HHS Public Access

Author manuscript

*Nat Chem*. Author manuscript; available in PMC 2024 January 01.

Published in final edited form as:

*Nat Chem*. 2023 July ; 15(7): 1012–1021. doi:10.1038/s41557-023-01210-4.

# Transient Water Wires Mediate Selective Proton Transport in Designed Channel Proteins

**Huong T. Kratochvil**[1,a,*], **Laura C. Watkins**[2,b], **Marco Mravic**[1,c], **Jessica L. Thomaston**[1], **John M. Nicoludis**[1,d], **Noah H. Somberg**[3], **Lijun Liu**[4], **Mei Hong**[3], **Gregory A. Voth**[2,*], **William F. DeGrado**[1,*]

[1]Department of Pharmaceutical Chemistry, University of California-San Francisco, San Francisco, CA 94158

[2]Department of Chemistry, Chicago Center for Theoretical Chemistry, Institute for Biophysical Dynamics, and James Franck Institute, The University of Chicago, Chicago, IL 60637

[3]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

[4]DLX Scientific, Lawrence, KS, USA

## Abstract

Selective proton transport through proteins is essential for forming and utilizing proton gradients in cells. Protons are conducted along hydrogen-bonded “wires” of water molecules and polar sidechains, which, somewhat surprisingly, are often interrupted by dry apolar stretches in the conduction pathways inferred from static protein structures. We hypothesize that protons are conducted through such dry spots by forming transient water wires, often highly correlated with the presence of the excess proton itself in the water wire. To test this hypothesis, we used 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 mechanisms of biological proton conduction and principles for engineering proton-conductive materials.

## Graphical Abstract

*Authors to whom correspondences should be addressed: huong.kratochvil@unc.edu, gavoth@uchicago.edu, william.degrado@ucsf.edu.

[a]Present address: Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

[b]Present address: Kemper Insurance, Chicago, IL, USA

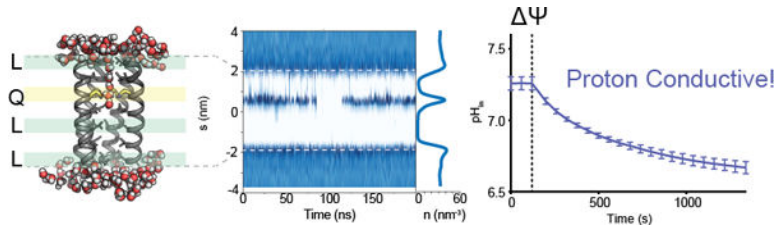
[c]Present address: Department of Integrative Structural Scripps Research Institute, La Jolla, CA. 92307

[d]Present address: Genentech, San Francisco, CA, USA

**Author contributions**: H.T.K. designed the channels, ran flux measurements, crystallized and collected the X-ray diffraction data, and ran and analyzed classical MD simulations. L.C.W. ran the MS-RMD simulations, L.C.W. and G.A.V. analyzed 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. analyzed experimental data. All authors contributed to data analysis and writing the manuscript.

**Competing interests**: The authors declare no competing interests.

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




## Introduction:

The controlled diffusion of protons through transmembrane proteins is critical for many aspects of physiological function, including substrate transport [1], control of cellular and organelle pH [2], the creation and utilization of pH gradients required for bioenergetics [3,4], and cellular signaling [5]. Proteins conduct protons along precisely defined pathways that prevent 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 significantly 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 waters. 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 “down-stream” 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 sidechains, which 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 cryo-electron 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 the 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 transmembrane potential, 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 where transient spectroscopy and serial crystallography have identified water molecules that form during portions of its photo-cycle [13,21–23]. However, this system represents only a single example in which large changes are induced by photo-isomerization.

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 where 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 also favor tight and efficient packing are likely to dampen 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. Here, we not only consider water explicitly, but we also account for the dynamic formation and breaking of covalent bonds as protons are passed from one water to the next. Finally, despite a few successes [27,37–40], the design and high-resolution structural characterization of membrane proteins remains a difficult endeavor. 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 anti-porting 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 significant 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 M2 and OTOP proton channels[41,42] Indeed, Gln is also flexible so that it should be able to stabilize multiple polarizations of water wires that are created during a conduction cycle.

Our constructs began with 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 and Ile residues (Fig. 2a, pdb 6mct) [43]. A narrow, fully hydrophobic pore (app. 2 – 3 Å diameter), which 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 pore without steric clashes. The variants are designated as LLLL, QLLL, QQLL, etc., according to the identity of the Leu-to-Gln substitutions (Fig. 2). Seven variants with 0, 1, or 2 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 it was held constant at approximately 18 Å to 20 Å for the remaining five variants (Fig. 2c). Both single and double proton-loading sites were evaluated by varying the positions of the introduced Gln residues. A peptide with 3 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, Supplementary Figs. S1–S6). The pore of the starting LLLL (Fig. 3a) pentamer was devoid of significant water density throughout the trajectories. The length of the longest apolar path in its pore ($l_{obs}$ =33.9 ± 0.3 Å Fig. 3a, Extended Table 1) was in good agreement with design ($l_{exp}$ of 33 Å). QLLL (Fig. 3b) showed strong water density near its Gln sidechain near the entry 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) has strong hydration near the Gln carboxamide sidechain, which communicates with the bulk water at the top of the channel via infrequent, flickering water-wires. The remaining C-terminal pore region is fully dehydrated over approximately length 21 Å ($l_{obs}$ = 20.7 ± 0.1 Å). LLQL (Fig. 3d) shows inverse behavior with hydrated Gln residues, fluctuating water wires near the bottom of the channel, and a 22.4 ± 0.04 Å dehydrated pore at the top. The doubly-substituted pentamers had 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 having dry regions of similar length to the longest stretch seen in the single-Gln variants ($l_{obs}$ = 20.9 ± 0.04 Å 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 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 backbone a RMSD to the starting LLLL pentamer ranging from 0.25 to 0.32 Å (Fig. 4a). The Gln sidechains converge in layers near the center of the bundle, where they form sidechain-sidechain and sidechain-backbone hydrogen bonds (Fig. 4 and Supplementary Figs. S7–S16). As expected, they also surround puncta of electron density, which were well modeled as water molecules at full occupancy (Extended Data Fig. 1). The Gln sidechains are not fully symmetric and different conformations are seen in the individual monomeric units of the pentamers. They also have higher Debye Waller factors than the surrounding main-chain and pore-lining sidechains. These findings are in good agreement with the hydration of Gln sidechains observed in MD simulations.

To confirm the hydration observed from MD simulations and crystal structures, we examined water-edited $^{13}C$ solid state NMR (ssNMR) spectra of a series of LLLL and LQLL peptides uniformly labeled with $^{13}C$, $^{15}N$ 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, Extended Data Fig. 2). Water magnetization transfer is observed for both LLLL and LQLL channels because magnetization is also transferred from the membrane-surface water to 13C-labeled residues via relayed 1H 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 labeled Gln10 also significant 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 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 to water. To simulate this process, we turned to RMD simulations.

## Multiscale reactive MD 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 with the Grotthuss shuttling mechanism caused by protons permeating the channel [8,10]. We therefore turned to multiscale reactive MD (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 from between water molecules [45–47]. While MS-RMD is significantly faster than explicit quantum mechanical (*ab initio* MD) calculations, the calculation of proton transport through an entire TM protein via MS-RMD is still computationally very intensive so we confined our focus to a comparison of LLLL and LQLL from the crystal structures. We used two collective variables (CVs) to enable enhanced free energy sampling in the RMD simulations: the position of the “center of excess charge” (CEC), which tracks the translocation of a proton charge defect along the pore [45–47], and a water wire connectivity parameter, $\phi$, which quantifies the number and connectivity of hydrogen-bonded waters in the pore and also associated with the excess proton structure [8]. A value of $\phi$ of 0 corresponds to a dry pore, and 1 to a pore with a fully connected water wire spanning the pore and containing an excess proton.

