LLQL	22.4 ± 0.1 Å	9.3 ± 0.5 Å
QQLL	20.9 ± 0.1 Å	8.2 ± 0.2 Å
QLQL	20.3 ± 0.2 Å	9.2 ± 0.2 Å

The longest hydrophobic length for all peptides and shorter hydrophobic length ($l_{obs,sh}$) for each peptide LQLL, LLQL, QQLL, and QLQL were calculated from the classical MD simulations (see Methods).

https://doi.org/10.1038/s41557-023-01210-4

Extended Data Table 2 | Data collection and refinement statistics for all structures

Property	QLLL PDB: 7UDY APS 23-ID-D	LQLL PDB: 7UDZ APS 23-ID-B	LLQL PDB: 7UDV APS 23-ID-D	QQLL PDB: 7UDW ALS 8.3.1	QLQL PDB: 7UDX ALS 8.3.1
Data Collection					
Space Group	C 2 2 2_1	C2	P 2_1 2_1 2_1	P2_1	C 2 2 2_1
Cell dimensions					
a, *b*, *c* (Å)	56.39 84.84 149.56	86.77 46.66 66.32	52.53 55.36 82.71	48.246 100.175 50.498	55.72 83.68 147.03
α, *β*, *γ* (°)	90.0 90.0 90.0	90.00 108.96 90.00	90.0 90.0 90.0	90.0 110.506 90.0	90.0 90.0 90.0
Resolution (Å)	74.78 – 2.40 (2.53-2.40)	41.03 – 2.47 (2.61-2.47)	46.00 – 2.40 (2.46-2.40)	42.77 – 3.00 (3.18 – 3.00)	73.24 – 2.99 (3.04-2.99)
R_{merge}	0.155 (1.249)	0.114 (0.708)	0.111 (0.826)	0.222 (1.207)	0.145 (0.884)
< I /σI >	8.1 (2.2)	5.5 (1.8)	7.6 (2.0)	6.6 (1.8)	13.4 (3.5)
Completeness (%)	99.5 (97.7)	98.1 (93.4)	91.2 (97.3)*	98.8 (98.5)	99.9 (100.0)
CC1/2	0.998 (0.831)	0.991(0.800)	0.999 (0.591)	0.998 (0.815)	0.999 (0.921)
Redundancy	8.5 (8.5)	3.2 (3.0)	3.3 (3.6)	6.7 (6.4)	13.1 (13.9)
Refinement					
Resolution (Å)	2.40	2.47	2.40	3.00	2.99
No. reflections	14307	8896	8048	8355	7234
R_{work}/R_{free}	0.230/0.251	0.248/0.269	0.246/0.268	0.251/0.286	0.231/0.273
No. atoms					
Protein	3098	2038	2035	4093	3093
Ligand	63	21	0	0	14
Water	0	4	4	6	0 (6)*
***B*-Factors**					
Protein	67.13	55.05	52.91	62.97	55.87
Ligand/ion	82.94	61.56	N/A	N/A	48.21
Water	N/A	56.71	43.92	41.17	N/A (30.74)*
R.M.S. deviations					
Bond lengths (Å)	0.001	0.002	0.003	0.004	0.003
Bond angles (°)	0.324	0.414	0.498	0.583	0.490
Ramachandran statistics					
Outliers (%)	0	0	0	0	0
Allowed (%)	0	0	0	0.43	0
Favored (%)	100	100	100	99.57	100

All structures were determined from single protein crystals. Data in parentheses denote statistics for outermost shell. All listed data represent isotropic statistics unless denoted by *, which is anisotropic.

nature portfolio

Corresponding author(s): William F. DeGrado

Last updated by author(s): Apr 15, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All fluorescence experiments were performed on Biotek instrumentation with software Gen5 v 1.11.5 and v 3.10. All data collection information for crystal structures reported in the Methods section.
Data analysis	All fluorescence data analyzed using Matlab 2018b and GraphPad Prism 9. All data analysis for crystal structures reported in Methods section. All software for analysis of MD simulations reported in Methods section. Protein structures were modeled and visualized using Pymol 2.3.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Coordinates and data files have the PDB with accession codes 7UDY (QLLL), 7UDZ (LQLL), 7UDV (LLQL), 7UDW (QQLL), and 7UDX (QLQL). Data and materials are available from the authors on request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All data for proton flux experiments were measured in triplicate on n > 6 samples. All classical molecular dynamics (MD) simulations were run in triplicate from random initial configurations.
Data exclusions	Sample outliers in the fluorescence measurements were excluded as per statistical analysis. Minimization and equilibration MD simulation runs were excluded from analysis.
Replication	All functional data collected on samples prepared on different days were collected in triplicate and successfully reproduced (all samples used in analysis shown in Extended Data and Supplementary Information). All deposited crystal structures were collected from single protein crystals. For classical MD simulations, runs were initiated with different starting equilibrated structures.
Randomization	This is not a study where a hypothesis is tested through statistical analyses of the results/observations of individuals, therefore randomization is not needed. For MD simulations, different starting configurations were randomly generated.
Blinding	This is not a study where a hypothesis is tested through statistical analyses of the results/observations of individuals, therefore blinding is not needed.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

nature chemistry

Article

https://doi.org/10.1038/s41557-023-01210-4

Transient water wires mediate selective proton transport in designed channel proteins

In the format provided by the authors and unedited

Supplementary Materials for

Transient Water Wires Mediate Selective Proton Transport in Designed Channel Proteins

Huong T. Kratochvil, Laura C. Watkins, Marco Mravic, Noah H. Somberg, Jessica L. Thomaston, John M. Nicoludis, Lijun Liu, Mei Hong, Gregory A. Voth, William F. DeGrado

Correspondence to: huong.kratochvil@unc.edu, gavoth@uchicago.edu, william.degrado@ucsf.edu,

Materials and Methods

Synthesis of Proton Channel Peptides

Proton channel peptides, shown in Fig. 2 and Supplementary Table 1, were synthesized using Fmoc-based solid-phase peptide synthesis on either a Biotage Initiator Alstra microwave synthesizer or a Syro II parallel peptide synthesizer. All peptides were synthesized with a free amine N-terminus and a C-terminal carboxamide using Tentagel S-RAM resin (Chem-Impex Int'l Inc) with 0.22-0.24 mmol/g. Following synthesis, the peptides were cleaved from the resin with a trifluoroacetic acid: triisopropylsilane: water (TFA:TIPS:H_2O, 95:2.5:2.5) solution, precipitated out with cold diethyl ether, redissolved in 50% 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) and water, and purified using reverse phase HPLC on a C4 prep column (Vydac) with a gradient of solvents A (water, 0.1% TFA) and B (isopropanol:acetonitrile:water:TFA, 60:30:9.9:0.1) at a flow rate of 10 mL/min. Peptides were then lyophilized until dry. Following successful purification, peptides were confirmed to be >90% pure with analytical HPLC over a C4 column (Phenomenex) and confirmed to have the correct mass with MALDI mass spectrometry (Shimadzu AXIMA Performance) using α-cyano-4-hydroxycinnamic acid (Sigma) as the matrix (Supplementary Table 1). Peptides were then dissolved in ethanol to the appropriate stock concentrations and subsequently used for all experiments as described.

Expression of Full-Length Influenza A M2 Polypeptide

The Influenza A M2 polypeptide from previous work [49] was expressed and purified from *E. coli* cells. Briefly, the gene containing the A/Udorn/72 Cys-free W15F variant (W15F, C17S, C19S, C50S) and a C-terminal 6x His tag was cloned into a pEXP-5-NT plasmid and transformed into chemically competent *E. coli* BL21(DE3) cells (Invitrogen) using heat shock methods. The cells were grown in TB media (Invitrogen) with ampicillin at 37°C. When an OD_{600} of 0.6-0.8 was achieved, the expression culture was induced with 1 mM isopropyl ß-D-1-thiogalactopyranoside (IPTG). The cells were then harvested no more than 2.5 hours after induction and centrifuged down at 4°C for 10 minutes. Pellets were then frozen prior to purification. For purification, the cells were resuspended in lysis buffer (50 mM Tris pH 8, 300 mM NaCl, 5% glycerol, 25 mM imidazole, 0.1% decyl maltose neopentyl glycol (DMNG), 8 M urea) and sonicated with a microtip sonicator for 5 minutes at 20% amplitude (2 s on/off). Lysed cells were then spun down for 15 minutes at 18000 rpm (~41000 x g) and purified using Ni-NTA beads using batch purification methods. Wash buffer was the same as the lysis buffer. The protein was eluted in elution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5% glycerol, 250 mM imidazole, 0.1% DMNG, 8 M urea) and mixed with HFIP

prior to loading onto the C4 column for secondary purification by HPLC. Polypeptide was lyophilized until dry and reconstituted in ethanol to appropriate stock concentrations. SDS-PAGE and MALDI-MS were used to confirm the expression of the polypeptide.

SDS-PAGE Gels of Proton Channels

25-50 µg of peptide were lyophilized from the peptide stocks and dissolved in 25 mM Tris pH 7.5 (Fisher Scientific) with 50 mM octyl-β-glucopyranoside (OG, Carbosynth Limited) for a final concentration of 2.5 µg/µL. An equal volume of 2x LDS (Invitrogen) was added to each sample and the sample was subsequently boiled for 15 minutes at 95ºC. 5 µg of peptide was loaded in each well of a 12% Bis-Tris NuPAGE gel (Thermo Fisher Scientific) with Precision Plus Protein Dual Xtra prestained protein standards (Bio-Rad Laboratories); the gel was subsequently run for 30 minutes at 200 V.

Proton Channel Liposomal Flux Assays and Analysis

The proton channel liposomal flux assays were adapted from references [48,49], and the exact methodologies for preparation and running of these assays are described below.

Preparation of Proteoliposomes

All proteoliposome samples were made as a stock solution of 1:500 peptide:lipids with final concentrations of 10 µM and 5 mM, respectively. To prepare the proteoliposome stock, 10 uL of 100 mM of a ratiometric pH-sensitive dye, 8-hydroxypyrene-1,3,6-trisulfonic acid trisodium salt (HPTS) was added to total of 2 mL of K^+ buffer (50 mM K_2SO_4, 30 mM K_2HPO_4, pH 7.5). The solution was then added to a dried film of 3:1 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine:1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (POPC:POPG, Avanti Lipids) to make a 10 mM lipid solution, which was subsequently vortexed for 1 minute before sonicating with microtip for 5 min at 20% amplitude for 2 s on/off. The sonicated solution was then divided into 200 uL aliquots where 1M OG was added to a final concentration of 26 mM OG. After one hour on the rotisserie at RT, peptide in 26 mM OG in K^+ buffer (or, in the case of empty, 15 µL of 26 mM OG in K^+ buffer) was added to the solution of detergent-solubilized lipids for a final concentration of 20 µM peptide. The peptide-detergent-lipid (PDL) solution was equilibrated for 1 hour at RT on the rotisserie. Following the incubation of the PDL solution, 50 µL of XAD biobeads solution (Sigma-Aldrich) in K^+ buffer was added to the PDL mixture every 5-10 minutes five times for a final volume of around 500 µL and final peptide and lipid concentrations as described above. After the final addition of

biobeads, the solution was left to rotate in the 4°C rotisserie overnight. In the following morning, the sample was spun down in an ultracentrifuge at 96k rpm (~328000 x g) for 10 minutes to pellet the liposomes. The liposome pellet was resuspended in the same amount of dye-free K^+ buffer to afford the stock solution used for the liposomal assays. Dynamic light scattering (Malvern Panalytical) of the vesicle samples enabled size determination prior to proton flux assays to ensure liposome formation.

Preparation of Working Solutions for Proton Flux Assays

A working 3 µM valinomycin solution was made by dissolving 1 mg/mL (~900 mM) valinomycin stock in DMSO (purchased from Sigma-Aldrich) into Na^+ buffer (50 mM Na_2SO_4, 30 mM Na_2HPO_4, pH 7.5). Similarly, a 3 µM CCCP working solution was generated by dissolving 600 µM CCCP stock in DMSO into Na^+ buffer.

For the assay, 17.3 µL of the stock proteoliposome solution was diluted into 632.7 µL of Na^+ buffer mixed with 10 mM p-xylene-bis-pyridinium bromide (DPX, Invitrogen) for a total of 650 uL of sample at the working concentrations. The presence of membrane-impermeable DPX quenches any extraliposomal HPTS fluorescence. Therefore, the fluorescence measured in this experiment comes only from intraliposomal HPTS. The sample was allowed to equilibrate for 20 minutes before 190 µL was aliquoted into three wells on a black, u-shaped bottom 96-well plate (Greiner) for fluorescence measurements (collected every 55 s).

Instrumentation

All data reported in Fig. 6, Extended Figs. 5-7, Supplementary Figs. S19-S23 were collected using the Biotek Synergy 2 equipped with a 405 nm/20 nm bandwidth and 460 nm/40 nm bandwidth excitation filters and a 528 nm/20 nm bandwidth emission filter (Biotek). The valinomycin and CCCP solutions (10 uL each) were added to the wells using an injector system at 120 s and 1350 s into the assay, respectively, with 3 s of vigorous shaking following addition. The fluorescence was measured at the two excitation wavelengths every 12 s.

For the long-time kinetics run (Extended Data Fig. 4), the data was collected on a Biotek Synergy H1 using monochromators. The excitation wavelengths were set at 417 nm and 460 nm, and the emission wavelength was set to 515 nm.

Calibration and Analysis of Flux Data

For all data reported in Figs. 6, Extended Data Figs. 5-7, and Supplemental Figs. S19-S23, the fluorescence signals were calibrated using the ratio of the fluorescence signals of the

deprotonated dye, F^- ($\lambda_{excitation}$ = 460 nm, $\lambda_{emission}$ = 528 nm) and the isosbestic point, F_{iso} ($\lambda_{excitation}$ = 405 nm, $\lambda_{emission}$ = 528 nm) as a function of pH from pH 8 to pH 4 (Extended Data Fig. 3). Similarly, for the long-time kinetics runs, a calibration curve of the ratio of the fluorescence signals of F^- ($\lambda_{excitation}$ = 460 nm, $\lambda_{emission}$ = 515 nm) and F_{iso} ($\lambda_{excitation}$ = 417 nm, $\lambda_{emission}$ = 515 nm) was derived as a function of pH (Extended Data Fig. 1). All fluorescence measurements for each reported sample was taken in triplicate and were then converted into values of pH_{in} using the defined calibration curves.

For data reported in Fig. 6, Extended Data Fig. 5-7, and Supplemental Figs. S19-S23, initial rates of proton conduction were determined for each sample using data collected within the first sixty seconds following the addition of valinomycin. Similarly, for the long-time kinetics runs (Extended Data Fig. 4), initial rates were determined by fitting the first 220 seconds of data following the addition of valinomycin. All data were fit using linear regression and the initial rates were extracted from the fitted slope.

Orientation Determination of Peptides in Vesicles

Liposomes containing 108 uM of peptide (P:L ratio of 1:50) were used for the following experiment. Calibration curves were prepared as follows: 75 uL of liposome solution was added to 37.5 uL of 1 M OG and 37.5 uL of water (for non-reacted or NR samples) or 37.5 uL of 20 mM methyltetrazine-sulfo NHS ester (Click Chemistry Tools) in water (for reacted or R samples) for 30 minutes. Samples were then quenched with 16.67 uL of 1 M Tris pH 8 for 5 minutes. Non-reacted and reacted samples were mixed in differing ratios and diluted in 50% HFIP in water before running on reverse phase HPLC on an analytical C4 column with a gradient of solvents A (water, 0.1% TFA) and B (isopropanol:acetonitrile:water:TFA, 60:30:9.9:0.1) at a flow rate of 1 mL/min. The area under the curves for the non-reacted and reacted peaks were calculated for each sample to generate the calibration curve (Extended Data Fig. 8). Traces in figures are background-subtracted and offset for ease of viewing.

To determine orientation of the samples, 50 uL of the liposome sample was mixed with 50 uL of 10 mM methyltetrazine sulfo-NHS ester in water for 30 minutes and subsequently quenched with 5.56 uL of 1 M Tris pH 8 for 5 minutes. The samples were then diluted with 50:50 1 M OG and HFIP before running on HPLC. The area under the curve was calculated for the non-reacted and reacted peaks and used with the calibration curves to estimate the orientation of the peptide (Extended Data Fig. 8d,e). All samples were measured in duplicates.

Lipidic Cubic Phase (LCP) Crystallography of Proton Channels

The methods described for LCP crystallization are originally described in detail in Caffrey and Cherezov and have been slightly modified as follows [72,73]. Peptide from the ethanol stock was mixed with 60 mg of monoolein (Sigma-Aldrich) until clear. The solution containing peptide and monoolein was then dried under a gentle stream of N_2 and lyophilized overnight. To prepare the LCP, the monoolein-peptide mixture was heated to 42ºC until it became liquid, and subsequently extruded ~200 times with 2/3 times the volume of 50 mM OG in coupled gastight Hamilton syringes at room temperature. Successful LCP formation was confirmed as the solution became clear and was not birefringent in the cross-polarizer. The final concentration of peptide for each sample is listed in Supplementary Table 2.