The computed potentials of mean force (PMFs) for LLLL vs. 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 (Figs. 5a,b and Supplementary Fig. S17). The landscape and its saddle region for LQLL (rescaled in 5c to allow easier viewing of its contours) describes 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 N-terminus (right side in Fig. 5a,b, top in Fig. 5d) and a "long apolar path" near the C-terminus (left in Fig. 5a,b, 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 and 5d) features water-wires that are fully formed through only this region of the channel and then "flip" to translocate the proton on through. 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 and 5d, and Supplementary Fig. S18). The transmission of protons through this region was thus computed to proceed via two energetically similar pathways (Fig. 5c), which differ in whether water wires are present or absent in the shorter path as the proton moves through the longer path (i.e., 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 point 4A to 4B in Fig. 5c) indicates an large positive entropy change for the excess proton moving into the saddle region. This is a highly favorable entropy of activation and so one must therefore not interpret the "barrier" for the proton transport as simply coming 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 coming from an average over many such paths crossing along that vertical saddle, resulting in a quite favorable proton translocation behavior as seen in the experiments. In summary, RMD of LLLL versus LQLL predicts that the introduction of a Gln-rich PLS one third of the way through the pore should dramatically enhance 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 used a vesicle assay (adapted from [48,49]) that uses a chemiosmotically-induced electrical potential to drive 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 the 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 the $Na^+$ will diffuse into the vesicle dissipating the electrical potential (Extended Data Fig. 3). Figure 5c 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 addition of valinomycin (t = 120 s) leads to a change in $pH_{in}$ as expected from M2's proton-selectivity, but no change in $pH_{in}$ for empty vesicles which contain no protein (Fig. 6b). Following the addition of valinomycin,

the protonophore, carbonyl cyanide m-chlorophenyl hydrazone (CCCP), is added (Extended Data Fig. 3); the observed additional bolus of proton flux assures that an electrical and pH gradients have been maintained throughout the experiment, even on the order of hours (Extended Data Figs. 4–7, Supplementary Figs. S19–S23).

Using this assay, we determined that LLLL ($I_{obs}$ = 33.9 ± 0.3 Å) had no significant conductance over background, and QLLL ($I_{obs}$ = 32.0 ± 0.6 Å) had a very low conduction that was not significantly different from 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 par with M2 at equivalent symmetrical $pH_{in}$ and $pH_{out}$, and peptide/lipid ratios (Fig. 6g). The conductance rate of these four functional proton channels were all highly similar (within a factor of two), which is consistent with their similar hydrophobic lengths. This finding was confirmed in 6–9 independent replicates for each channel ($p < 0.0001$) (Table S3). 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.

## Discussion

We have successfully engineered minimalist proton channels that enable us to explore the roles transient water wires through apolar regions play in proton selectivity and conduction. In keeping with our minimalist strategy, a ring of Gln sidechains 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 concerning 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 show here can be as long as ~20 Å. While this was predicted based on MS-RMD, our classical MD simulations, which do not explicitly consider a mobile Grotthuss shuttling excess proton, predicted that water wires could only span on the order 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 for identifying water wires that are induced as a proton enters the channel (Fig. 1d). The M2 proton channel from influenza A virus [53,54] and Hv1 [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 release of protons that induce formation of water wires. Similarly, proton transport to the proton loading site of cytochrome c oxidase traverses an apolar region [58–62]; classical and MS-RMD simulations have 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 other larger ions have apolar pores that are dry in the “off” state but become significantly more enlarged and hydrated in response to a larger conformational gating transition [16–19,65].

During proton conduction, each water in a water wire changes the direction of its dipole as protons transfer from one water molecule to the next. The orientation of the waters must be reset to regenerate the initial polarization before a new proton can be transported. Water wires in restricted hydrophobic environments are well suited for stabilizing both polarizations because they do not form strong dipolar interactions with the pore that would bias their orientations. Also, the Gln sidechains are relatively mobile in our crystal structures as assessed from the B-factors, MD and MS-RMD (Supplementary Fig. S24), which shows rapid conformational fluctuations. This behavior contrasts with the requirements for water channels, like aquaporins, which feature stable water-binding sites with strong polarization that undermine 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 the static structures to include the dynamic processes required for proton migration, including both protein and water dynamics, which *de novo* design generally ignores in favor of structural stability and computational efficiency. MS-RMD allowed even deeper consideration of the bond-making/breaking steps required for Grotthuss proton migration through water wires. While computational speed is currently too slow to incorporate RMD into early stages of protein design, it can clearly provide an important filter to assess potential designs. Indeed, Mondal and coworkers 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 showed that close positioning of a PLS proximal to a dry pore resulted in channels that are highly selective for protons. Our minimalist designs show that the PLS need not be elaborate in design. Indeed, the present 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 design of soft materials proton-selective membranes. For example, Jiang and coworkers have designed proton-selective copolymers consisting of apolar segments interspersed with occasional polar ethylene glycol units [71], bearing 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 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 utilization.

## Materials and Methods

### Synthesis of Proton Channel Peptides

Proton channel peptides, shown in Fig. 2 and Supplementary Table S1, 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 S1). 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 × 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 95°C. 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 × 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 42°C 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 S2.

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 20°C. Crystals of each construct were harvested from conditions noted in Supplementary Table S2 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 [5] 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 [8] 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 $T\ (K) = 96.9 \times (7.83\ ppm - \delta_{H2O})$

[79]. Typical radiofrequency field strengths were 50–80 kHz for 1H 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 × 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 × 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]:

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

Where $\vec{r}^{\,i}_{COC}$ 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.

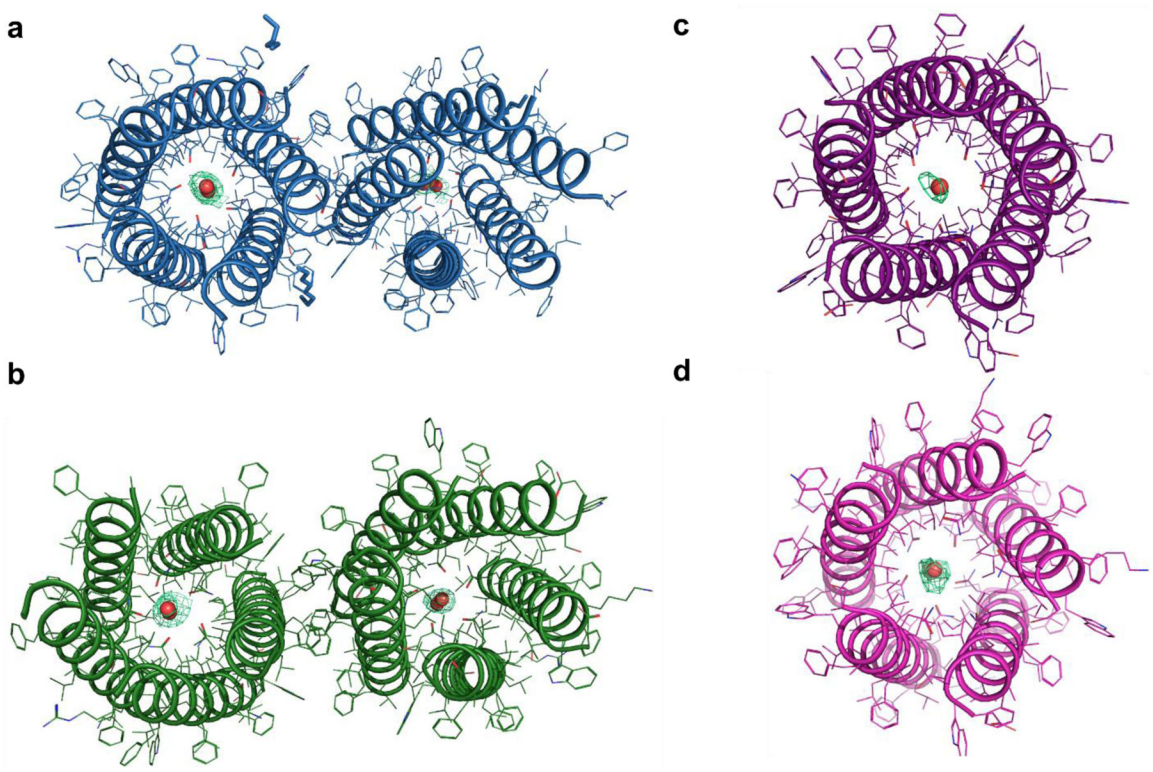
The open-source, community developed PLUMED library [100,101] was used to define the umbrella bias potentials. Umbrella sampling windows were set up every 0.5Å in the range [-22, 22] Å for Z'CEC and every 0.035 in the range [0.140,0.980] and [0.245,0.980] for φ for LLLL and LQLL, respectively, for a 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.