For crystallization, 50 nL of the LCP mixture was dispensed onto 96-well Laminex plastic sandwich plates (Molecular Dimensions) with 0.5 to 1 uL of precipitant solution using the TTP Labtech LCP Mosquito robot at room temperature. Plates were sealed using plastic coverslips (Molecular Dimensions) and monitored using a Formulatrix RockImager at 20ºC. Crystals of each construct were harvested from conditions noted in Supplementary Table 2 with and cryoprotected with 30% v/v PEG 400, if necessary, and flash-frozen with liquid nitrogen.

X-Ray Diffraction Data Collection and Analysis

Crystals were mounted under a cryostream at 100K and data was collected at both the 8.3.1 beamline at the Advanced Light Source at Lawrence Berkeley National Lab at a wavelength of 1.115832 Å or at 23-ID-B/D at Argonne National Laboratory at a wavelength of 1.03318 Å. The data was processed with the XDS package[74] and reduced with AIMLESS within the CCP4 suite [75]. The structures were determined by molecular replacement with Phaser [76] using a previously designed de novo protein (pdb code: 6mct) as a model. The models were rebuilt in Coot [77] and the structures were then refined with PHENIX [78]. The data processing and structural refinement statistics are described in Extended Data Table 2.

Solid-state NMR (ssNMR) Experiments

LQLL and LLLL peptides with the appropriate site-specific ^{13}C, ^{15}N labels at the Ile6, Ile13 or Gln10 positions were made using SPPS methods as described previously. The purified peptides were reconstituted into d54-DMPC lipids at a protein/lipid molar ratio of 1:12 with a final peptide weight of ~5 mg. The samples are concentrated to a hydration level of ~40% (w/w) and pelleted into 3.2 mm rotors for NMR experiments.

Magic-angle-spinning (MAS) solid-state nuclear magnetic resonance (NMR) spectra were measured on a Bruker Avance III HD 600 MHz (14.1 T) spectrometer with a Bruker 3.2 mm HXY MAS probe operating in double-resonance $^{1}H/^{13}C$ mode. All ^{13}C spectra were externally referenced to to the adamantane CH_2 resonance at 38.48 ppm on the trimethylsilane scale. All spectra were recorded under 10.5 kHz MAS at a sample temperature of 277 K. The sample temperature was estimated based on the ^{1}H chemical shift of bulk water at 4.97 ppm, according to the equation T (K) = 96.9 × (7.83 ppm – δ_{H2O}) [79]. Typical radiofrequency field strengths were 50-80 kHz for ^{1}H and 50-60 kHz for ^{13}C. All spectra used a recycle delay of 3 s.

Water-edited ^{13}C cross-polarization (CP) experiments were used to probe the water accessibility and hydration of the channel[80-84]. These consisted of a ^{1}H excitation pulse at 71.4 kHz, followed by a 52 rotor period (5.0 ms) Hahn echo with a selective Gaussian 180˚ pulse of 4.8 ms placed on resonance with water. This water-selective echo period is followed by a ^{1}H spin diffusion period, whose duration (t_{SD}) was varied from 1 to 225 ms. The ^{1}H spin diffusion period was followed by a 500 µs ^{1}H-^{13}C cross-polarization (CP) for ^{13}C detection, during which ^{1}H two-phase modulation (TPPM) decoupling was applied at 71.4 kHz.

Water-edited ^{13}C spectral intensities were analyzed as integrated intensities from 0-75 ppm for all samples and as peak heights for the resolved Ile13 signals in LQLL and LLLL samples. These intensities are corrected for water ^{1}H T_1 relaxation by dividing each value by $\exp(-t_{SD}/T_1)$. The T_1 corrected intensities are then normalized to the value at 225 ms. The water ^{1}H T_1 values were measured using the inversion-recovery experiment and ranged from 1.0–1.5 s.

Classical Molecular Dynamics Simulations

The model for the engineered pentameric membrane protein LLLL was taken from its crystal structure (pdb: 6mct). Molecular models of the Leu-to-Gln (LQ) variant proteins were built and sidechains repacked to the global energetic minima using Scrwl4 [85], with the LLLL crystal structure as the input template. The simulations were performed prior to LCP X-ray structures being solved, and thus modeled *a priori*. The initial transmembrane orientation of LLLL in a lipid bilayer was predicted by the OPM PPM 2.0 server [86], and the LQ variant proteins were modeled with identical geometry as LLLL relative to the membrane by structural superposition.

The MD system was built through an automatic script merging protein and membrane components using VMD [87] and the GROMACS engine [88]. First, a pre-assembled 7.5 x 7.5 nm 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer was modeled using VMD's membrane builder application and aligned with the implicit bilayer used to predict the membrane

proteins' insertion geometry. The oriented membrane protein was merged with the lipid bilayer and lipid molecules which clashed with protein atoms (<1.2 Å overlap) were removed. The system was treated as a periodic box, 7.0 nm in the Z direction, and hydrated with TIP3P water (ca. 20 Å of water regions above and below the bilayer). KCl was added to the system to neutralize protein charge and to yield a final 0.15 M ion concentration. The system used CHARMM36 parameters [89] and the GROMACS 2018 engine was used for minimization and dynamics simulations. The recommended CHARMM36 cut-offs (rcoulomb, rvdw = 1.2 nm), switching (1.0 nm), and Particle-Mesh Ewald distances were used. A 2 fs time step was used for Langevin dynamics.

The system was minimized with the steepest decent algorithm (5000 steps max, tolerance of < 1000.0 kJ/mol/nm) without atomic position restraints. Harmonic positional restraints of 1 kcal mol^{-1} $Å^{-2}$ on all non-hydrogen protein atoms. A 50 ps NVT dynamics simulation was initiated from the minimized model using harmonic position restraints on all non-hydrogen protein atoms (1 kcal mol^{-1} $Å^{-2}$) using the Velocity-rescale thermostat fixed at 298.15° K with a 0.1 ps coupling constant. Next, a 15 ns restraint NPT equilibration simulation was run using a semiisotropic Berendsen barostat (P = 1 bar, pressure coupling time constant = 5 ps, compressibility = 4.5E-5 bar-1) and a Berendsen thermostat (T = 298.15° K, 1.0 ps time constant) while maintaining 1 kcal mol-1 Å-2 harmonic restraints on protein Cα atoms relative to the input structure. Unrestrained production dynamics simulations were then run for 200 ns with a Nose-Hoover thermostat at 298.15° K with a 1.0 ps time constant and a semiisotropic Parrinello-Rahman barostat fixed at 1 bar with a pressure coupling constant of 5 ps. Coordinate frames were extracted at 20 ps intervals in these production simulation trajectories. Three independent simulation trajectories were launched for each unique protein sequence, using different initial atomic velocities.

The backbone atoms from every frame in the trajectory versus the initial frame had an atomic RMSD in the range of 0.6-1.2 Å for all variant studies, and each triplicate trajectory. Likewise, across all variants, the average RMSD between any two frames throughout the MD trajectory was <1 Å. By these backbone RMSD metrics we can determine the protein fold and tertiary structure is very stable for all variants and deviates from the LLLL model on the same order as thermal fluctuations. The backbone RMSD of the medoid frame of each triplicate trajectory versus the LCP X-ray structures for the LQ variants were all <1.2 Å.

Analysis of Classical MD Simulations

Analysis of the channels and their water content was done using the Channel Annotation Package (CHAP, https://www.channotation.org) available from the Sansom Lab [44]. The trajectories from the production runs were centered to the protein using GROMACS software and

sampled every 200 ps to create a compressed xtc trajectory file for CHAP analysis. Water density plots (like those seen in Figs. 3 and 6 and Supplementary Figs. S1-S6) were generated for each 3 x 200 ns trajectory for each of the pentameric bundle designs. Similarly, the time-averaged water density profiles (i.e. panel c of Supplemental Figs. S1-S6) were generated by calculating the average water density for a given s over the course of the 200 ns simulation. These time-averaged water density profiles were then used to calculate the observed hydrophobic lengths (l_{obs}) reported in Fig. 2c and Extended Data Table 1 using a second derivative method.

Multiscale Reactive MD Simulations

The X-ray crystal structures for LQLL (pdb 7udz) and LLLL (pdb 6mct) were used as the starting structures for simulation. Classical simulations were first performed to equilibrate the protein structures and the simulation systems. Each protein was embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer and solvated with water using the CHARMM GUI [90-93], and the membrane and water were equilibrated using a standard equilibration protocol. Classical equilibration with no restraints was performed for 500ns. The CHARMM36 forcefield [94] was used to model all interactions, and simulations were run at 298K in the NPT ensemble using GROMACS [95].

MS-RMD [45-47, 96] was subsequently used to model the water and excess proton in all simulations used in our analyses. The MS-RMD method captures proton delocalization in water by allowing hydrogen-oxygen bonds to break and form. This is done by taking a linear combination of possible bonding topology states at every timestep. See our previous work for a detailed description and theory. The MS-EVB 3.2 parameters were used to describe the hydrated excess proton [97]. The excess proton center of excess charge (CEC) is defined as [98]:

$$\vec{r}_{CEC} = \sum_{i}^{N} c_i^2 \, \vec{r}_{COC}^{\,i} \ , \qquad (1)$$

Where $\vec{r}_{COC}^{\,i}$ is the coordinate of the center of excess charge of the ith diabatic state and c_i^2 is the amplitude of that state. The CEC defines the position of the delocalized excess proton. The CEC defines the position of the delocalized excess proton. The CHARMM36 forcefield was used to model the remaining interactions. Simulations were run at 298K in the NVT ensemble using LAMMPS [99] with the MS-RMD package.

Umbrella sampling simulations were performed in two dimensions to model the PT process. The first collective variable (CV) used is the CEC position along the channel axis, Z'CEC. The channel axis for each system was defined as the average principal component of the protein from

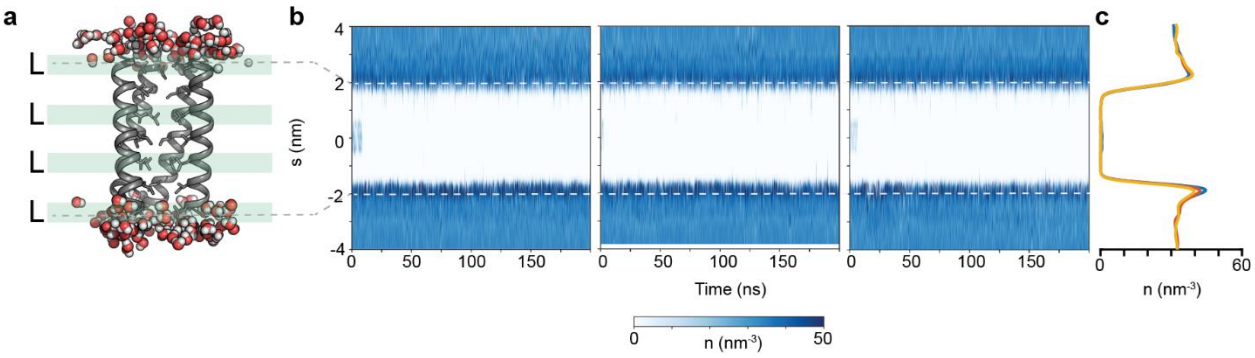
a 750ps MS-RMD simulation after equilibration. The position along this axis is calculated in reference to the center of mass of the Ile13 alpha-carbons, such that Z'CEC = 0Å at that point. The second CV used, φ, is a recently developed CV that measures the water connectivity within a channel using graph theory [8]. This CV is a significant improvement over water density, which does not directly bias the formation of a continuous water wire and can result in unphysical water "clumps" as the bias increases. Instead, the new water connectivity CV measures the length of transient water wire formations on a scale from 0 to 1, where 0 is no water and 1 is water fully connected throughout the channel, agnostic to the number of water molecules. We refer the reader to reference [35] for further theoretical details.

The open-source, community developed PLUMED library [100,101] was used to define the umbrella bias potentials. Umbrella sampling windows were set up every 0.5Å in the range [-22, 22] Å for Z'CEC and every 0.035 in the range [0.140,0.980] and [0.245,0.980] for φ for LLLL and LQLL, respectively, for a total of ~2000 windows for each system. Initial windows were generated by using steered MD to pull water into the channel, and the excess proton was placed at each point along the channel. Subsequent windows were pulled from nearby windows. Windows were equilibrated for 100ps, and then run for 0.5-3.5ns. A few additional windows were added to ensure sampling overlap.

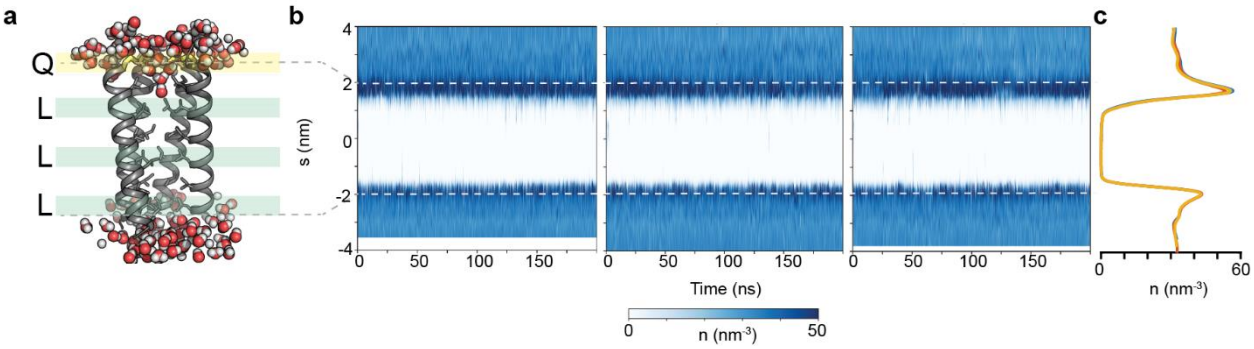
The 2D PMFs were calculated using WHAM-2D [102]. Error bars were calculated using the block method with four blocks. The minimum free energy path (MFEP) was calculated using the string method.

Figs. 5, Supplemental Figs. S17, S18, and S24 were generated with Matplotlib [103].

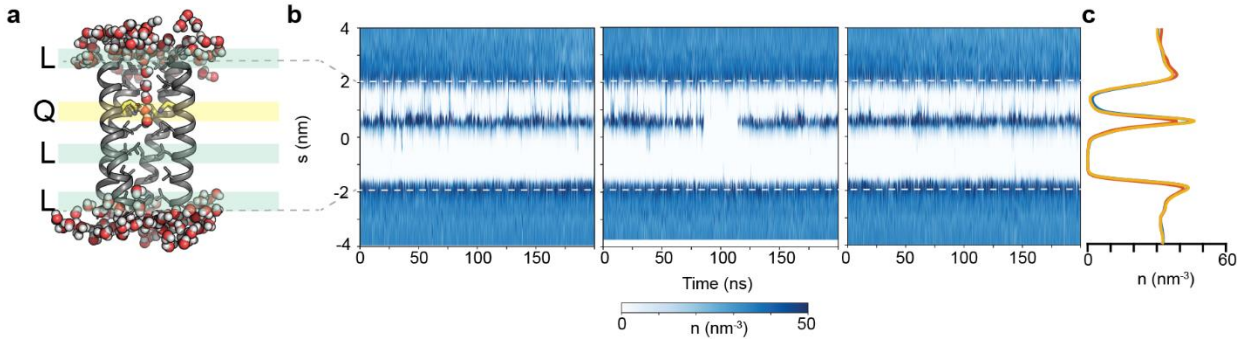
Fig. 5 was generated using Visual Molecular Dynamics (VMD) [87].



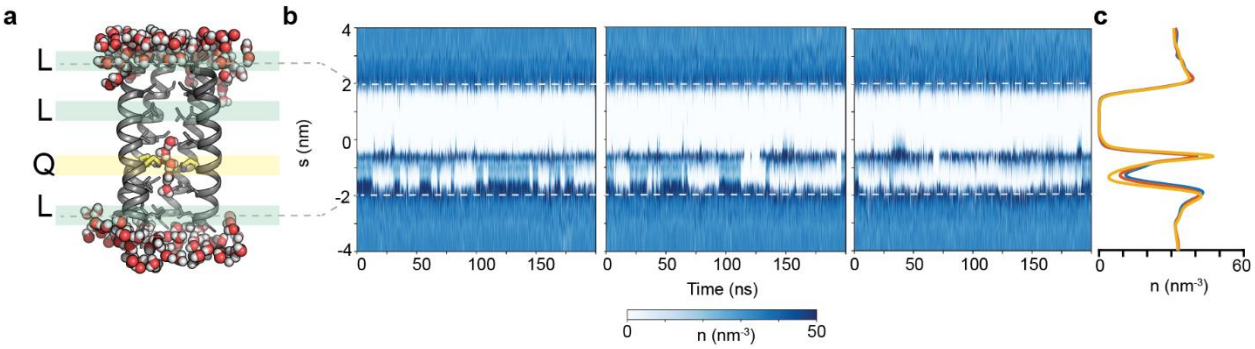
Supplementary Fig. 1. Analysis of classical MD simulations of LLLL. a, Snapshot taken from production run #1 of 3 of LLLL simulations. Fifth helix removed for clarity. **b,** Water density traces for all 3 x 200 ns trajectories show the channel (which spans s = -2 to 2 nm) devoid of water over the course of the simulation. **c,** Time-averaged water density as a function of s corresponding to the three independent simulations (blue, red, and yellow are productions #1, #2, and #3, respectively). From the second derivative of the trace, the l_{obs} of the channel was calculated to be 33.9 ± 0.3 Å. The leftmost water density plot is shown in Figs. 3a and 6a of the main text.



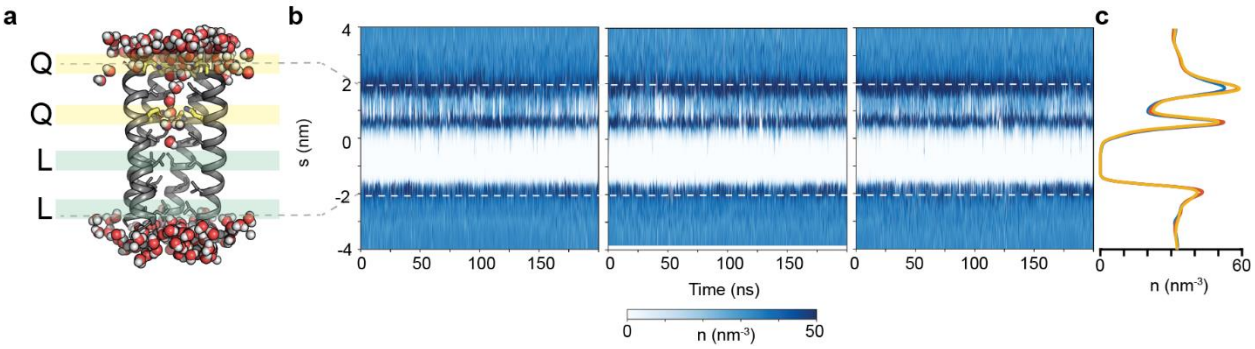
Supplementary Fig. 2. Analysis of classical MD simulations of QLLL. a, Snapshot taken from production run #1 of 3 of QLLL simulations. Fifth helix removed for clarity. **b,** Water density traces for all 3 x 200 ns trajectories show the channel (which spans s = -2 to 2 nm), like LLLL, devoid of water over the course of the simulation. Notably, there is an increase in water density near the Gln site, at s = +2 nm. **c,** Time-averaged water density as a function of s corresponding to the three independent simulations (blue, red, and yellow are productions #1, #2, and #3, respectively). From the second derivative of the trace, the l_{obs} of the channel was calculated to be 32.0 ± 0.6 Å. The leftmost water density plot is shown in Fig. 3b.



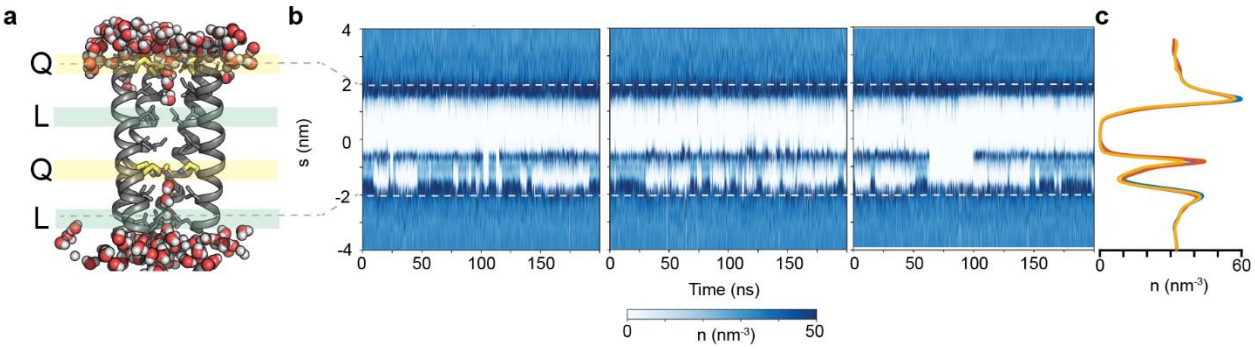
Supplementary Fig. 3. Analysis of classical MD simulations of LQLL. a, Snapshot taken from production run #2 of 3 of LQLL simulations. Fifth helix removed for clarity. **b,** Water density traces for all 3 x 200 ns trajectories show the channel (which spans s = -2 to 2 nm) with increased water densities within the pore near the Gln site, s = +0.5 nm. Moreover, there seems to be events in which multiple waters span the shorter segment. **c,** Time-averaged water density as a function of s corresponding to the three independent simulations (blue, red, and yellow are productions #1, #2, and #3, respectively). From the second derivative of the trace, the l_{obs} of the channel was calculated to be 20.7 ± 0.1 Å. The shorter hydrophobic path at the N-terminal end of the channel was calculated to be 11.6 ± 0.4 Å. The middle water density plot is shown in Figs. 3c and 6e.



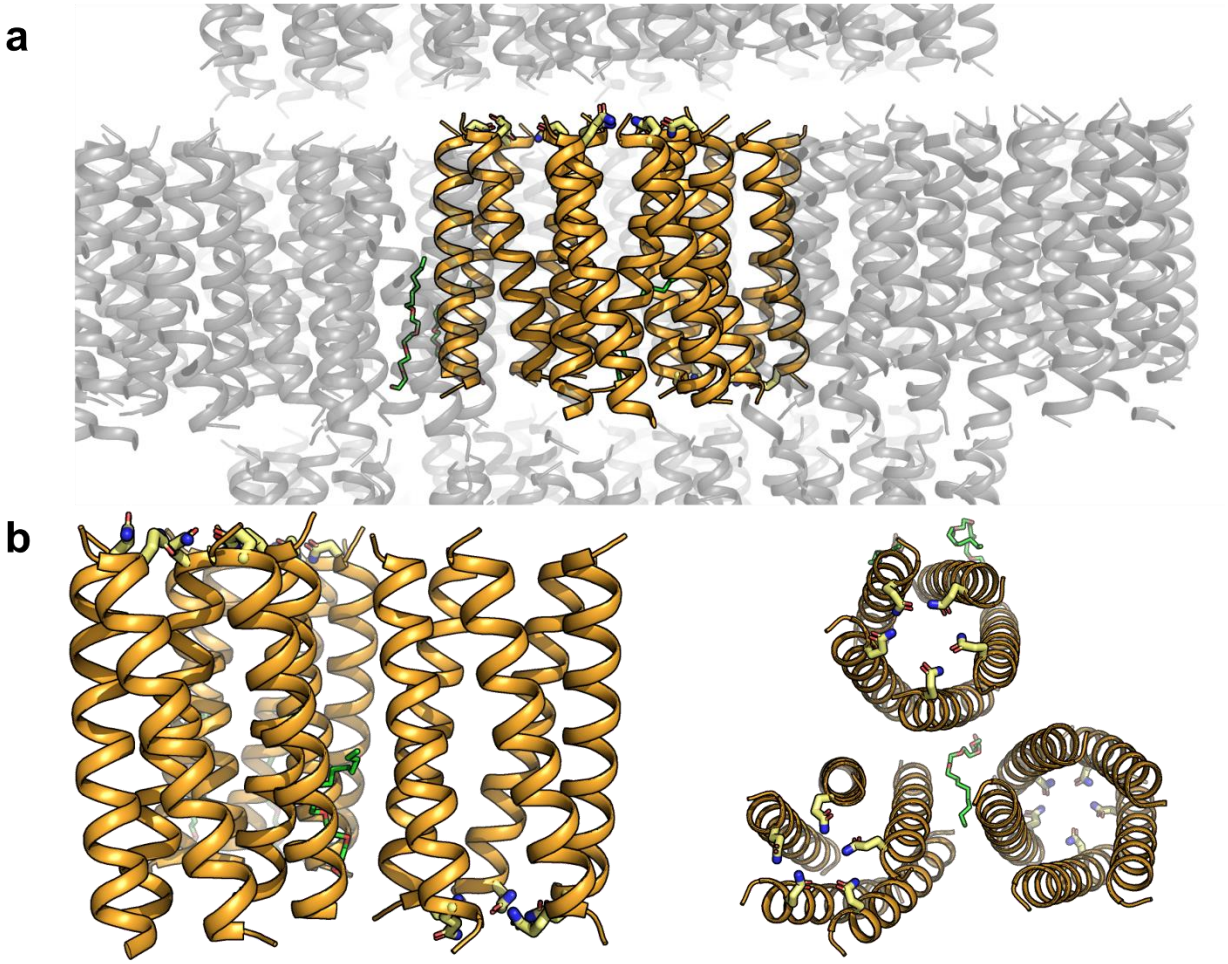
Supplementary Fig. 4. Analysis of classical MD simulations of LLQL. a, Snapshot taken from production run #1 of 3 of LLQL simulations. Fifth helix removed for clarity. **b,** Water density traces for all 3 x 200 ns trajectories show the channel (which spans s = -2 to 2 nm) with increased water densities within the pore near the Gln site, s = -0.5 nm. Moreover, there seems to be events in which multiple waters span the shorter segment. **c,** Time-averaged water density as a function of s corresponding to the three independent simulations (blue, red, and yellow are productions #1, #2, and #3, respectively). From the second derivative of the trace, the l_{obs} of the channel was calculated to be 22.4 ± 0.1 Å. The shorter hydrophobic path at the C-terminal end of the channel was calculated to be 9.3 ± 0.5 Å. The rightmost water density plot is shown in Fig. 3d.



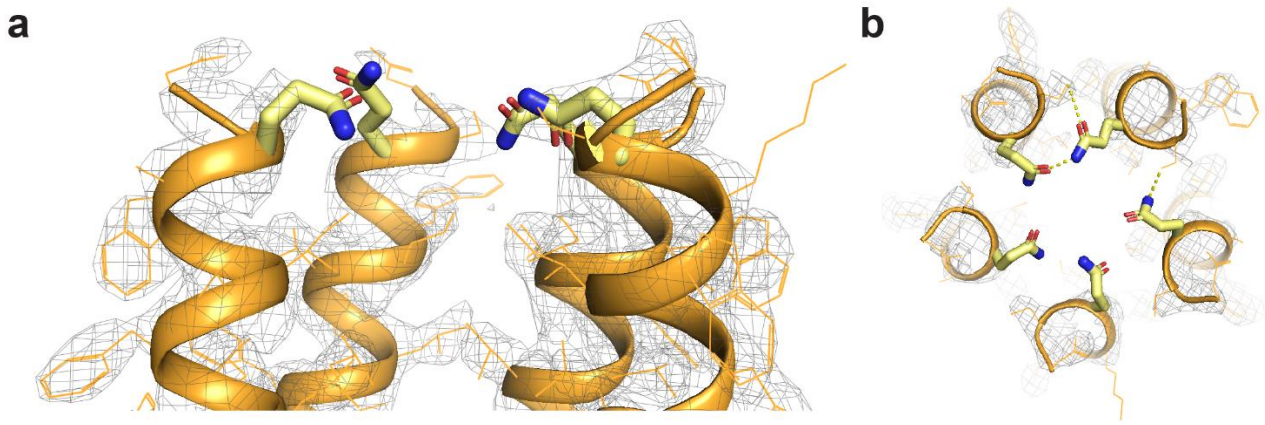
Supplementary Fig. 5. Analysis of classical MD simulations of QQLL. a, Snapshot taken from production run #1 of 3 of QQLL simulations. Fifth helix removed for clarity. **b,** Water density traces for all 3 x 200 ns trajectories show the channel (which spans s = -2 to 2 nm) with increased water densities within the pore near the Gln sites, at s = +2 and +0.5 nm. Moreover, there seems to be events in which multiple waters span the shorter segment. **c,** Time-averaged water density as a function of s corresponding to the three independent simulations (blue, red, and yellow are productions #1, #2, and #3, respectively). From the second derivative of the trace, the l_{obs} of the channel was calculated to be 20.9 ± 0.1 Å. The shorter hydrophobic path at the C-terminal end of the channel was calculated to be 8.2 ± 0.2 Å. The leftmost water density plot is shown in Figs. 3e and 6f.



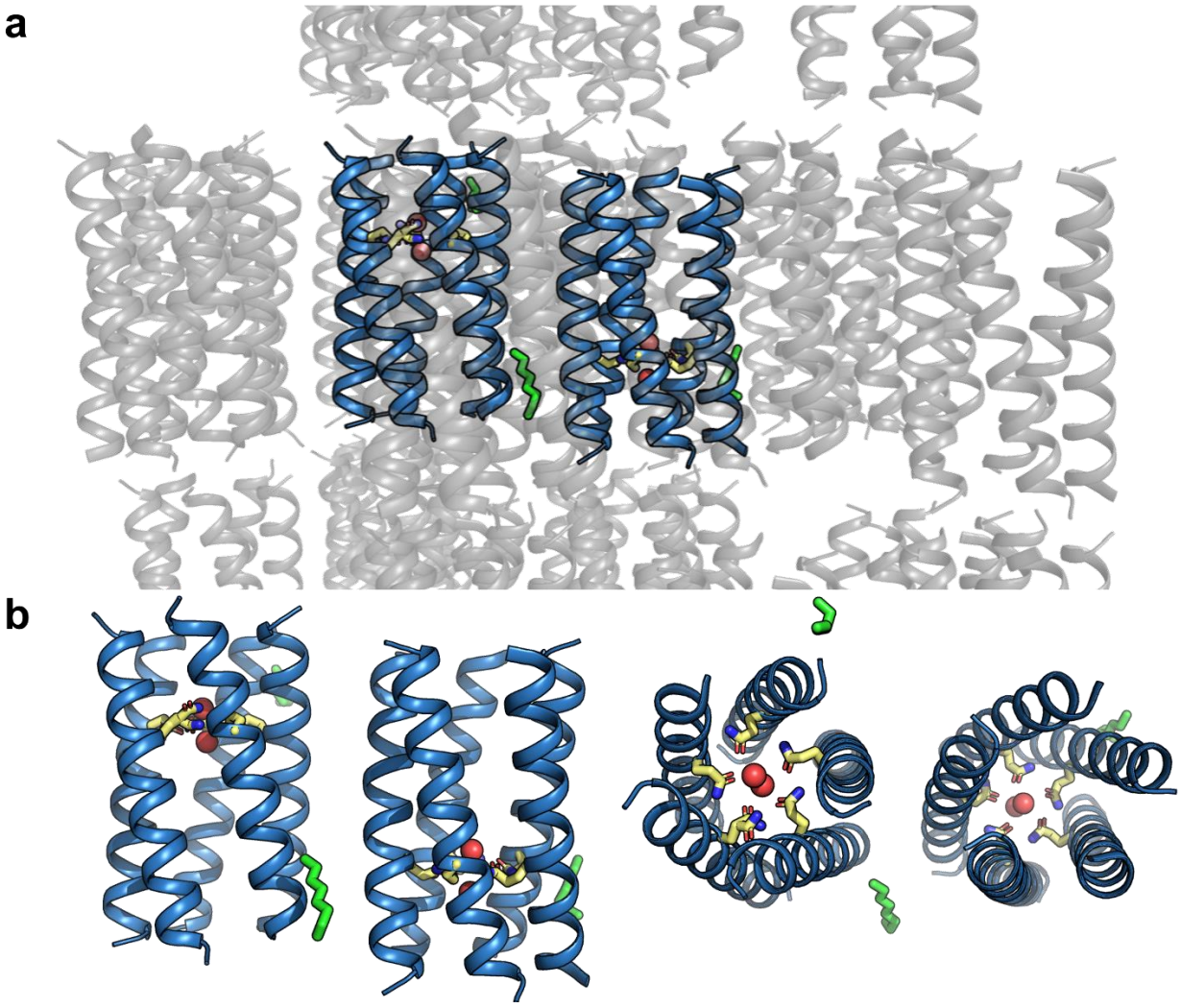
Supplementary Fig. 6. Analysis of classical MD simulations of QLQL. a, Snapshot taken from production run #1 of 3 of QLQL simulations. Fifth helix removed for clarity. **b,** Water density traces for all 3 x 200 ns trajectories show the channel (which spans s = -2 to 2 nm) with increased water densities within the pore near the Gln sites, at s = +2 and -0.5 nm. Moreover, there seems to be events in which multiple waters span the shorter segment. **c,** Time-averaged water density as a function of s corresponding to the three independent simulations (blue, red, and yellow are productions #1, #2, and #3, respectively). From the second derivative of the trace, the l_{obs} of the channel was calculated to be 20.3 ± 0.2 Å. The shorter hydrophobic path at the C-terminal end of the channel was calculated to be 9.2 ± 0.2 Å. The rightmost water density plot is shown in Fig. 3f.