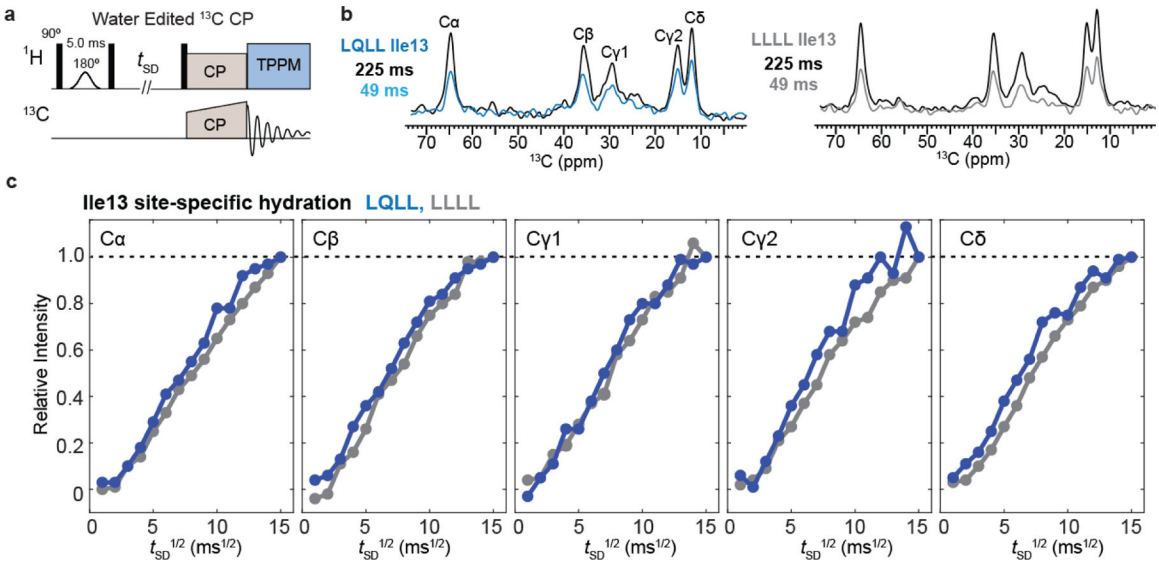
Figs. 5, Supplemental Figs. S17, S18, and S24 were generated with Matplotlib [103].

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

## Extended Data

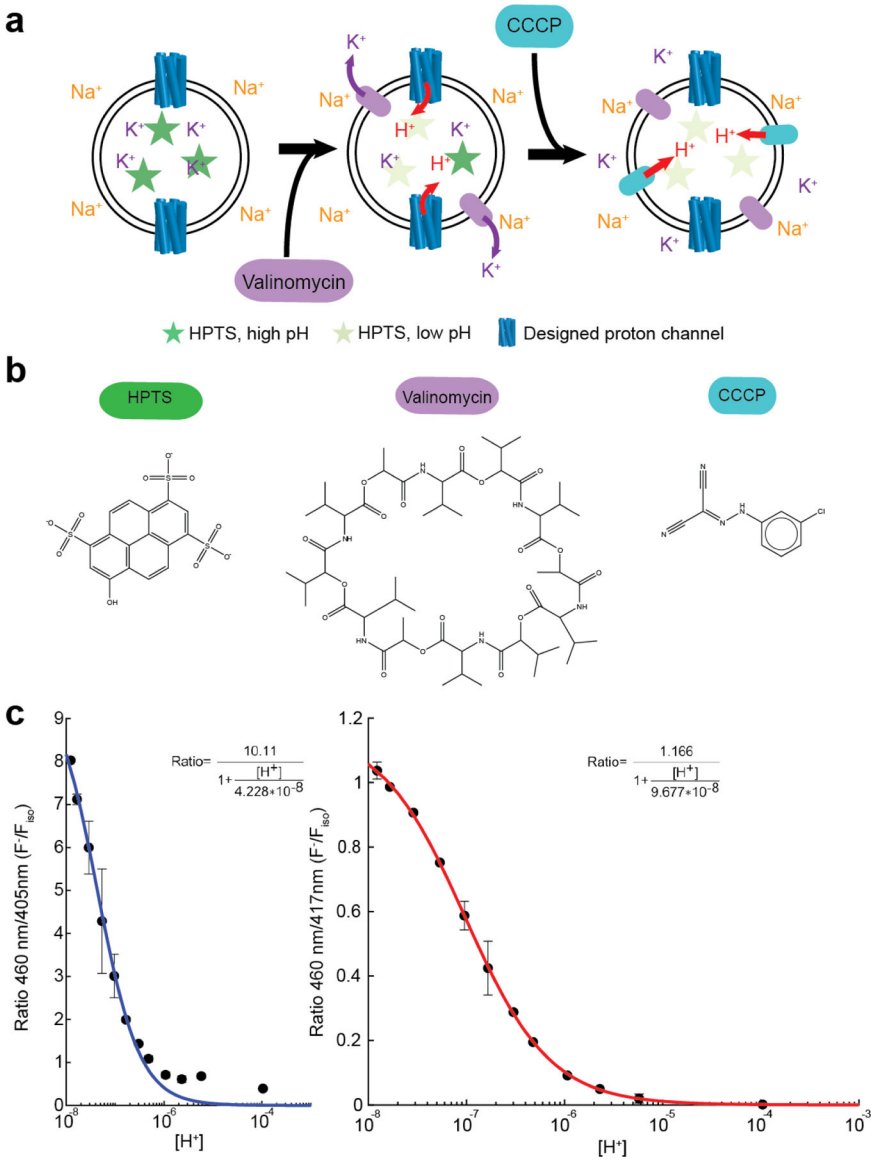



**Extended Data Fig. 1. Composite omit maps (2mFo-DFc) of designed proton channels.**
Composite omit maps of the asymmetric unit for **a,** LQLL, **b**, LLQL, and one pentamer from the asymmetric unit for **c,** QQLL, and **d,** QLQL (shown only for the waters for clarity). All contours at σ = 1.0. Omit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



**Extended Data Fig. 2. Pulse diagram and water-edited 13C spectra, water buildup curves of membrane-bound LQLL and LLLL peptides.**