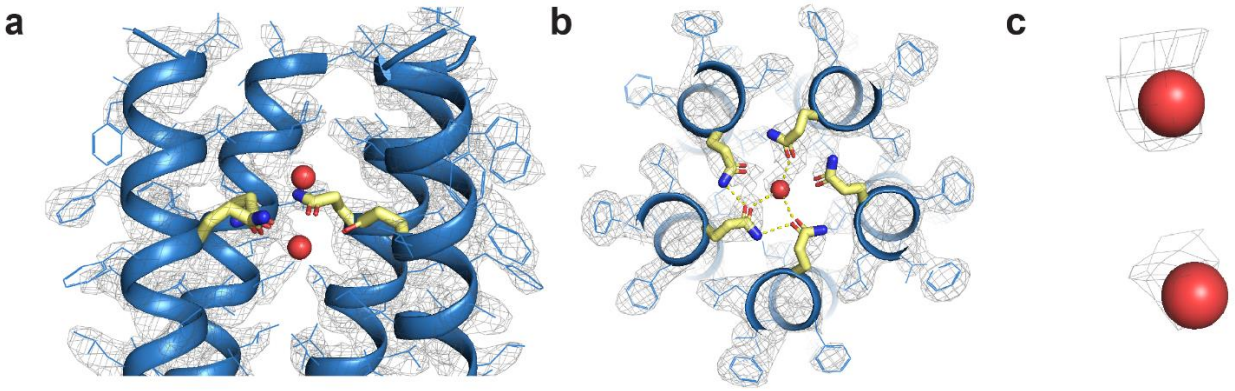
Supplementary Fig. 7. Crystal lattice of QLLL. a, Asymmetric unit (orange) in crystal lattice reveals three pentameric QLLL assemblies. **b,** Asymmetric unit is composed of three QLLL pentamers (2 up and 1 down). The three pentamers are structurally similar with rmsd values from 0.252 to 0.336 Å.



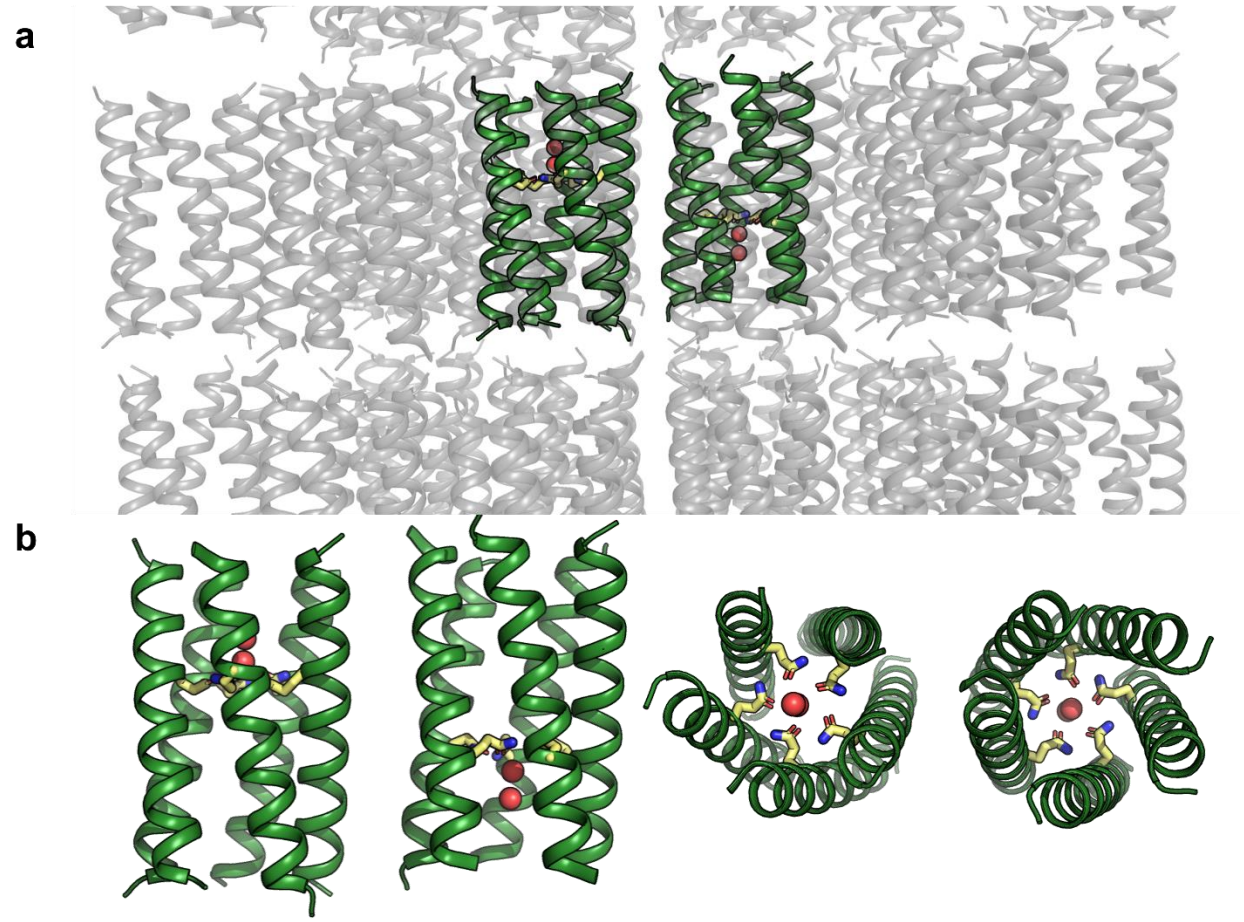
Supplementary Fig. 8. Structure of QLLL with electron density. a, Close-up of sideview with electron density (σ = 1.0) of one of the QLLL assemblies. **b,** Top view of Gln layer. Hydrogen bonds of < 3.2 Å shown in yellow dashes.



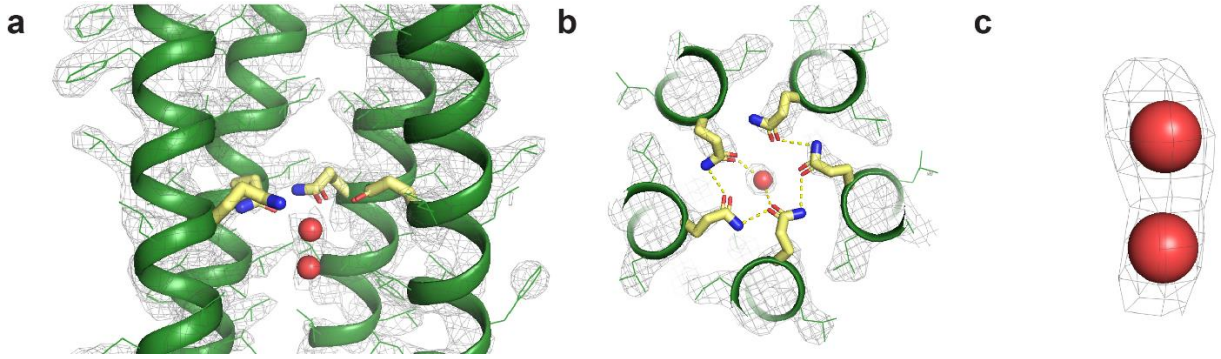
Supplementary Fig. 9. Crystal lattice of LQLL. a, Asymmetric unit (blue) in crystal lattice reveals two pentameric LQLL assemblies. **b,** Asymmetric unit is composed of two LQLL pentamers (one up and one down). The two structures in the asymmetric unit have an rmsd of 0.189 Å.



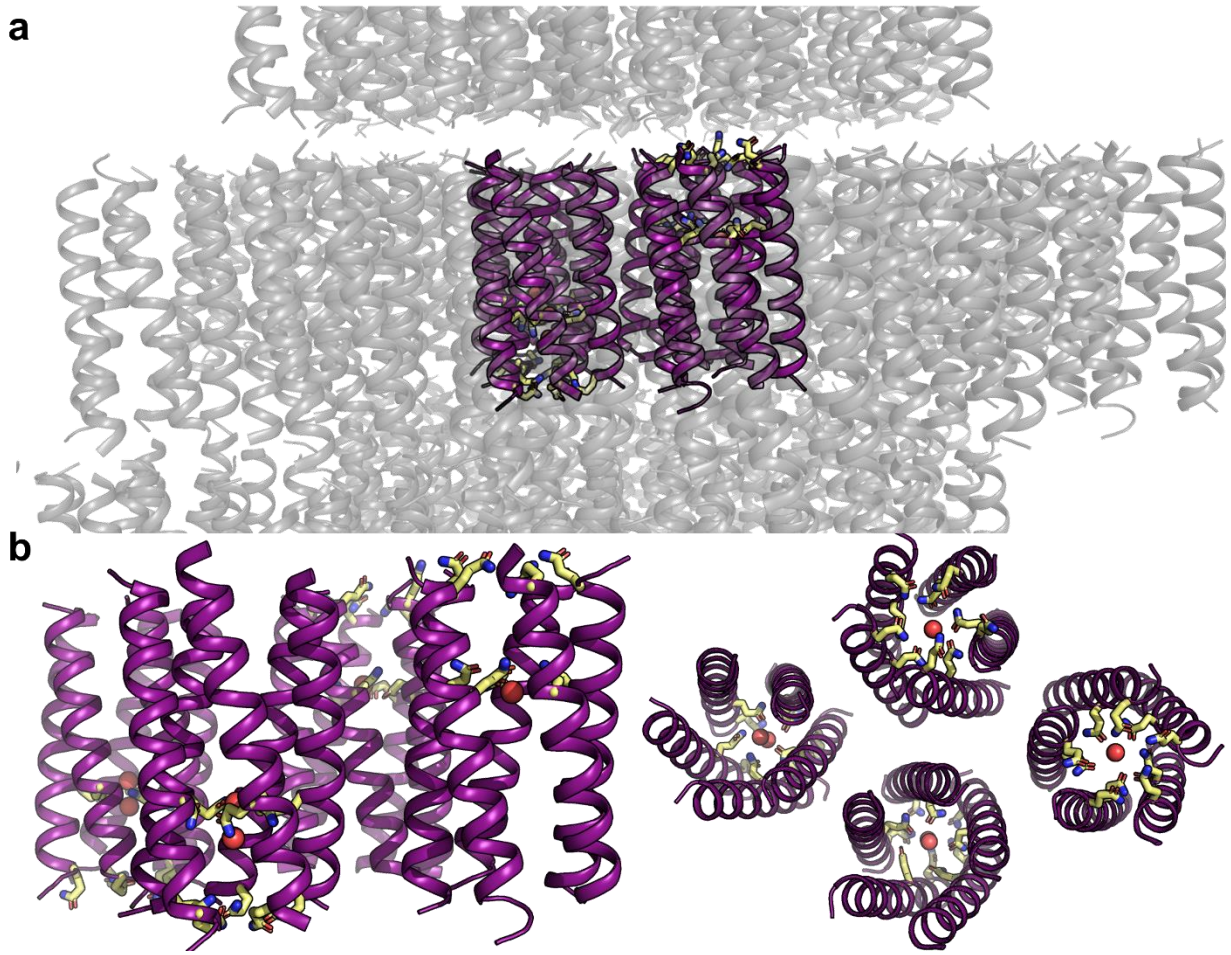
Supplementary Fig. 10. Structure of LQLL with electron density. a, Close-up of sideview with electron density (σ = 1.0) of one of the LQLL assemblies. **b,** Top view of Gln layer. Hydrogen bonds of < 3.2 Å shown in yellow dashes. **c,** Electron density (σ = 1.0) around waters in crystal structure.



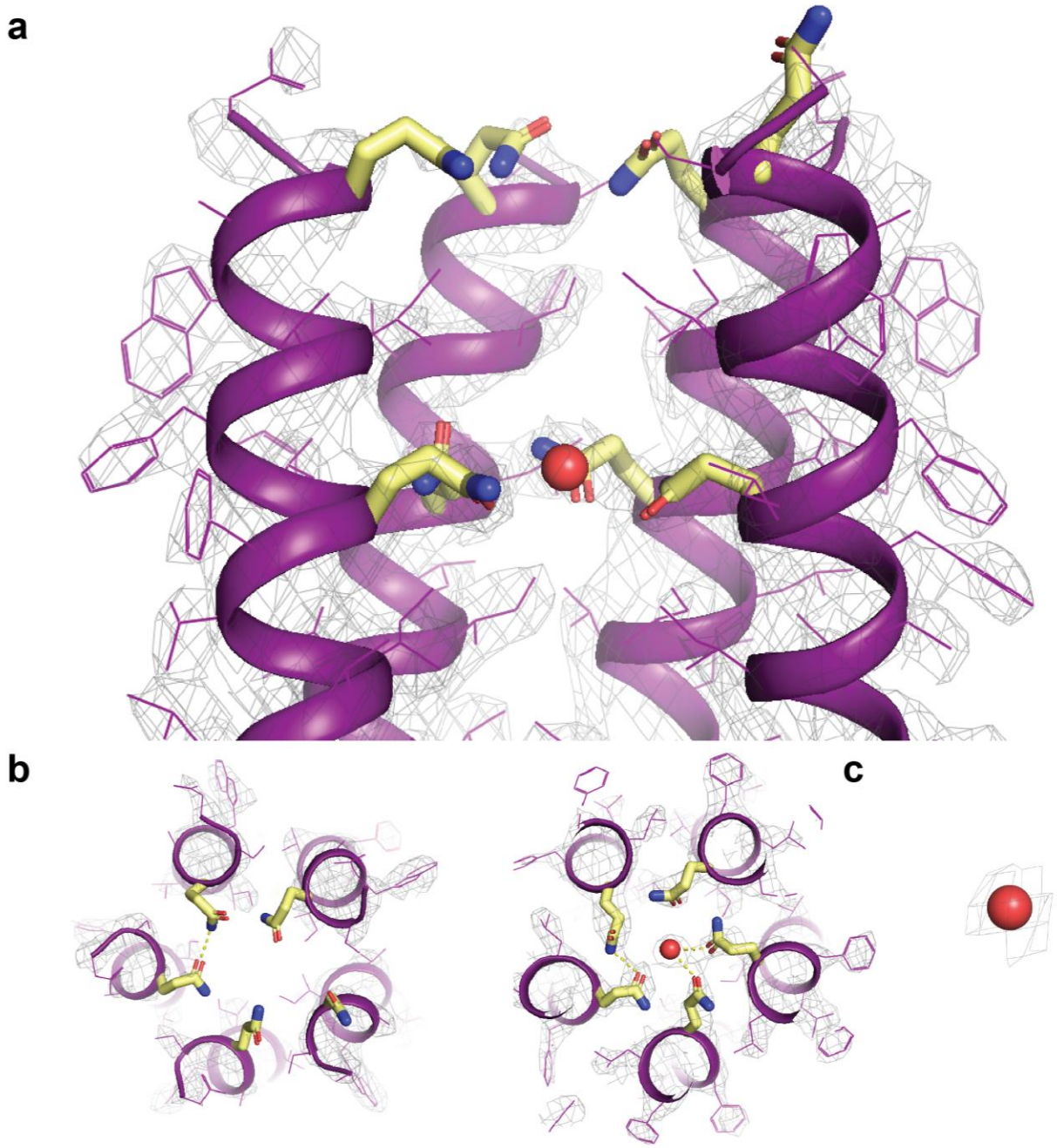
Supplementary Fig. 11. Crystal lattice of LLQL. a, Asymmetric unit (green) in crystal lattice reveals two pentameric LLQL assemblies. **b,** Asymmetric unit is composed of two LLQL pentamers (one up and one down). The structures of the two pentamers are within 0.337 Å rmsd of each other.



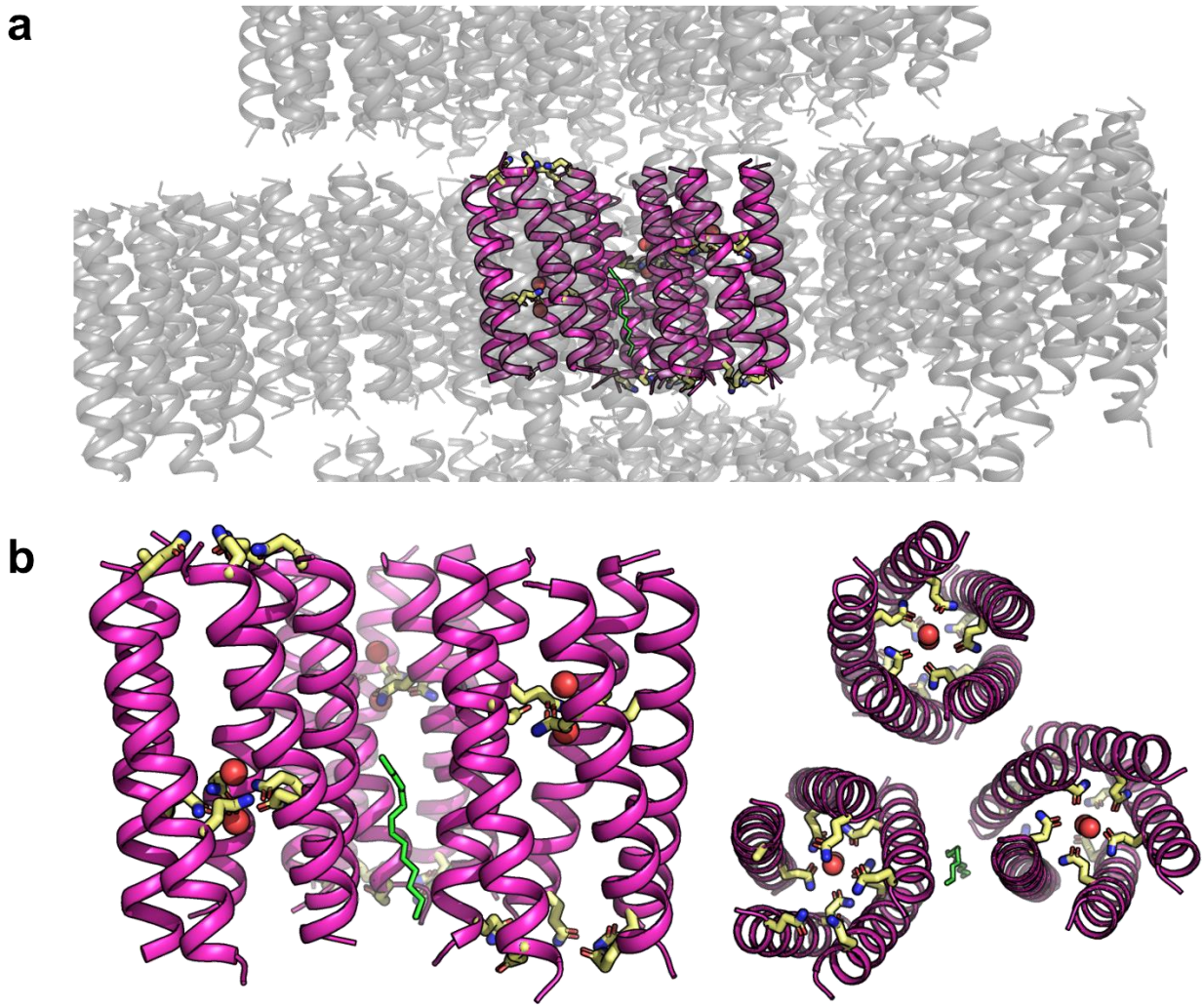
Supplementary Fig. 12. Structure of LLQL with electron density. a, Close-up of sideview with electron density (σ = 1.0) of one of the LLQL assemblies. **b,** Top view of Gln layer. Hydrogen bonds of < 3.2 Å shown in yellow dashes. **c,** Electron density (σ = 1.0) around waters in crystal structure.



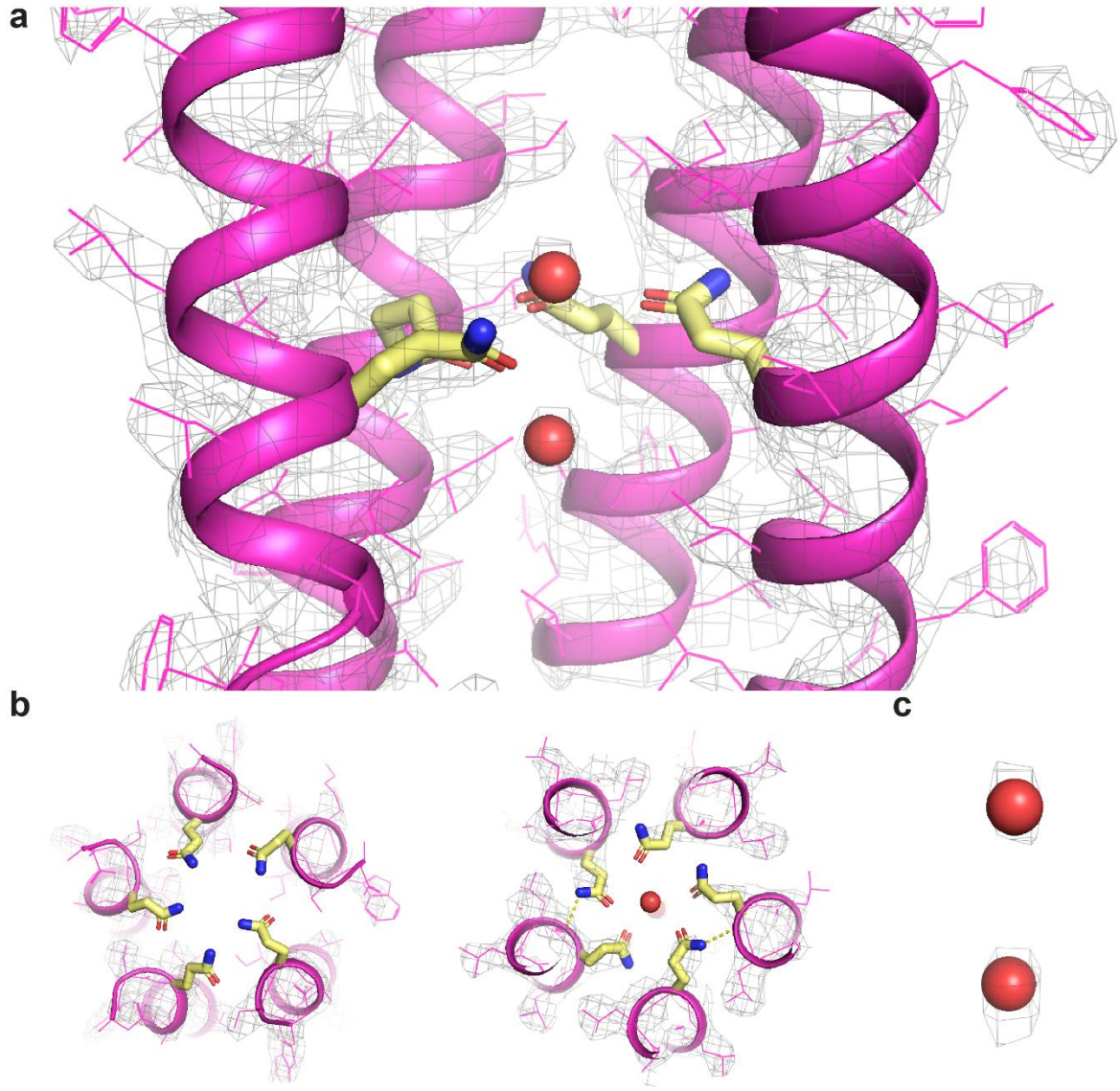
Supplementary Fig. 13. Crystal lattice of QQLL. a, Asymmetric unit (purple) in crystal lattice reveals four pentameric QQLL assemblies. **b,** Asymmetric unit is composed of four QQLL pentamers (2 up and 2 down). The backbone rmsd of the four bundles in the asymmetric unit vary from 0.275 to 1.62 Å.



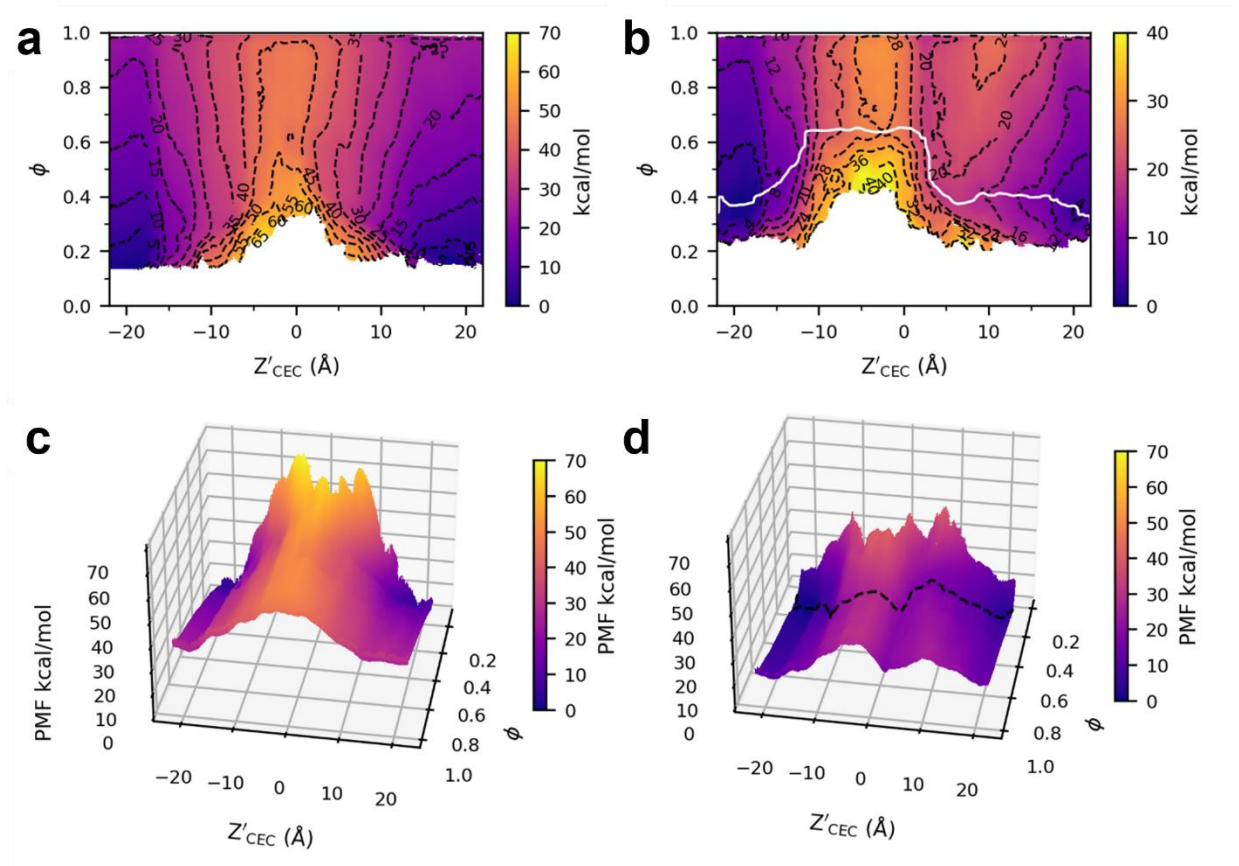
Supplementary Fig. 14. Structure of QQLL with electron density. a, Close-up of sideview with electron density of one of the QQLL assemblies. **b,**. Top view of Gln layer 1 (left) and Gln layer 2 (right). **c,** Electron density at σ = 1.0 for water.



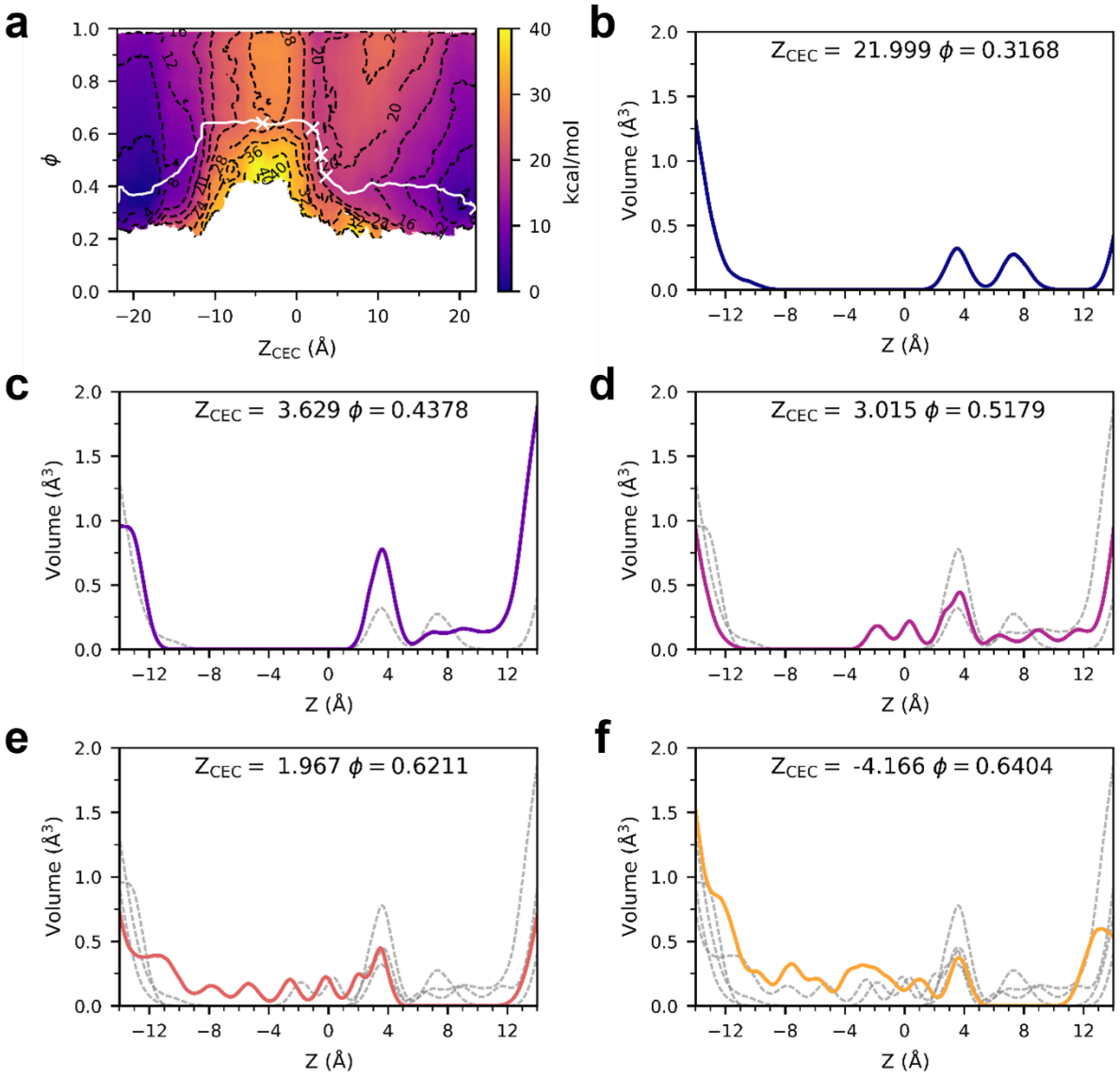
Supplementary Fig. 15. Crystal lattice of QLQL. a, Asymmetric unit (pink) in crystal lattice reveals three pentameric QLQL assemblies. **b,** Asymmetric unit is composed of three QLQL pentamers (2 up and 1 down). The three pentamers in the asymmetric unit have an rmsd between 0.376-0.427 Å.



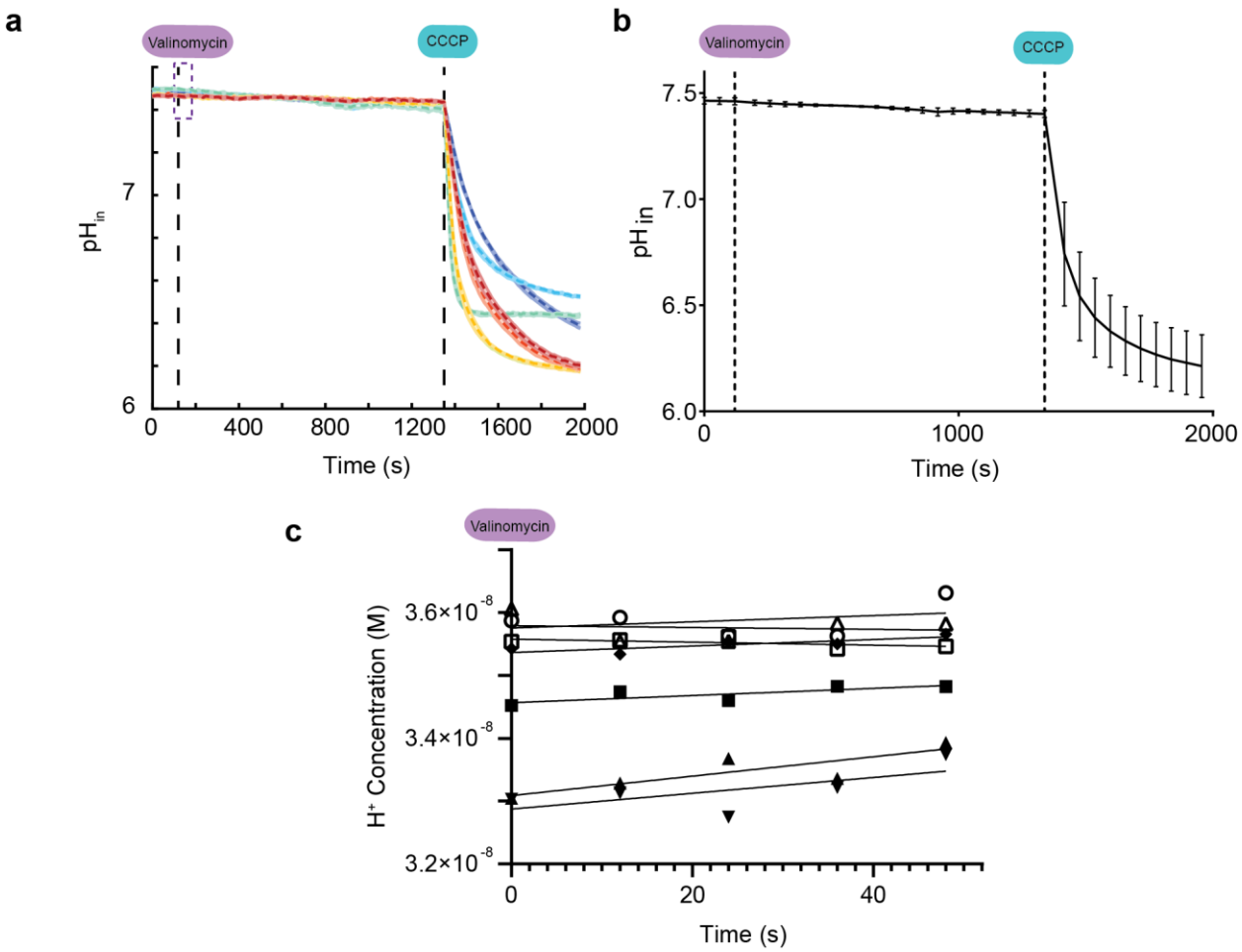
Supplementary Fig. 16. Structure of QLQL with electron density. a, Close-up of sideview with electron density (σ = 1.0) of one of the QLQL assemblies. **b,** Top view of Gln layer 1 (left) and Gln layer 2 (right). Hydrogen bonds of < 3.2 Å shown in yellow dashes. **c,** Electron density (σ = 1.0) around waters in crystal structure.