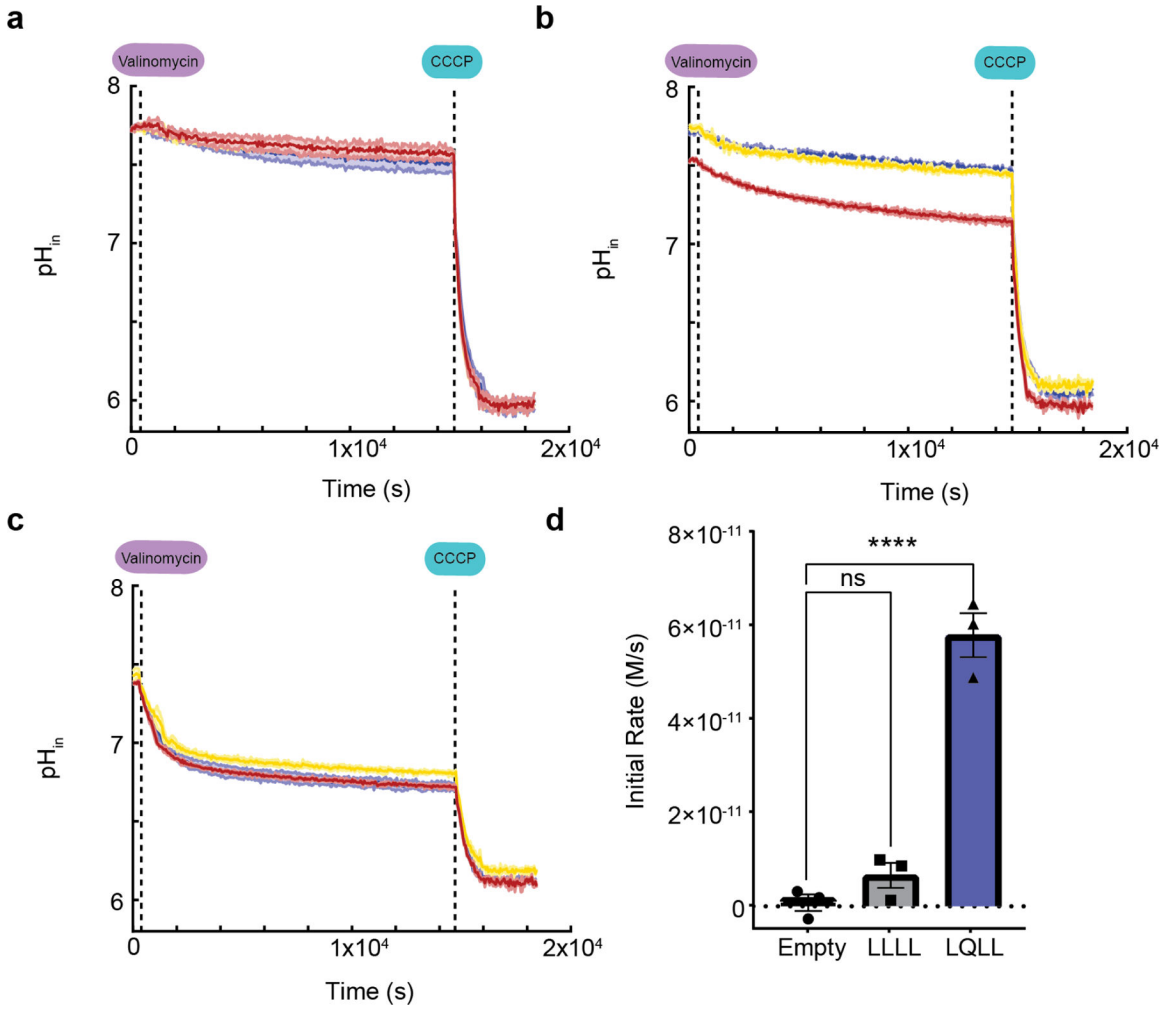
**a**, Pulse diagram of water-edited 13C CP experiment. **b**, Representative water-edited 13C spectra of I13 in LQLL and LLLL, measured with 225 ms and 49 ms $^{1}$H mixing. The relative intensities of the 49 ms spectrum to the 225 ms spectrum are higher for LQLL Ile13 than LLLL Ile13, especially for 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.




**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

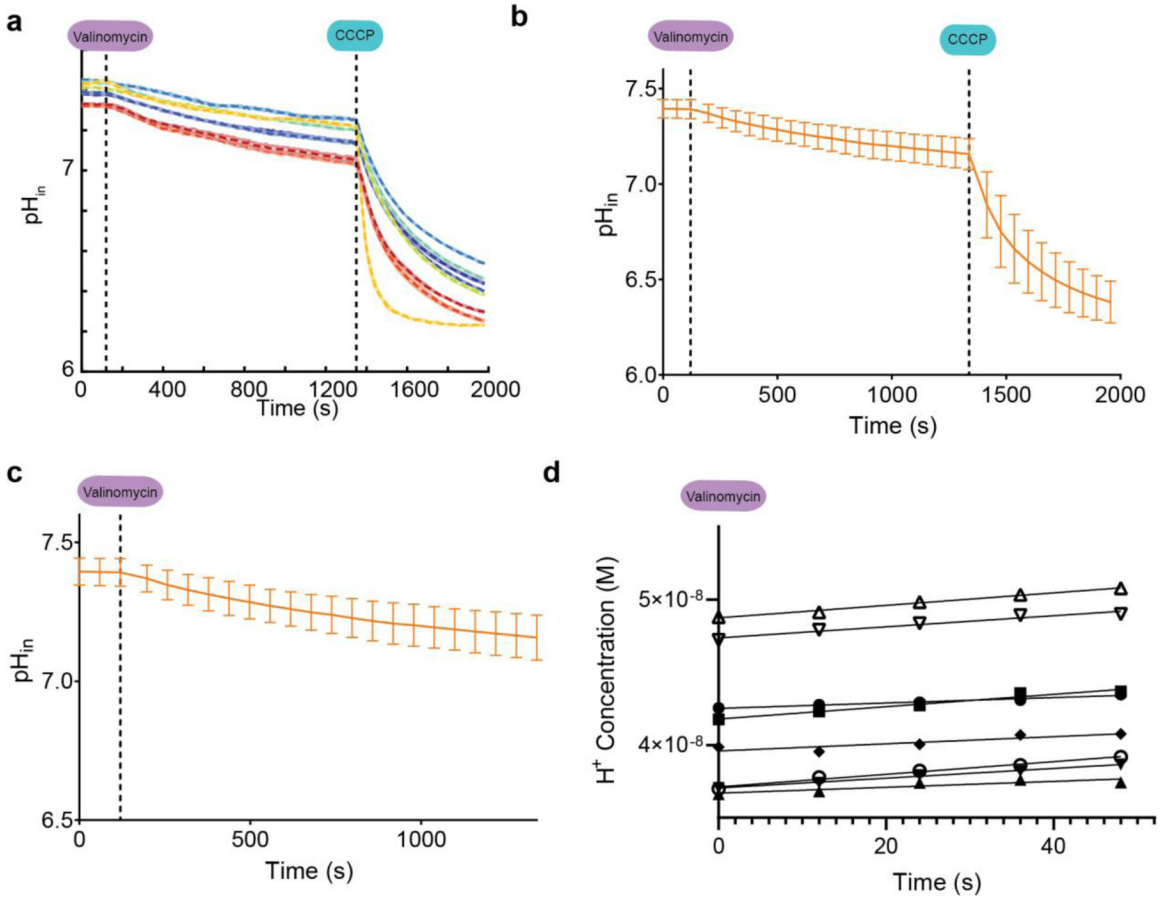
of 50 mM $K_2SO_4$, 30 mM K2HPO4 at different pH values for two plate readers used in data collection process. Unless stated, all data collected with instrument that generated the blue calibration curve. Fits used for downstream data processing shown for each of the two instruments with adjusted R-squared 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.




**Extended Data Fig. 4. All proton flux assay data for long kinetics runs.**

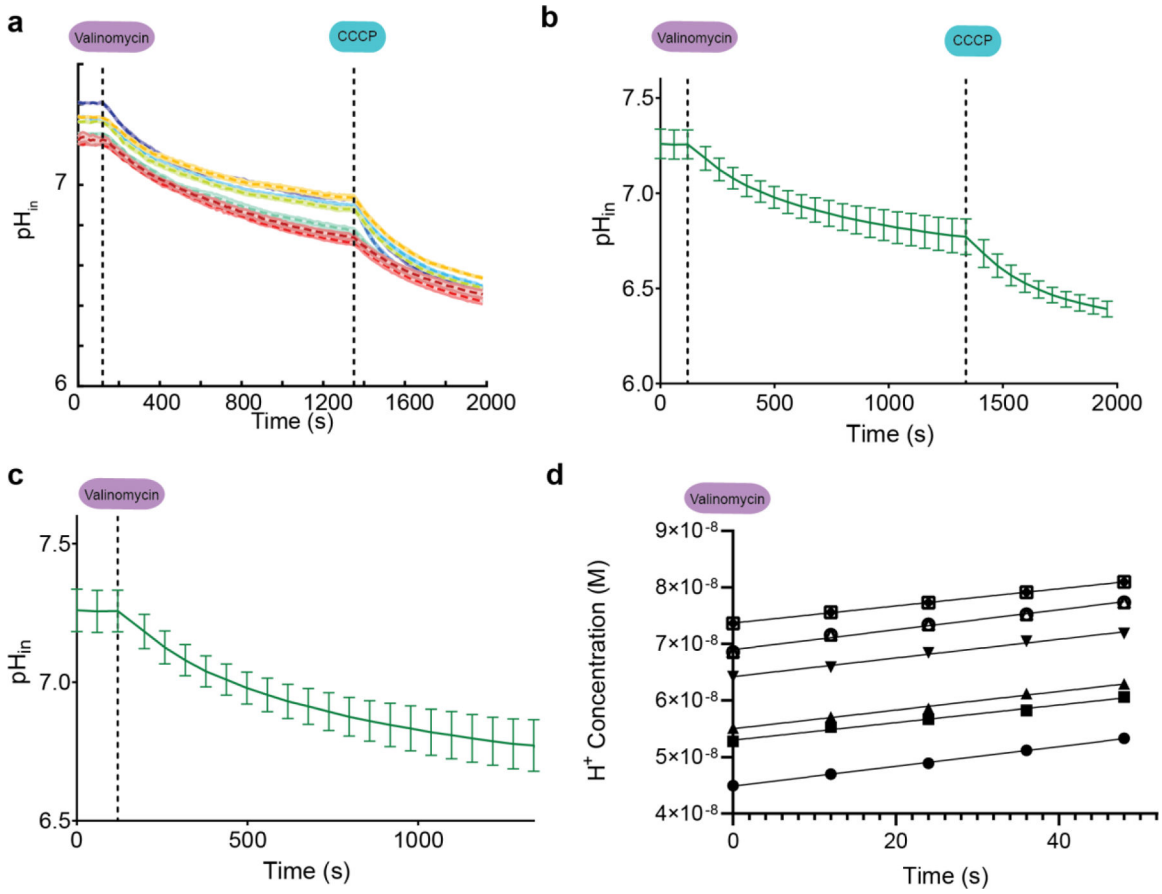
Long kinetics runs of about 5 hours total for **a,** empty, **b,** LLLL, and **c,** LQLL vesicles. Dotted lines denote times in experiment when valinomycin and CCCP were added to the samples. Three samples of the different conditions were measured in triplicate. These long-time measurements reveal that the vesicles are not significantly 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.