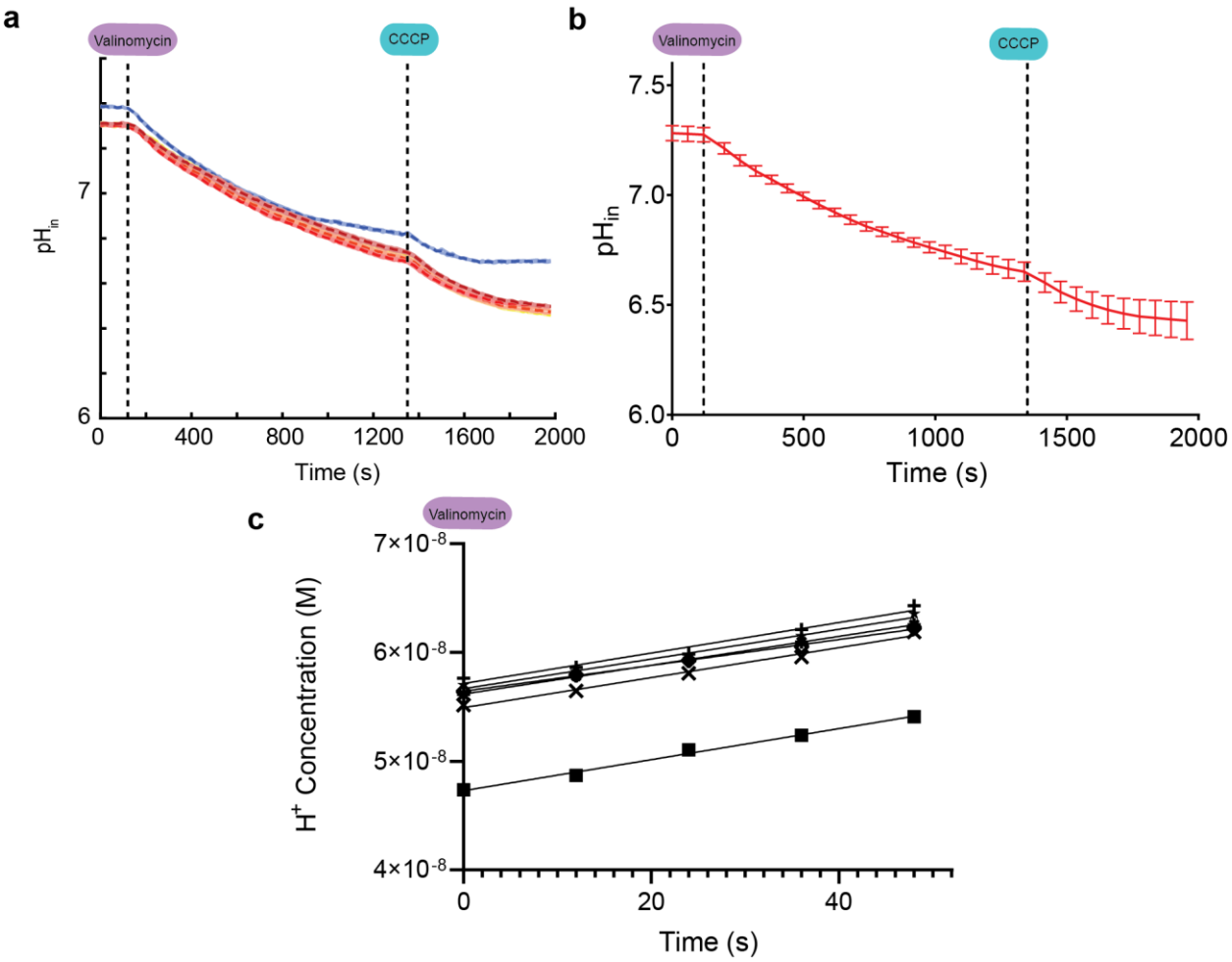
Supplementary Fig. 17. Generating MFEP from 2D PMFs. 2D PMFs for **a,** LLLL and **b,** LQLL, respectively. Note the different color scheme used for each for better contrast. Contours shown and labeled in each. In **b,**, the white line is the MFEP calculated using string theory. 3D depictions of the same PMFs for **c,** LLLL and **d,** LQLL, respectively, using the same scaling. The black dashed line in **d,** represents the MFEP.



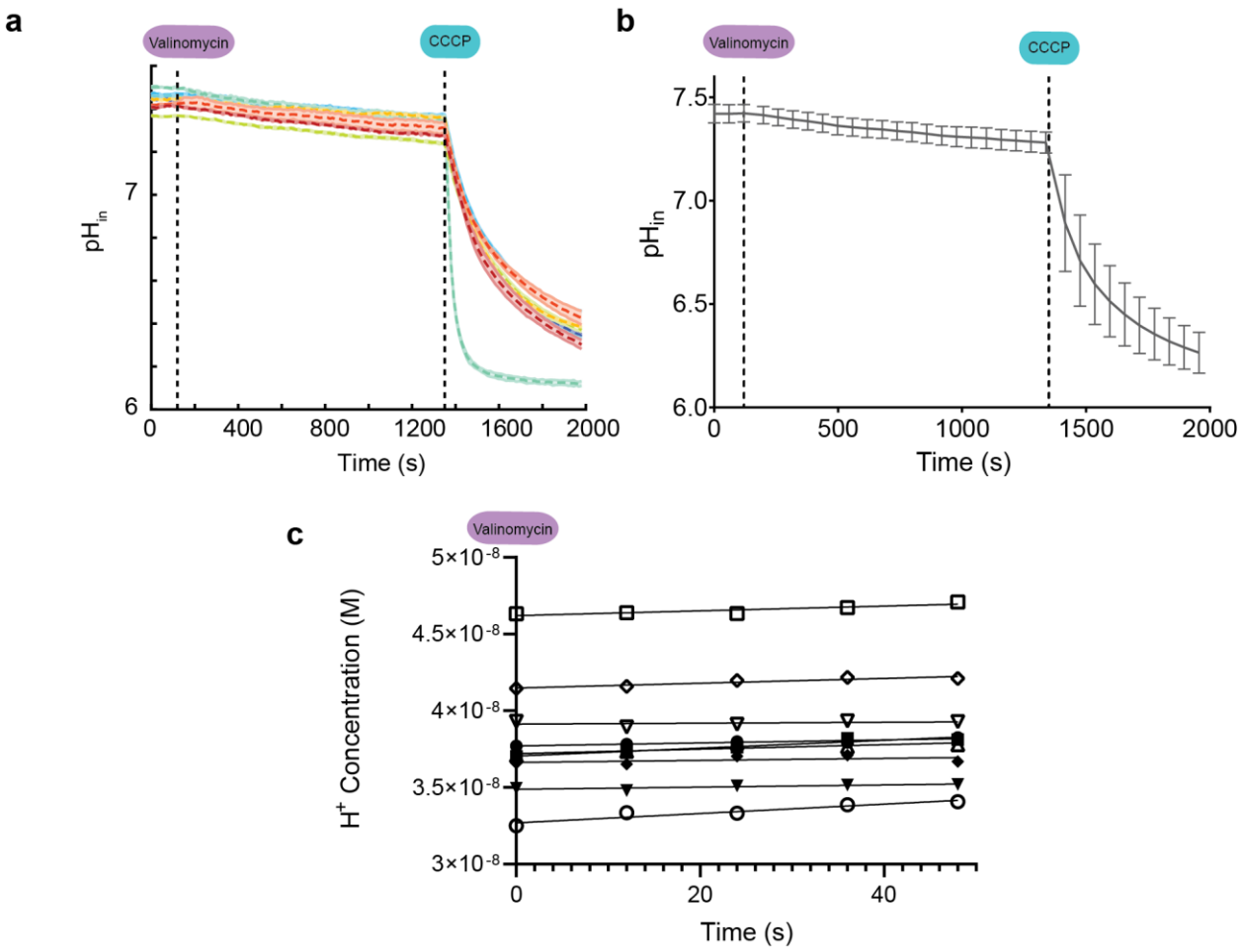
Supplementary Fig. 18. Average water volume as a function of Z'$_{CEC}$ and ϕ. a, Average volume of water in the channel shown for 5 different points along the LQLL MFEP, as indicated by white X's. **b,** The excess proton is out of the channel and the few waters in the channel are not connected to bulk. **c,** The proton is just below Gln and water connects to the top. **d,** The water structure transitions from the top to the bottom while the proton remains below Gln. **e,** The water is full connected in the bottom. **f,** The proton moves through the bottom-connected water wire. The dashed lines in C-F are the curves shown in all previous panels, b-e.



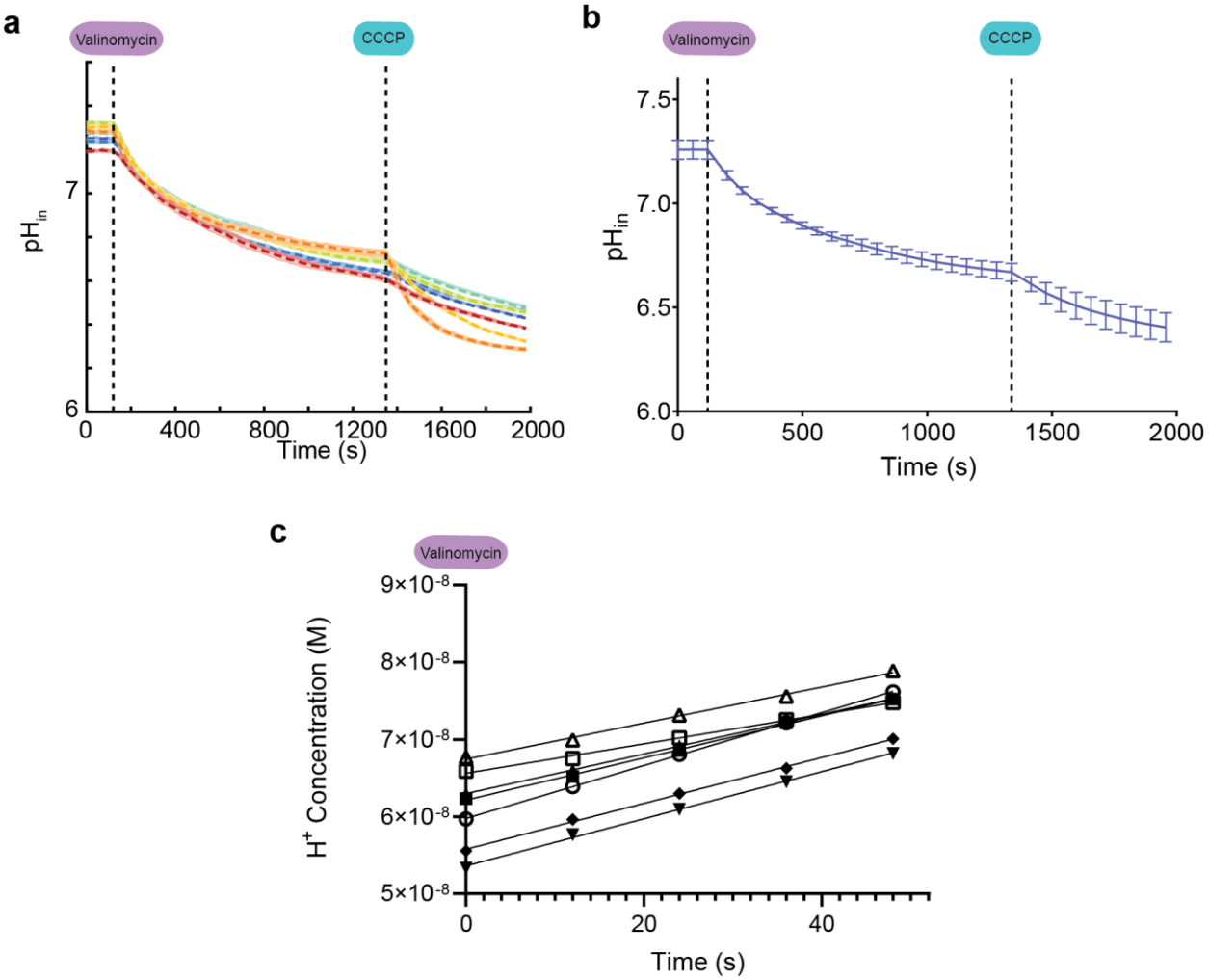
Supplementary Fig. 19. All proton flux assay data for "Empty" vesicles. Seven samples (each run in triplicate with shaded error bars shown) for empty (non-proteinaceous) vesicles that were run independently in assay. **a,** pH_{in} as a function of time throughout the measurement for each independent sample. Purple box shows time points used in analysis of initial rates in **c**. **b,** Mean and standard deviation for data collection. Data prior to CCCP step shown in Fig. 6. **c,** Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.



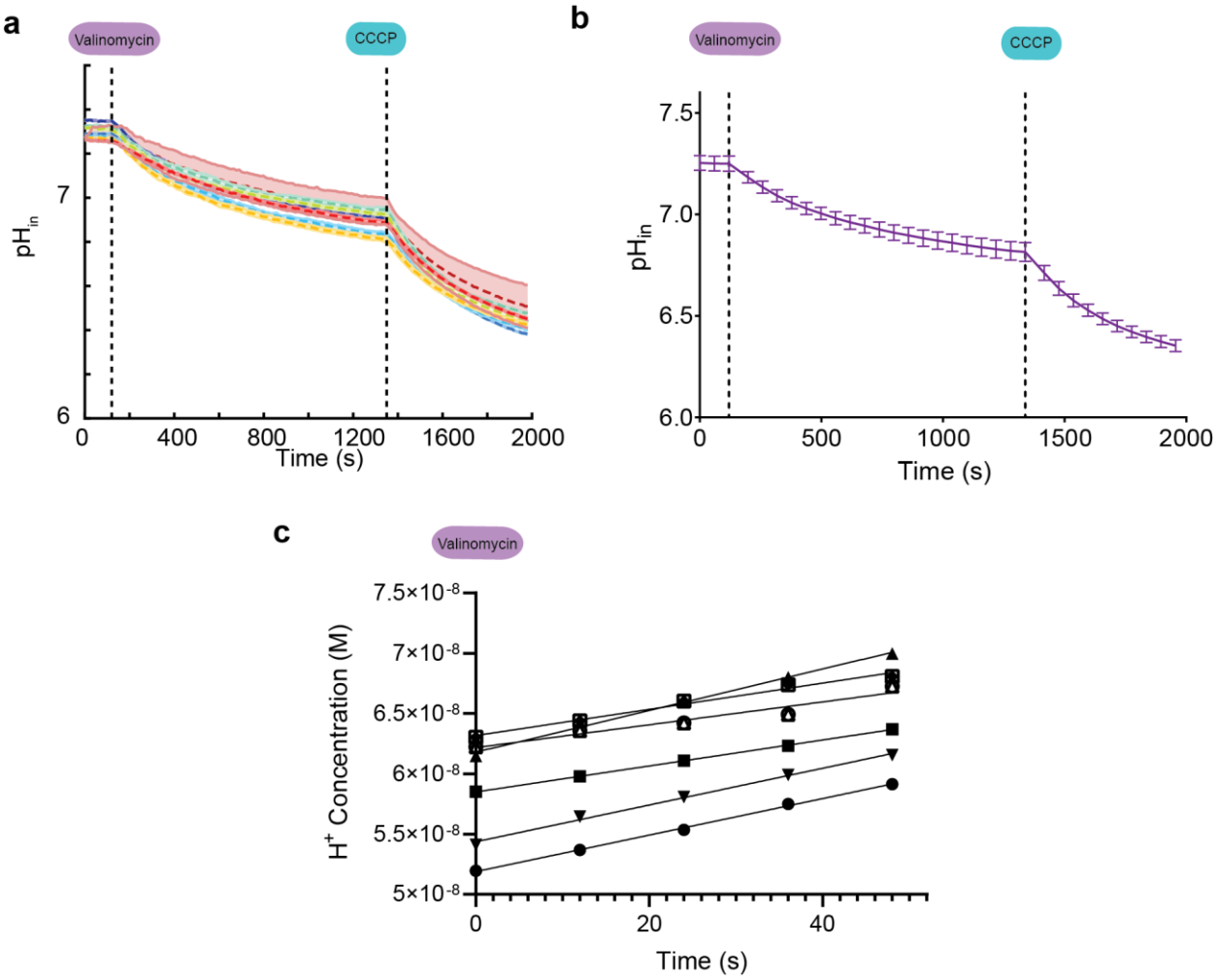
Supplementary Fig. 20. All proton flux assay data for Influenza A M2 channel vesicle samples. Six samples (each run in triplicate with shaded error bars shown) containing 1:500 polypeptide:lipid ratio; samples were run independently in the assay. **a,** pH_{in} as a function of time throughout the measurement for each independent sample. **b,** Mean and standard deviation for data collection. Data prior to CCCP step shown in Fig. 6. **c,** Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.



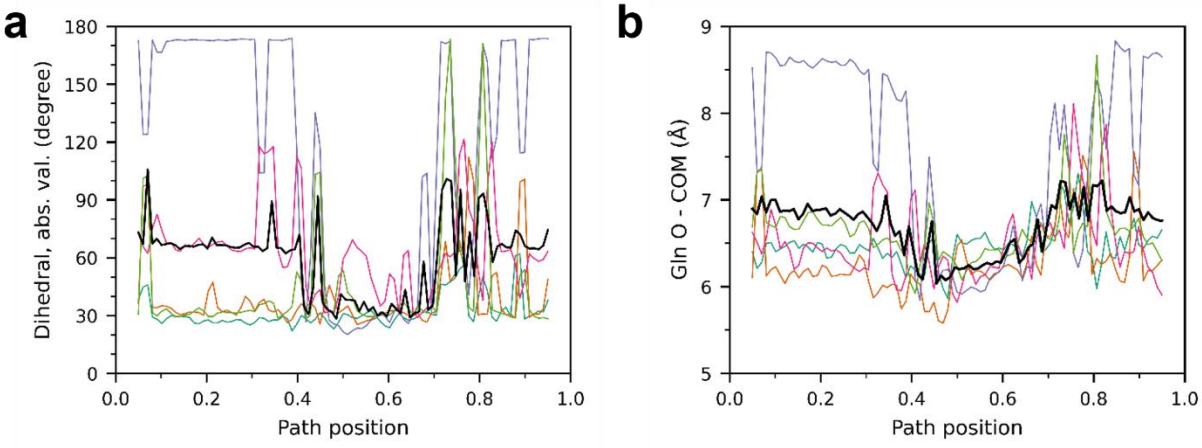
Supplementary Fig. 21. All proton flux assay data for LLLL vesicle samples. Nine samples (each run in triplicate with shaded error bars shown) containing 1:500 peptide:lipid ratio; samples were run independently in the assay. **a,** pH_{in} as a function of time throughout the measurement for each independent sample. **b,** Mean and standard deviation for data collection. Data prior to CCCP step shown in Fig. 6. **c,** Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.



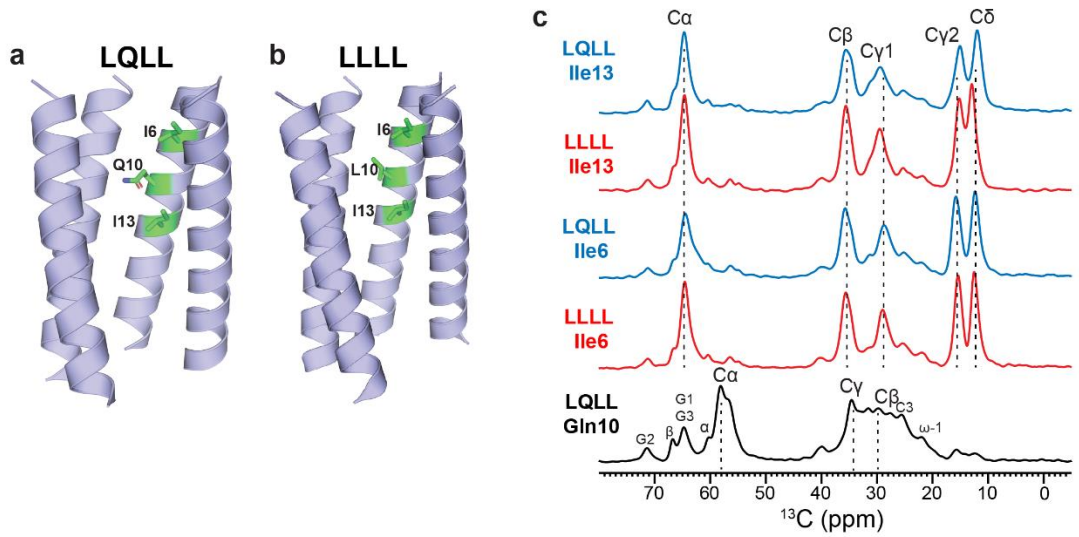
Supplementary Fig. 22. All proton flux assay data for LQLL vesicle samples. Seven samples (each run in triplicate with shaded error bars shown) containing 1:500 peptide:lipid ratio; samples were run independently in the assay. **a,** pH_{in} as a function of time throughout the measurement for each independent sample. **b,** Mean and standard deviation for data collection. Data prior to CCCP addition shown in Fig. 4e. **d,** Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.



Supplementary Fig. 23. All proton flux assay data for QQLL vesicle samples. Eight samples (each run in triplicate with shaded error bars shown) containing 1:500 peptide:lipid ratio; samples were run independently in the assay. **a,** pH_{in} as a function of time throughout the measurement for each independent sample. **b,** Mean and standard deviation for data collection. Data prior to CCCP addition shown in Fig. 6. **d,** Fits for the initial 50 seconds following addition of valinomycin. From the linear regression fits, all slopes (which give the initial rates (in M/s)) were used to calculate mean and standard error presented in Fig. 6g and Supplementary Table 3.



Supplementary Fig. 24. Analysis of Gln sidechain dihedrals reveals asymmetry in Gln conformations. a, The absolute value of the Gln sidechain dihedral for each Gln, as a function of the system position along the path. The average over all 5 residues is shown in black. **b,** The distance between the Gln carbonyl oxygen and the Ile13 center of mass, shown for each Gln. The average over all 5 residues is shown in black



Supplementary Fig. 25. Structure and 1D ^{13}C NMR spectra of membrane-bound LQLL and LLLL peptides. (a) Crystal structure of pentameric LQLL (pdb: 7udz), highlighting Ile6, Gln10, Ile13, which are isotopically labeled in this study. Fifth helix removed for clarity. (b) Crystal structure of pentameric LLLL (PDB: 6MCT). Ile6 and Ile13 are isotopically labeled in this study. One chain is removed for clarity. (c) 1D ^{13}C CP-MAS spectra of five DMPC membrane-bound LQLL and LLLL peptides, each containing a single labeled residue. All Ile and Gln Cα and Cβ chemical shifts indicate α-helical conformation. Minor signals of the DMPC natural abundance headgroup and glycerol backbone carbons are also observed in the Gln10 spectra. Ile6 in LLLL shows narrower linewidths than the other Ile residues in the other peptides, indicating higher conformational homogeneity. Gln10 exhibits broader linewidths, indicating conformational heterogeneity.

Supplementary Table 1. Designed peptide sequences and observed masses.

Peptides used for the experiments were synthesized and purified using protocols described in Materials and Methods. Samples were calibrated to bovine insulin (Sigma Aldrich) $[M+H]^+$ and $[M+2H]^{2+}$peaks.

Design	Sequence	Expected Mass (Da)	Observed Mass (Da)
LLLL	DS ^{3}LKWIVFL ^{10}LFLIVLL ^{17}LLAIVFL LRG	3028.1	3030.4
QLLL	DS ^{3}QKWIVFL ^{10}LFLIVLL ^{17}LLAIVFL LRG	3043.1	3045.0
LQLL	DS ^{3}LKWIVFL ^{10}QFLIVLL ^{17}LLAIVFL LRG	3043.1	3045.2
LLQL	DS ^{3}LKWIVFL ^{10}LFLIVLL ^{17}QLAIVFL LRG	3043.1	3044.4
QQLL	DS ^{3}QKWIVFL ^{10}QFLIVLL ^{17}LLAIVFL LRG	3058.0	3059.2
QLQL	DS ^{3}QKWIVFL ^{10}LFLIVLL ^{17}QLAIVFL LRG	3058.0	3059.3

Supplementary Table 2. Protein crystallization conditions

Solutions and conditions for crystallization of all designed proton channels. All crystals prepped using LCP methods described in Materials and Methods at 30 mg/mL in MAG 9.9 with 50 mM OG.

Design	Crystallization Condition
QLLL	0.22 M sodium citrate, 0.1 M Tris pH 8.0, 35% v/v PEG 400
LQLL	0.1 M calcium chloride, 0.1 M MES, 29% v/v PEG 400
LLQL	0.2 M magnesium chloride, 0.1 M potassium chloride, 0.03 M sodium citrate pH 4, 33% v/v PEG 400
QQLL	0.22 M sodium citrate, 0.1 M Tris pH 8.0
QLQL	0.2 M ammonium sulfate, 0.1 M sodium chloride, 0.1 M sodium citrate pH 6.0, 20% w/v PEG 2000

Supplementary Table 3. Descriptive statistics for all vesicle samples in proton flux assays. Mean and standard deviation/error of initial rates for all vesicle samples run for data presented in Fig. 6, Extended Data Figs. 4-7, Supplementary Figs. S19-S23.

	Empty	Influenza A M2	LLLL	QLLL	LQLL	LLQL	QQLL	QLQL
Number of values	7	6	9	8	7	8	8	7
Minimum	-2.382e-012	1.190e-010	2.974e-012	1.855e-011	1.908e-010	1.518e-010	9.487e-011	1.350e-010
Maximum	1.546e-011	1.428e-010	3.056e-011	4.342e-011	3.432e-010	1.767e-010	1.726e-010	1.778e-010
Range	1.784e-011	2.380e-011	2.759e-011	2.487e-011	1.524e-010	2.490e-011	7.773e-011	4.280e-011
Mean	5.716e-012	1.350e-010	1.432e-011	3.297e-011	2.716e-010	1.646e-010	1.240e-010	1.551e-010
Std. Deviation	6.576e-012	8.494e-012	9.053e-012	1.046e-011	5.017e-011	1.077e-011	3.013e-011	1.497e-011
Std. Error of Mean	2.486e-012	3.468e-012	3.018e-012	3.699e-012	1.896e-011	3.807e-012	1.065e-011	5.659e-012
Lower 95% CI of mean	-3.663e-013	1.260e-010	7.358e-012	2.422e-011	2.252e-010	1.556e-010	9.884e-011	1.412e-010
Upper 95% CI of mean	1.180e-011	1.439e-010	2.127e-011	4.172e-011	3.180e-010	1.736e-010	1.492e-010	1.689e-010

Supplementary Table 4. Unpaired t-test for control samples: empty vesicle initial rates vs. Influenza A M2 proton channel. Comparison of initial rates for the control samples reveals that rates measured for vesicles containing the natural proton channel (influenza A M2 channel) are statistically significant from empty vesicles. Two-tailed p value reported as p = 4.78E-12.

Unpaired t test between Influenza A M2 channel and empty vesicles	
P value	<0.0001
P value summary	****
Significantly different (P < 0.05)?	Yes
One- or two-tailed P value?	Two-tailed
t, df	t=30.93, df=11

Supplementary Table 5. Summary of ordinary one-way ANOVA for multiple comparisons (Dunn's test) of LLLL vs. designed channels. Comparison of initial rates from Figure 6g shows that initial rates of LLLL and QLLL are not statistically significant. However, initial rates of conduction for LQLL, LLQL, QQLL and QLQL are statistically significant when compared to LLLL.

Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Summary	Adjusted P Value
LLLL vs. QLLL	-1.9E-11	-5.001E-011 to 1.270E-011	ns	0.405
LLLL vs. LQLL	-2.6E-10	-2.898E-010 to -2.248E-010	****	<0.0001
LLLL vs. LLQL	-1.5E-10	-1.816E-010 to -1.189E-010	****	<0.0001
LLLL vs. QQLL	-1.1E-10	-1.411E-010 to -7.836E-011	****	<0.0001
LLLL vs. QLQL	-1.4E-10	-1.696E-010 to -1.069E-010	****	<0.0001

Supplementary Table 6. Summary of ordinary one-way ANOVA for multiple comparisons (Dunn's test) of QLLL vs. designed channels. Comparing the initial rates of proton flux of all channels to QLLL. Similarly, this analysis reveals that there is a statistically significant difference between initial rates of the non-conductive QLLL/LLLL with the other four proton-conductive designed channels.

Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Summary	Adjusted P Value
QLLL vs. LLLL	1.87E-11	-1.258E-011 to 4.988E-011	ns	0.3947
QLLL vs. LQLL	-2.4E-10	-2.719E-010 to -2.054E-010	****	<0.0001
QLLL vs. LLQL	-1.3E-10	-1.638E-010 to -9.948E-011	****	<0.0001
QLLL vs. QQLL	-9.1E-11	-1.232E-010 to -5.893E-011	****	<0.0001
QLLL vs. QLQL	-1.2E-10	-1.517E-010 to -8.747E-011	****	<0.0001

Supplementary Movie 1. Snapshots from MS-RMD simulations along MFEP (as shown in Fig. 5) depict the proton (represented as a yellow hydronium in the movie) inducing the formation of water wires in hydrophobic regions of the channel lumen. The proton moves along these transiently forming water wires to make its way down to a region just below the PLS formed by the green Gln residues. Once near this site, the presence of the proton enables the formation of a secondary water wire through the longer hydrophobic region. This allows the proton to traverse the rest of the way across the channel.

Complete references and notes:

1 Moriyama, Y. & Futai, M. H+-ATPase, a primary pump for accumulation of neurotransmitters, is a major constituent of brain synaptic vesicles. *Biochemical and biophysical research communications* **173**, 443-448, doi:10.1016/s0006-291x(05)81078-2 (1990).

2 Nishi, T. & Forgac, M. The vacuolar (H+)-ATPases--nature's most versatile proton pumps. *Nature reviews. Molecular cell biology* **3**, 94-103, doi:10.1038/nrm729 (2002).

3 Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144-148, doi:10.1038/191144a0 (1961).

4 Nicholls, D. G. Mitochondrial ion circuits. *Essays in biochemistry* **47**, 25-35, doi:10.1042/bse0470025 (2010).

5 Diering, G. H. & Numata, M. Endosomal pH in neuronal signaling and synaptic transmission: role of Na(+)/H(+) exchanger NHE5. *Frontiers in physiology* **4**, 412, doi:10.3389/fphys.2013.00412 (2014).

6 Agmon, N. The Grotthuss mechanism. *Chemical Physics Letters* **244**, 456-462, doi:10.1016/0009-2614(95)00905-j (1995).

7 Calio, P. B., Li, C. & Voth, G. A. Resolving the Structural Debate for the Hydrated Excess Proton in Water. *J. Am. Chem. Soc.* **143**, 18672-18683, doi:10.1021/jacs.1c08552 (2021).

8 Li, C. & Voth, G. A. A quantitative paradigm for water-assisted proton transport through proteins and other confined spaces. *Proceedings of the National Academy of Sciences of the United States of America* **118**, doi:10.1073/pnas.2113141118 (2021).

9 Wraight, C. A. Chance and design--proton transfer in water, channels and bioenergetic proteins. *Biochimica et biophysica acta* **1757**, 886-912, doi:10.1016/j.bbabio.2006.06.017 (2006).

10 Decoursey, T. E. Voltage-gated proton channels and other proton transfer pathways. *Physiol Rev* **83**, 475-579, doi:10.1152/physrev.00028.2002 (2003).

11 Peng, Y., Swanson, J. M., Kang, S. G., Zhou, R. & Voth, G. A. Hydrated Excess Protons Can Create Their Own Water Wires. *The journal of physical chemistry. B* **119**, 9212-9218, doi:10.1021/jp5095118 (2015).