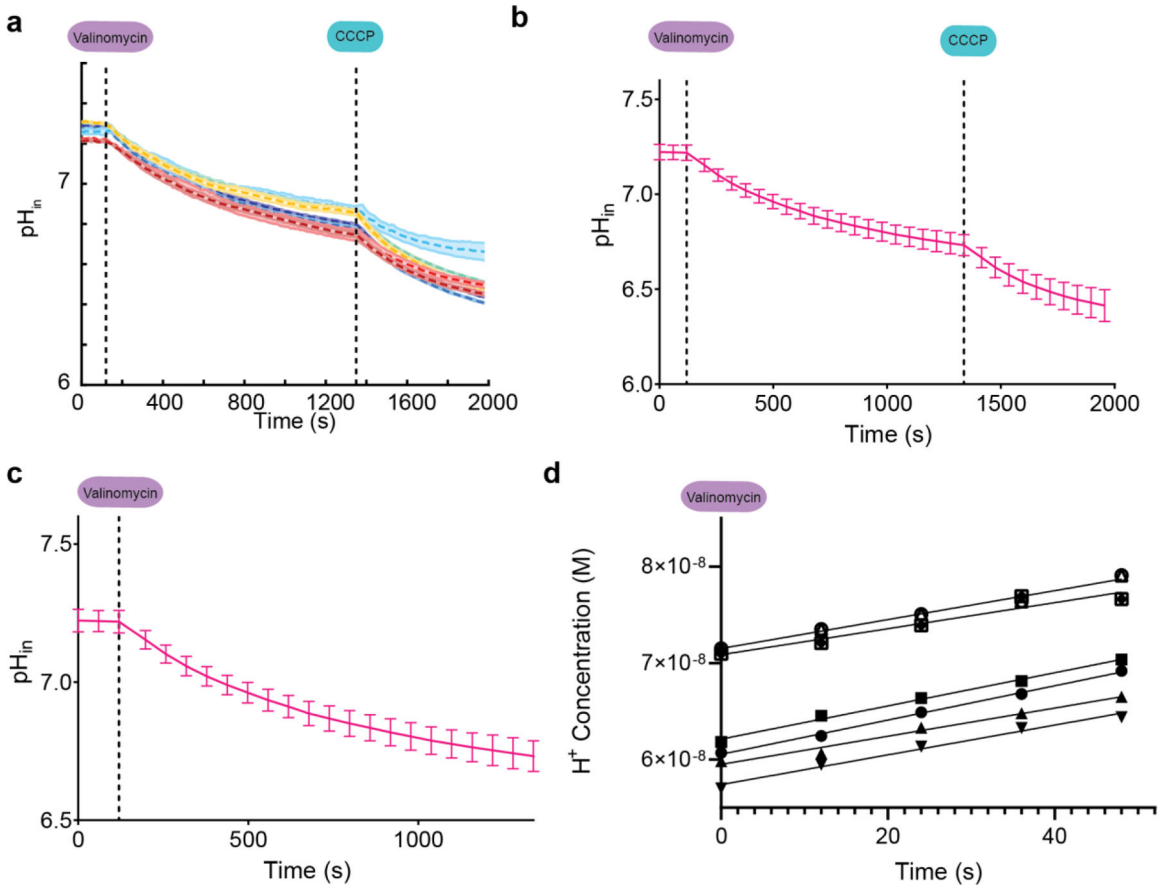

**Extended Data Fig. 5. All proton flux assay data for QLLL vesicle samples.**

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 S3.



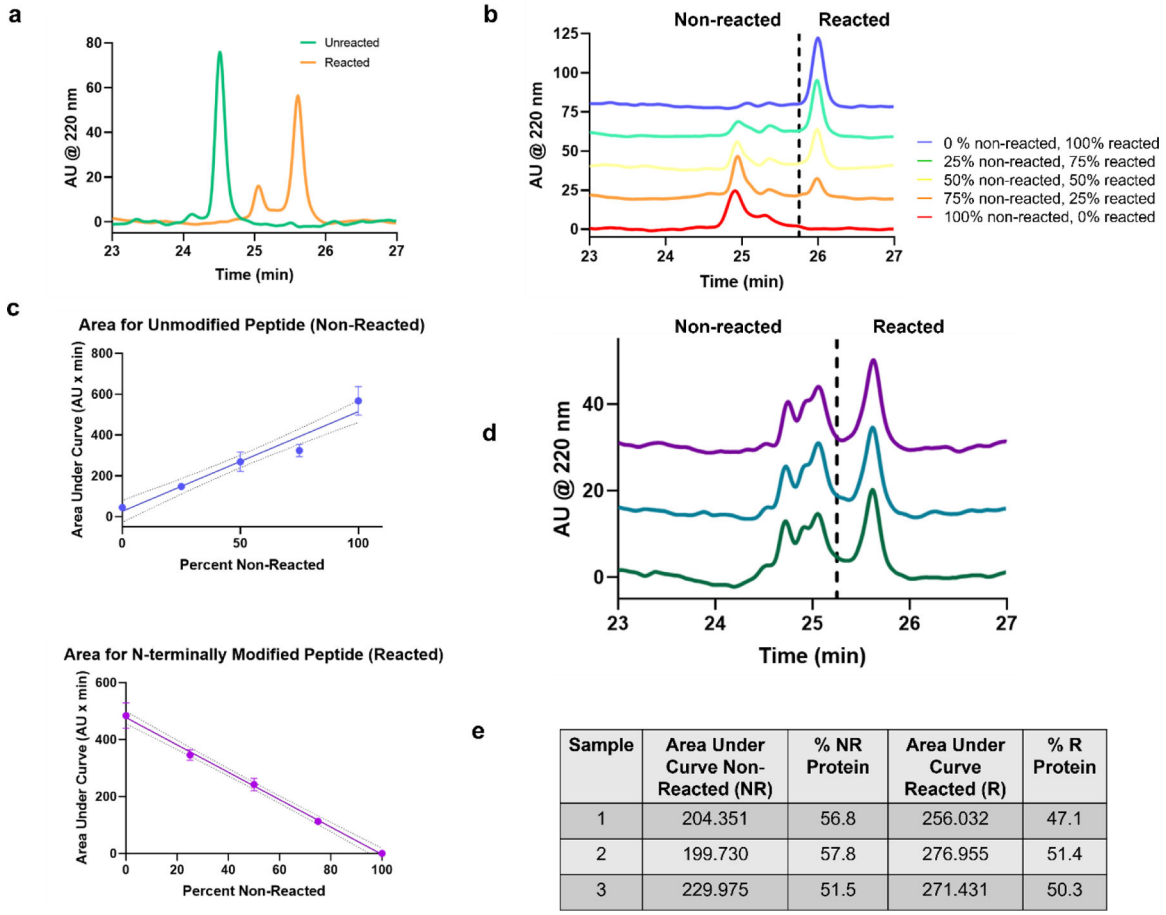

**Extended Data Fig. 6. All proton flux assay data for LLQL vesicle samples.**

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 significant 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 S3.