12 Banh, R. *et al.* Hydrophobic gasket mutation produces gating pore currents in closed human voltage-gated proton channels. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 18951-18961, doi:10.1073/pnas.1905462116 (2019).

13 Garczarek, F. & Gerwert, K. Functional waters in intraprotein proton transfer monitored by FTIR difference spectroscopy. *Nature* **439**, 109-112, doi:10.1038/nature04231 (2006).

14 Kaur, D., Khaniya, U., Zhang, Y. & Gunner, M. R. Protein Motifs for Proton Transfers That Build the Transmembrane Proton Gradient. *Frontiers in chemistry* **9**, 660954, doi:10.3389/fchem.2021.660954 (2021).

15 Kalra, A., Garde, S. & Hummer, G. Osmotic water transport through carbon nanotube membranes. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10175-10180, doi:10.1073/pnas.1633354100 (2003).

16 Ben-Abu, Y., Zhou, Y., Zilberberg, N. & Yifrach, O. Inverse coupling in leak and voltage-activated K+ channel gates underlies distinct roles in electrical signaling. *Nature structural & molecular biology* **16**, 71-79, doi:10.1038/nsmb.1525 (2009).

17 Jensen, M. O. *et al.* Principles of conduction and hydrophobic gating in K+ channels. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 5833-5838, doi:10.1073/pnas.0911691107 (2010).

18 Aryal, P., Sansom, M. S. & Tucker, S. J. Hydrophobic gating in ion channels. *J Mol Biol* **427**, 121-130, doi:10.1016/j.jmb.2014.07.030 (2015).

19 Zhu, F. & Hummer, G. Drying transition in the hydrophobic gate of the GLIC channel blocks ion conduction. *Biophysical journal* **103**, 219-227, doi:10.1016/j.bpj.2012.06.003 (2012).

20 Rasaiah, J. C., Garde, S. & Hummer, G. Water in nonpolar confinement: from nanotubes to proteins and beyond. *Annual review of physical chemistry* **59**, 713-740, doi:10.1146/annurev.physchem.59.032607.093815 (2008).
21 Wang, T. *et al.* Deprotonation of D96 in bacteriorhodopsin opens the proton uptake pathway. *Structure* **21**, 290-297, doi:10.1016/j.str.2012.12.018 (2013).
22 Weinert, T. *et al.* Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography. *Science* **365**, 61-65, doi:10.1126/science.aaw8634 (2019).
23 Freier, E., Wolf, S. & Gerwert, K. Proton transfer via a transient linear water-molecule chain in a membrane protein. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11435-11439, doi:10.1073/pnas.1104735108 (2011).
24 Regan, L. & DeGrado, W. F. Characterization of a helical protein designed from first principles. *Science* **241**, 976-978, doi:10.1126/science.3043666 (1988).
25 Walsh, S. T., Cheng, H., Bryson, J. W., Roder, H. & DeGrado, W. F. Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 5486-5491, doi:10.1073/pnas.96.10.5486 (1999).
26 Kuhlman, B. *et al.* Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364-1368, doi:10.1126/science.1089427 (2003).
27 Vorobieva, A. A. *et al.* De novo design of transmembrane beta barrels. *Science* **371**, doi:10.1126/science.abc8182 (2021).
28 Yang, C. *et al.* Bottom-up de novo design of functional proteins with complex structural features. *Nat Chem Biol* **17**, 492-500, doi:10.1038/s41589-020-00699-x (2021).
29 Polizzi, N. F. & DeGrado, W. F. A defined structural unit enables de novo design of small-molecule-binding proteins. *Science* **369**, 1227-1233, doi:10.1126/science.abb8330 (2020).
30 Cao, L. *et al.* De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *bioRxiv : the preprint server for biology*, doi:10.1101/2020.08.03.234914 (2020).
31 Fleishman, S. J. *et al.* Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816-821, doi:10.1126/science.1202617 (2011).
32 Jiang, L. *et al.* De novo computational design of retro-aldol enzymes. *Science* **319**, 1387-1391, doi:10.1126/science.1152692 (2008).
33 Lassila, J. K., Privett, H. K., Allen, B. D. & Mayo, S. L. Combinatorial methods for small-molecule placement in computational enzyme design. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16710-16715, doi:10.1073/pnas.0607691103 (2006).
34 Polizzi, N. F. *et al.* De novo design of a hyperstable non-natural protein-ligand complex with sub-A accuracy. *Nature chemistry* **9**, 1157-1164, doi:10.1038/nchem.2846 (2017).
35 Leaver-Fay, A. *et al.* ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods in enzymology* **487**, 545-574, doi:10.1016/B978-0-12-381270-4.00019-6 (2011).
36 Koga, N. *et al.* Principles for designing ideal protein structures. *Nature* **491**, 222-227, doi:10.1038/nature11600 (2012).
37 Scott, A. J. *et al.* Constructing ion channels from water-soluble alpha-helical barrels. *Nature chemistry* **13**, 643-650, doi:10.1038/s41557-021-00688-0 (2021).
38 Xu, C. *et al.* Computational design of transmembrane pores. *Nature* **585**, 129-134, doi:10.1038/s41586-020-2646-5 (2020).
39 Joh, N. H. *et al.* De novo design of a transmembrane Zn2+-transporting four-helix bundle. *Science* **346**, 1520-1524, doi:10.1126/science.1261172 (2014).
40 Lu, P. *et al.* Accurate computational design of multipass transmembrane proteins. *Science* **359**, 1042-1046, doi:10.1126/science.aaq1739 (2018).

41 Thomaston, J. L. *et al.* X-ray Crystal Structure of the Influenza A M2 Proton Channel S31N Mutant in Two Conformational States: An Open and Shut Case. *J. Am. Chem. Soc.* **141**, 11481-11488, doi:10.1021/jacs.9b02196 (2019).
42 Saotome, K. *et al.* Structures of the otopetrin proton channels Otop1 and Otop3. *Nature structural & molecular biology* **26**, 518-525, doi:10.1038/s41594-019-0235-9 (2019).
43 Mravic, M. *et al.* Packing of apolar side chains enables accurate design of highly stable membrane proteins. *Science* **363**, 1418-1423, doi:10.1126/science.aav7541 (2019).
44 Klesse, G., Rao, S., Sansom, M. S. P. & Tucker, S. J. CHAP: A Versatile Tool for the Structural and Functional Annotation of Ion Channel Pores. *J Mol Biol* **431**, 3353-3365, doi:10.1016/j.jmb.2019.06.003 (2019).
45 Lee, S., Liang, R., Voth, G. A. & Swanson, J. M. Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. *Journal of chemical theory and computation* **12**, 879-891, doi:10.1021/acs.jctc.5b01109 (2016).
46 Knight, C., Lindberg, G. E. & Voth, G. A. Multiscale reactive molecular dynamics. *J Chem Phys* **137**, 22A525, doi:10.1063/1.4743958 (2012).
47 Yamashita, T., Peng, Y., Knight, C. & Voth, G. A. Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. *Journal of chemical theory and computation* **8**, 4863-4875, doi:10.1021/ct3006437 (2012).
48 Moffat, J. C. *et al.* Proton transport through influenza A virus M2 protein reconstituted in vesicles. *Biophysical journal* **94**, 434-445, doi:10.1529/biophysj.107.109082 (2008).
49 Ma, C. *et al.* Identification of the functional core of the influenza A virus A/M2 proton-selective ion channel. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 12283-12288, doi:10.1073/pnas.0905726106 (2009).
50 Leiding, T., Wang, J., Martinsson, J., DeGrado, W. F. & Arskold, S. P. Proton and cation transport activity of the M2 proton channel from influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15409-15414, doi:10.1073/pnas.1009997107 (2010).
51 Slope, L. N. & Peacock, A. F. De Novo Design of Xeno-Metallo Coiled Coils. *Chemistry, an Asian journal* **11**, 660-666, doi:10.1002/asia.201501173 (2016).
52 Pinter, T. B. J., Koebke, K. J. & Pecoraro, V. L. Catalysis and Electron Transfer in De Novo Designed Helical Scaffolds. *Angewandte Chemie* **59**, 7678-7699, doi:10.1002/anie.201907502 (2020).
53 Khurana, E. *et al.* Molecular dynamics calculations suggest a conduction mechanism for the M2 proton channel from influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1069-1074, doi:10.1073/pnas.0811720106 (2009).
54 Yi, M., Cross, T. A. & Zhou, H. X. A secondary gate as a mechanism for inhibition of the M2 proton channel by amantadine. *The journal of physical chemistry. B* **112**, 7977-7979, doi:10.1021/jp800171m (2008).
55 Ramsey, I. S. *et al.* An aqueous H+ permeation pathway in the voltage-gated proton channel Hv1. *Nature structural & molecular biology* **17**, 869-875, doi:10.1038/nsmb.1826 (2010).
56 Chamberlin, A. *et al.* Hydrophobic plug functions as a gate in voltage-gated proton channels. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E273-282, doi:10.1073/pnas.1318018111 (2014).
57 Takeshita, K. *et al.* X-ray crystal structure of voltage-gated proton channel. *Nature structural & molecular biology* **21**, 352-357, doi:10.1038/nsmb.2783 (2014).
58 Wikstrom, M., Krab, K. & Sharma, V. Oxygen Activation and Energy Conservation by Cytochrome c Oxidase. *Chemical reviews* **118**, 2469-2490, doi:10.1021/acs.chemrev.7b00664 (2018).

59 Hofacker, I. & Schulten, K. Oxygen and proton pathways in cytochrome c oxidase. *Proteins: Structure, Function, and Genetics* **30**, 100-107, doi:10.1002/(sici)1097-0134(199801)30:1<100::aid-prot9>3.0.co;2-s (1998).
60 Wikström, M., Verkhovsky, M. I. & Hummer, G. Water-gated mechanism of proton translocation by cytochrome c oxidase. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1604**, 61-65, doi:10.1016/s0005-2728(03)00041-0 (2003).
61 Tashiro, M. & Stuchebrukhov, A. A. Thermodynamic properties of internal water molecules in the hydrophobic cavity around the catalytic center of cytochrome c oxidase. *The journal of physical chemistry. B* **109**, 1015-1022, doi:10.1021/jp0462456 (2005).
62 Goyal, P., Lu, J., Yang, S., Gunner, M. R. & Cui, Q. Changing hydration level in an internal cavity modulates the proton affinity of a key glutamate in cytochrome c oxidase. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 18886-18891, doi:10.1073/pnas.1313908110 (2013).
63 Liang, R., Swanson, J. M. J., Wikstrom, M. & Voth, G. A. Understanding the essential proton-pumping kinetic gates and decoupling mutations in cytochrome c oxidase. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 5924-5929, doi:10.1073/pnas.1703654114 (2017).
64 Liang, R., Swanson, J. M., Peng, Y., Wikstrom, M. & Voth, G. A. Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 7420-7425, doi:10.1073/pnas.1601982113 (2016).
65 Lynch, C. I., Rao, S. & Sansom, M. S. P. Water in Nanopores and Biological Channels: A Molecular Simulation Perspective. *Chemical reviews*, doi:10.1021/acs.chemrev.9b00830 (2020).
66 Chen, H. *et al.* Charge delocalization in proton channels, I: the aquaporin channels and proton blockage. *Biophysical journal* **92**, 46-60, doi:10.1529/biophysj.106.091934 (2007).
67 Murata, K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599-605, doi:10.1038/35036519 (2000).
68 Mondal, D., Kolev, V. & Warshel, A. Combinatorial Approach for Exploring Conformational Space and Activation Barriers in Computer-Aided Enzyme Design. *ACS Catal* **10**, 6002-6012, doi:10.1021/acscatal.0c01206 (2020).
69 Tunuguntla, R. H., Allen, F. I., Kim, K., Belliveau, A. & Noy, A. Ultrafast proton transport in sub-1-nm diameter carbon nanotube porins. *Nature nanotechnology* **11**, 639-644, doi:10.1038/nnano.2016.43 (2016).
70 Geng, J. *et al.* Stochastic transport through carbon nanotubes in lipid bilayers and live cell membranes. *Nature* **514**, 612-615, doi:10.1038/nature13817 (2014).
71 Jiang, T. *et al.* Single-chain heteropolymers transport protons selectively and rapidly. *Nature* **577**, 216-220, doi:10.1038/s41586-019-1881-0 (2020).
72 Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nature protocols* **4**, 706-731, doi:10.1038/nprot.2009.31 (2009).
73 Caffrey, M. Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annual review of biophysics* **38**, 29-51, doi:10.1146/annurev.biophys.050708.133655 (2009).
74 Kabsch, W. XDS. *Acta crystallographica. Section D, Biological crystallography* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).
75 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
76 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of applied crystallography* **40**, 658-674, doi:10.1107/S0021889807021206 (2007).

77 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
78 Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **68**, 352-367, doi:10.1107/S0907444912001308 (2012).
79 Böckmann, A. *et al.* Characterization of different water pools in solid-state NMR protein samples. *J. Biomol. NMR* **45**, 319-327 (2009).
80 Luo, W. & Hong, M. Conformational changes of an ion channel detected through water-protein interactions using solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **132**, 2378-2384, doi:10.1021/ja9096219 (2010).
81 Williams, J. K. & Hong, M. Probing membrane protein structure using water polarization transfer solid-state NMR. *Journal of magnetic resonance* **247**, 118-127, doi:10.1016/j.jmr.2014.08.007 (2014).
82 Mandala, V. S. *et al.* Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nature structural & molecular biology* **27**, 1202-1208, doi:10.1038/s41594-020-00536-8 (2020).
83 Gelenter, M. D. *et al.* Water orientation and dynamics in the closed and open influenza B virus M2 proton channels. *Communications biology* **4**, 338, doi:10.1038/s42003-021-01847-2 (2021).
84 Hong, M. *et al.* Coupling amplification in 2D MAS NMR and its application to torsion angle determination in peptides. *Journal of magnetic resonance* **129**, 85-92, doi:10.1006/jmre.1997.1242 (1997).
85 Krivov, G. G., Shapovalov, M. V. & Dunbrack, R. L., Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **77**, 778-795, doi:10.1002/prot.22488 (2009).
86 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**, D370-376, doi:10.1093/nar/gkr703 (2012).
87 Humphrey, W., Dalke, A. & Schulten, K. VMD-Visual Molecular Dynamics. *J. Molec. Graphics* **14**, 33-38 (1996).
88 Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *Journal of computational chemistry* **26**, 1701-1718, doi:10.1002/jcc.20291 (2005).
89 Huang, J. & MacKerell, A. D., Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of computational chemistry* **34**, 2135-2145, doi:10.1002/jcc.23354 (2013).
90 Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry* **29**, 1859-1865, doi:10.1002/jcc.20945 (2008).
91 Jo, S., Kim, T. & Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PloS one* **2**, e880, doi:10.1371/journal.pone.0000880 (2007).
92 Wu, E. L. *et al.* CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *Journal of computational chemistry* **35**, 1997-2004, doi:10.1002/jcc.23702 (2014).
93 Lee, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of chemical theory and computation* **12**, 405-413, doi:10.1021/acs.jctc.5b00935 (2016).
94 Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2)

dihedral angles. *Journal of chemical theory and computation* **8**, 3257-3273, doi:10.1021/ct300400x (2012).
95 Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25, doi:10.1016/j.softx.2015.06.001 (2015).
96 Nelson, J. G., Peng, Y., Silverstein, D. W. & Swanson, J. M. Multiscale Reactive Molecular Dynamics for Absolute pKa Predictions and Amino Acid Deprotonation. *Journal of chemical theory and computation* **10**, 2729-2737, doi:10.1021/ct500250f (2014).
97 Biswas, R., Tse, Y. L., Tokmakoff, A. & Voth, G. A. Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. *The journal of physical chemistry. B* **120**, 1793-1804, doi:10.1021/acs.jpcb.5b09466 (2016).
98 Day, T. J. F., Soudackov, A. V., Čuma, M., Schmitt, U. W. & Voth, G. A. A second generation multistate empirical valence bond model for proton transport in aqueous systems. *The Journal of Chemical Physics* **117**, 5839-5849, doi:10.1063/1.1497157 (2002).
99 Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics* **117**, 1-19, doi:10.1006/jcph.1995.1039 (1995).
100 Bonomi, M. B., G.; Camilloni, C.; et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nature methods* **16**, 670-673, doi:10.1038/s41592-019-0506-8 (2019).
101 Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **185**, 604-613, doi:10.1016/j.cpc.2013.09.018 (2014).
102 Grossfield, A. "WHAM: the weighted histogram analysis method", version 2.0.9, http://membrane.urmc.rochester.edu/wordpress/?page_id=126.
103 Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput Sci Eng* **9**, 90-95, doi: 10.1109/Mcse.2007.55 (2007).
104 Hodel, A., Kim, S.-H., Brünger, A.T. *Acta Cryst.* A**48**, 851-858. doi: 10.1107/S0108767392006044 (1992).