**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 significant 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 S3.




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 Materials and 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.

## Extended Data Table 1.

**Length ($l_{obs}$) of longer and shorter ($l_{obs,sh}$) hydrophobic stretches in designed channels.**

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 Materials and Methods).

| 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 |

## Extended Data Table 2.

**Data collection and refinement statistics for all structures.**

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.

| 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 | $C222_1$ | C2 | $P2_12_12_1$ | $P2_1$ | $C222_1$ |
| **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 |
| $\alpha, \beta, \gamma$ (°) | 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) |
| $\langle I/\sigma I \rangle$ | 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** | | | | | |

| 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 |
|---|---|---|---|---|---|
| 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 |

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements:

Diffraction data was 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, P30GM138396), and the Eiger 16M detector by NIH S10 OD012289. We also acknowledge the Advanced Photon Source, supported by U.S. 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 National Institutes of Health (NIH; R01 GM124149 and P30 GM124169), Plexxikon Inc. and the Integrated Diffraction Analysis Technologies program (US Department of Energy Office of Biological and Environmental Research. The Advanced Light Source at Lawrence Berkeley National Laboratory is supported by DE-AC02-05CH11231.

**Funding:**

H.T.K. was supported by NIH (K99GM138753). W.F.D. by 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 NIH (R01 GM053148), and J.M.N. by 5T32HL007731 and F32GM133085. J.L.T. was supported by the NIH R35GM122603.

## 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. Biochem. Biophys. Res. Comm 173, 443–448, doi:10.1016/s0006-291x(05)81078-2 (1990). [PubMed: 1979489]
2. Nishi T & Forgac M The vacuolar (H+)-ATPases--nature's most versatile proton pumps. Nat. Rev. Mol. Cell. Biol 3, 94–103, doi:10.1038/nrm729 (2002). [PubMed: 11836511]
3. Mitchell P Coupling of phosphorylation to electron and hydrogen transfer by a chemiosmotic type of mechanism. Nature 191, 144–148, doi:10.1038/191144a0 (1961). [PubMed: 13771349]
4. Nicholls DG Mitochondrial ion circuits. Essays Biochem 47, 25–35, doi:10.1042/bse0470025 (2010). [PubMed: 20533898]

5. Diering GH & Numata M Endosomal pH in neuronal signaling and synaptic transmission: role of Na(+)/H(+) exchanger NHE5. Front. Physiol 4, 412, doi:10.3389/fphys.2013.00412 (2014). [PubMed: 24454292]
6. Agmon N The Grotthuss mechanism. Chem. Phys. Lett 244, 456–462, doi:10.1016/0009-2614(95)00905-j (1995).
7. Calio PB, Li C & Voth GA Resolving the Structural Debate for the Hydrated Excess Proton in Water. J. Am. Chem. Soc 143, 18672–18683, doi:10.1021/jacs.1c08552 (2021). [PubMed: 34723507]
8. Li C & Voth GA A quantitative paradigm for water-assisted proton transport through proteins and other confined spaces. Proc. Natl. Acad. Sci. U.S.A 118, doi:10.1073/pnas.2113141118 (2021).
9. Wraight CA Chance and design--proton transfer in water, channels and bioenergetic proteins. Biochim. Biophys. Acta 1757, 886–912, doi:10.1016/j.bbabio.2006.06.017 (2006). [PubMed: 16934216]
10. Decoursey TE Voltage-gated proton channels and other proton transfer pathways. Physiol Rev 83, 475–579, doi:10.1152/physrev.00028.2002 (2003). [PubMed: 12663866]
11. Peng Y, Swanson JM, Kang SG, Zhou R & Voth GA Hydrated Excess Protons Can Create Their Own Water Wires. J. Phys. Chem. B 119, 9212–9218, doi:10.1021/jp5095118 (2015). [PubMed: 25369445]
12. Banh R et al. Hydrophobic gasket mutation produces gating pore currents in closed human voltage-gated proton channels. Proc. Natl. Acad. Sci. U.S.A 116, 18951–18961, doi:10.1073/pnas.1905462116 (2019). [PubMed: 31462498]
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). [PubMed: 16280982]
14. Kaur D, Khaniya U, Zhang Y & Gunner MR Protein Motifs for Proton Transfers That Build the Transmembrane Proton Gradient. Front. Chem 9, 660954, doi:10.3389/fchem.2021.660954 (2021). [PubMed: 34211960]
15. Kalra A, Garde S & Hummer G Osmotic water transport through carbon nanotube membranes. Proc. Natl. Acad. Sci. U.S.A 100, 10175–10180, doi:10.1073/pnas.1633354100 (2003). [PubMed: 12878724]
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, doi:10.1038/nsmb.1525 (2009). [PubMed: 19098918]
17. Jensen MO et al. Principles of conduction and hydrophobic gating in K+ channels. Proc. Natl. Acad. Sci. U.S.A 107, 5833–5838, doi:10.1073/pnas.0911691107 (2010). [PubMed: 20231479]
18. Aryal P, Sansom MS & Tucker SJ Hydrophobic gating in ion channels. J. Mol. Biol 427, 121–130, doi:10.1016/j.jmb.2014.07.030 (2015). [PubMed: 25106689]
19. Zhu F & Hummer G Drying transition in the hydrophobic gate of the GLIC channel blocks ion conduction. Biophys. J 103, 219–227, doi:10.1016/j.bpj.2012.06.003 (2012). [PubMed: 22853899]
20. Rasaiah JC, Garde S & Hummer G Water in nonpolar confinement: from nanotubes to proteins and beyond. Annu. Rev. Phys. Chem 59, 713–740, doi:10.1146/annurev.physchem.59.032607.093815 (2008). [PubMed: 18092942]
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). [PubMed: 23394942]
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). [PubMed: 31273117]
23. Freier E, Wolf S & Gerwert K Proton transfer via a transient linear water-molecule chain in a membrane protein. Proc. Natl. Acad. Sci. U.S.A 108, 11435–11439, doi:10.1073/pnas.1104735108 (2011). [PubMed: 21709261]
24. Regan L & DeGrado WF Characterization of a helical protein designed from first principles. Science 241, 976–978, doi:10.1126/science.3043666 (1988). [PubMed: 3043666]
25. Walsh ST, Cheng H, Bryson JW, Roder H & DeGrado WF Solution structure and dynamics of a de novo designed three-helix bundle protein. Proc. Natl. Acad. Sci. U.S.A 96, 5486–5491, doi:10.1073/pnas.96.10.5486 (1999). [PubMed: 10318910]

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). [PubMed: 14631033]
27. Vorobieva AA 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). [PubMed: 33398169]
29. Polizzi NF & DeGrado WF A defined structural unit enables de novo design of small-molecule-binding proteins. Science 369, 1227–1233, doi:10.1126/science.abb8330 (2020). [PubMed: 32883865]
30. Cao L et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. Science 370, 426–431, doi:10.1126/science.abd9909 (2020). [PubMed: 32907861]
31. Fleishman SJ et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science 332, 816–821, doi:10.1126/science.1202617 (2011). [PubMed: 21566186]
32. Jiang L et al. De novo computational design of retro-aldol enzymes. Science 319, 1387–1391, doi:10.1126/science.1152692 (2008). [PubMed: 18323453]
33. Lassila JK, Privett HK, Allen BD & Mayo SL Combinatorial methods for small-molecule placement in computational enzyme design. Proc. Natl. Acad. Sci. U.S.A 103, 16710–16715, doi:10.1073/pnas.0607691103 (2006). [PubMed: 17075051]
34. Polizzi NF 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. Meth. Enzymol 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). [PubMed: 23135467]
37. Scott AJ et al. Constructing ion channels from water-soluble alpha-helical barrels. Nat. Chem 13, 643–650, doi:10.1038/s41557-021-00688-0 (2021). [PubMed: 33972753]
38. Xu C et al. Computational design of transmembrane pores. Nature 585, 129–134, doi:10.1038/s41586-020-2646-5 (2020). [PubMed: 32848250]
39. Joh NH et al. De novo design of a transmembrane Zn2+-transporting four-helix bundle. Science 346, 1520–1524, doi:10.1126/science.1261172 (2014). [PubMed: 25525248]
40. Lu P et al. Accurate computational design of multipass transmembrane proteins. Science 359, 1042–1046, doi:10.1126/science.aaq1739 (2018). [PubMed: 29496880]
41. Thomaston JL 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). [PubMed: 31184871]
42. Saotome K et al. Structures of the otopetrin proton channels Otop1 and Otop3. Nat. Struct. Mol. Biol 26, 518–525, doi:10.1038/s41594-019-0235-9 (2019). [PubMed: 31160780]
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). [PubMed: 30923216]
44. Klesse G, Rao S, Sansom MSP & Tucker SJ 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). [PubMed: 31220459]
45. Lee S, Liang R, Voth GA & Swanson JM Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. J. Chem. Theory Comput 12, 879–891, doi:10.1021/acs.jctc.5b01109 (2016). [PubMed: 26734942]
46. Knight C, Lindberg GE & Voth GA Multiscale reactive molecular dynamics. J. Chem. Phys 137, 22A525, doi:10.1063/1.4743958 (2012).
47. Yamashita T, Peng Y, Knight C & Voth GA Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. J. Chem. Theory Comput 8, 4863–4875, doi:10.1021/ct3006437 (2012). [PubMed: 25100924]

48. Moffat JC et al. Proton transport through influenza A virus M2 protein reconstituted in vesicles. Biophys. J 94, 434–445, doi:10.1529/biophysj.107.109082 (2008). [PubMed: 17827230]
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. U.S.A 106, 12283–12288, doi:10.1073/pnas.0905726106 (2009). [PubMed: 19590009]
50. Leiding T, Wang J, Martinsson J, DeGrado WF & Arskold SP Proton and cation transport activity of the M2 proton channel from influenza A virus. Proc. Natl. Acad. Sci. U.S.A 107, 15409–15414, doi:10.1073/pnas.1009997107 (2010). [PubMed: 20713739]
51. Slope LN & Peacock AF De Novo Design of Xeno-Metallo Coiled Coils. Chem. Asian J 11, 660–666, doi:10.1002/asia.201501173 (2016). [PubMed: 26592205]
52. Pinter TBJ, Koebke KJ & Pecoraro VL Catalysis and Electron Transfer in De Novo Designed Helical Scaffolds. Angew. Chem., Int. Ed 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. Proc. Natl. Acad. Sci. U.S.A 106, 1069–1074, doi:10.1073/pnas.0811720106 (2009). [PubMed: 19144924]
54. Yi M, Cross TA & Zhou HX A secondary gate as a mechanism for inhibition of the M2 proton channel by amantadine. J. Phys. Chem. B 112, 7977–7979, doi:10.1021/jp800171m (2008). [PubMed: 18476738]
55. Ramsey IS et al. An aqueous H+ permeation pathway in the voltage-gated proton channel Hv1. Nat. Struct. Mol. Biol 17, 869–875, doi:10.1038/nsmb.1826 (2010). [PubMed: 20543828]
56. Chamberlin A et al. Hydrophobic plug functions as a gate in voltage-gated proton channels. Proc. Natl. Acad. Sci. U.S.A 111, E273–282, doi:10.1073/pnas.1318018111 (2014). [PubMed: 24379371]
57. Takeshita K et al. X-ray crystal structure of voltage-gated proton channel. Nat. Struct. Mol. Biol 21, 352–357, doi:10.1038/nsmb.2783 (2014). [PubMed: 24584463]
58. Wikstrom M, Krab K & Sharma V Oxygen Activation and Energy Conservation by Cytochrome c Oxidase. Chem. Rev 118, 2469–2490, doi:10.1021/acs.chemrev.7b00664 (2018). [PubMed: 29350917]
59. Hofacker I & Schulten K Oxygen and proton pathways in cytochrome c oxidase. Proteins: Struct. Funct. Genet 30, 100–107, doi:10.1002/(sici)1097-0134(199801)30:1<100::aid-prot9>3.0.co;2-s (1998). [PubMed: 9443344]
60. Wikström M, Verkhovsky MI & Hummer G Water-gated mechanism of proton translocation by cytochrome c oxidase. Biochim. Biophys. Acta Bioenerg 1604, 61–65, doi:10.1016/s0005-2728(03)00041-0 (2003).
61. Tashiro M & Stuchebrukhov AA 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, doi:10.1021/jp0462456 (2005). [PubMed: 16866474]
62. Goyal P, Lu J, Yang S, Gunner MR & 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. U.S.A 110, 18886–18891, doi:10.1073/pnas.1313908110 (2013). [PubMed: 24198332]
63. Liang R, Swanson JMJ, Wikstrom M & Voth GA Understanding the essential proton-pumping kinetic gates and decoupling mutations in cytochrome c oxidase. Proc. Natl. Acad. Sci. U.S.A 114, 5924–5929, doi:10.1073/pnas.1703654114 (2017). [PubMed: 28536198]
64. Liang R, Swanson JM, Peng Y, Wikstrom M & Voth GA Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase. Proc. Natl. Acad. Sci. U.S.A 113, 7420–7425, doi:10.1073/pnas.1601982113 (2016). [PubMed: 27339133]
65. Lynch CI, Rao S & Sansom MS P. Water in Nanopores and Biological Channels: A Molecular Simulation Perspective. Chem. Revs, 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). [PubMed: 17056733]
67. Murata K et al. Structural determinants of water permeation through aquaporin-1. Nature 407, 599–605, doi:10.1038/35036519 (2000). [PubMed: 11034202]

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). [PubMed: 34178420]

69. Tunuguntla RH, Allen FI, Kim K, Belliveau A & Noy A Ultrafast proton transport in sub-1-nm diameter carbon nanotube porins. Nat. Nanotechnol 11, 639–644, doi:10.1038/nnano.2016.43 (2016). [PubMed: 27043198]

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). [PubMed: 25355362]

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). [PubMed: 31915399]

## References for Materials and Methods

72. Caffrey M & Cherezov V Crystallizing membrane proteins using lipidic mesophases. Nat. Prot 4, 706–731, doi:10.1038/nprot.2009.31 (2009).

73. Caffrey M Crystallizing membrane proteins for structure determination: use of lipidic mesophases. Annu. Rev. Biophys 38, 29–51, doi:10.1146/annurev.biophys.050708.133655 (2009). [PubMed: 19086821]

74. Kabsch W XDS. Acta Crystallogr. D 66, 125–132, doi:10.1107/S0907444909047337 (2010). [PubMed: 20124692]

75. Winn MD et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D 67, 235–242, doi:10.1107/S0907444910045749 (2011). [PubMed: 21460441]

76. McCoy AJ et al. Phaser crystallographic software. J. Appl. Crystallogr 40, 658–674, doi:10.1107/S0021889807021206 (2007). [PubMed: 19461840]

77. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of Coot. Acta Crystallogr. D 66, 486–501, doi:10.1107/S0907444910007493 (2010). [PubMed: 20383002]

78. Afonine PV et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D 68, 352–367, doi:10.1107/S0907444912001308 (2012). [PubMed: 22505256]

79. Böckmann A et al. Characterization of different water pools in solid-state NMR protein samples. J. Biomol. NMR 45, 319–327 (2009). [PubMed: 19779834]

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). [PubMed: 20112896]

81. Williams JK & Hong M Probing membrane protein structure using water polarization transfer solid-state NMR. J. Magn. Reson 247, 118–127, doi:10.1016/j.jmr.2014.08.007 (2014). [PubMed: 25228502]

82. Mandala VS 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, doi:10.1038/s41594-020-00536-8 (2020). [PubMed: 33177698]

83. Gelenter MD et al. Water orientation and dynamics in the closed and open influenza B virus M2 proton channels. Commun. Biol 4, 338, doi:10.1038/s42003-021-01847-2 (2021). [PubMed: 33712696]

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, doi:10.1006/jmre.1997.1242 (1997). [PubMed: 9405219]

85. Krivov GG, Shapovalov MV & Dunbrack RL Jr. Improved prediction of protein side-chain conformations with SCWRL4. Proteins 77, 778–795, doi:10.1002/prot.22488 (2009). [PubMed: 19603484]

86. Lomize MA, Pogozheva ID, Joo H, Mosberg HI & Lomize AL 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). [PubMed: 21890895]

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. J. Comput. Chem 26, 1701–1718, doi:10.1002/jcc.20291 (2005). [PubMed: 16211538]
89. Huang J & MacKerell AD Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J. Comput. Chem 34, 2135–2145, doi:10.1002/jcc.23354 (2013). [PubMed: 23832629]
90. Jo S, Kim T, Iyer VG & Im W CHARMM-GUI: a web-based graphical user interface for CHARMM. J. Comput. Chem 29, 1859–1865, doi:10.1002/jcc.20945 (2008). [PubMed: 18351591]
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). [PubMed: 17849009]
92. Wu EL et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. J. Comput. Chem 35, 1997–2004, doi:10.1002/jcc.23702 (2014). [PubMed: 25130509]
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 Comput12, 405–413, doi:10.1021/acs.jctc.5b00935 (2016). [PubMed: 26631602]
94. Best RB 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. J. Chem. Theory Comput 8, 3257–3273, doi:10.1021/ct300400x (2012). [PubMed: 23341755]
95. Abraham MJ 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 JG, Peng Y, Silverstein DW & Swanson JM Multiscale Reactive Molecular Dynamics for Absolute pKa Predictions and Amino Acid Deprotonation. J. Chem. Theory Comput 10, 2729–2737, doi:10.1021/ct500250f (2014). [PubMed: 25061442]
97. Biswas R, Tse YL, Tokmakoff A & Voth GA 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). [PubMed: 26575795]
98. Day TJF, Soudackov AV, Čuma M, Schmitt UW & Voth GA A second generation multistate empirical valence bond model for proton transport in aqueous systems. J. Chem. Phys 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 MB, G.; Camilloni C; et al. Promoting transparency and reproducibility in enhanced molecular simulations. Nat. Methods 16, 670–673, doi:10.1038/s41592-019-0506-8 (2019). [PubMed: 31363226]
101. Tribello GA, Bonomi M, Branduardi D, Camilloni C & Bussi G PLUMED 2: New feathers for an old bird. Comput. Phys. Commun 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 JD 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 AT Acta Cryst. A48, 851–858. doi: 10.1107/S0108767392006044 (1992).

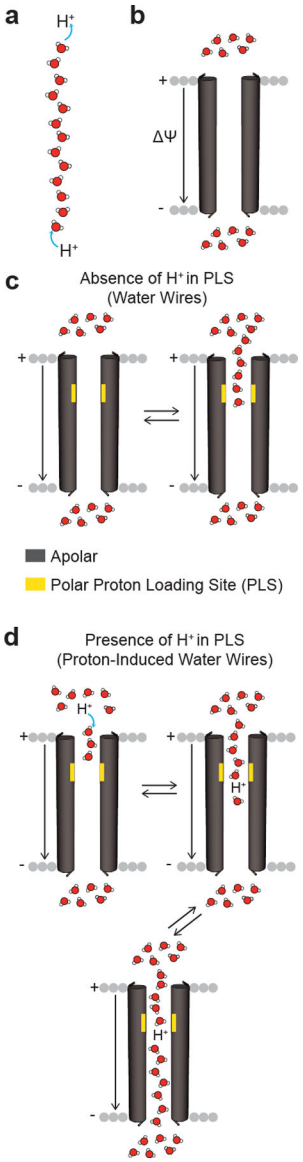
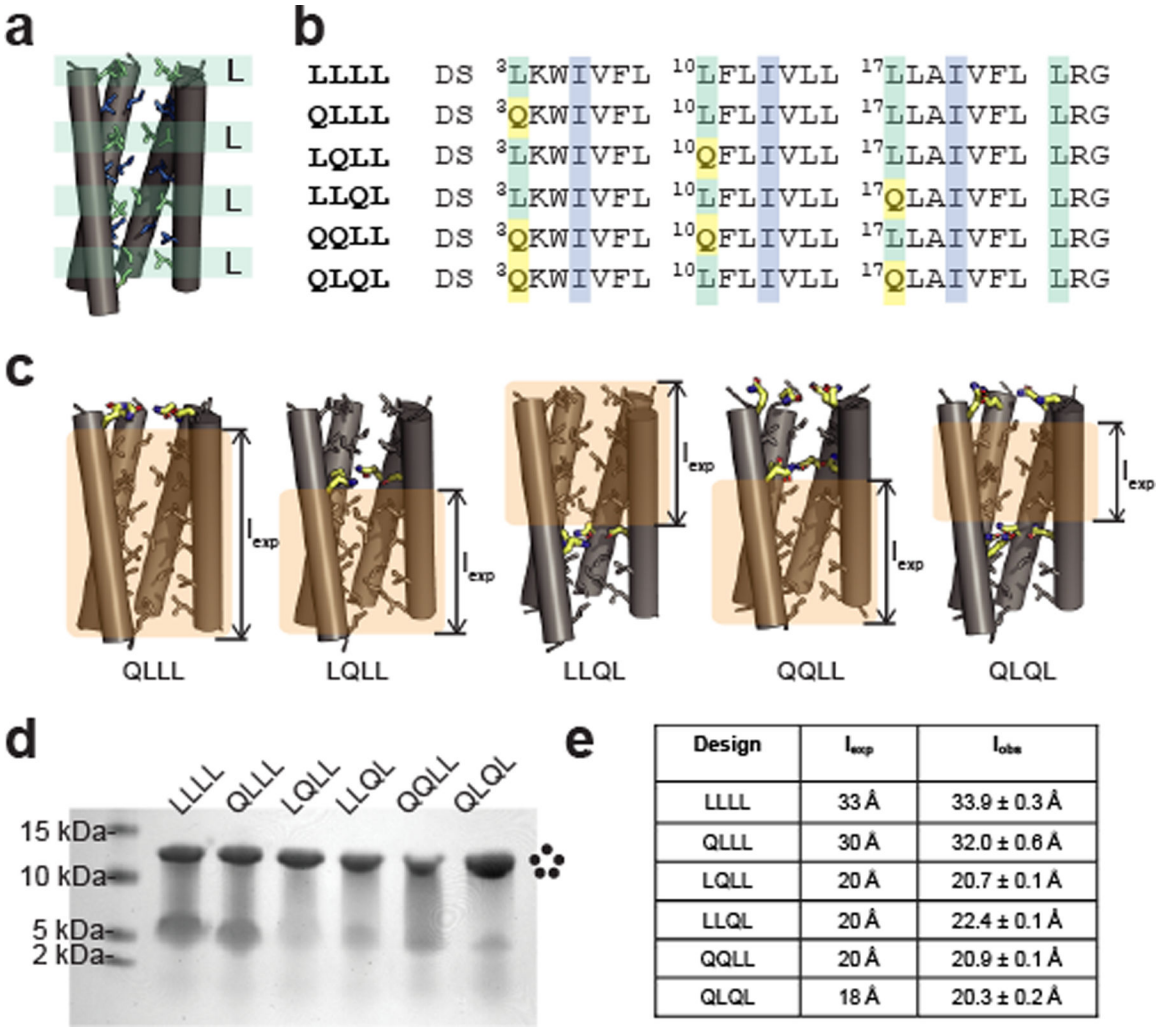



**Fig. 1. Hypothesis of proton-selective transport along transient water wires.**

**a**, Protons hop across dynamically rearranging hydrogen-bonded wires of water in the Grotthuss mechanism. **b**, In a wholly apolar channel, the pore remains devoid of water regardless of membrane polarization (ΔΨ), but a polar PLS, **c**, can mediate the flickering of neutral water in and out of a relatively short apolar sectors of the channel. **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.




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

**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. **c**, Models of the designed channels illustrates the different expected hydrophobic lengths ($l_{exp}$) of the channels relative to the parent scaffold. **d**, SDS-PAGE revealed that these designs formed stable pentamers. **e**, Comparison of the longest hydrophobic length expected ($l_{exp}$) and observed ($l_{obs}$, determined by classical MD simulations; see Methods and Extended Data Table 1). Side views of models in Fig. 2, 3, 4 and 6 have the fifth helix removed for clarity.

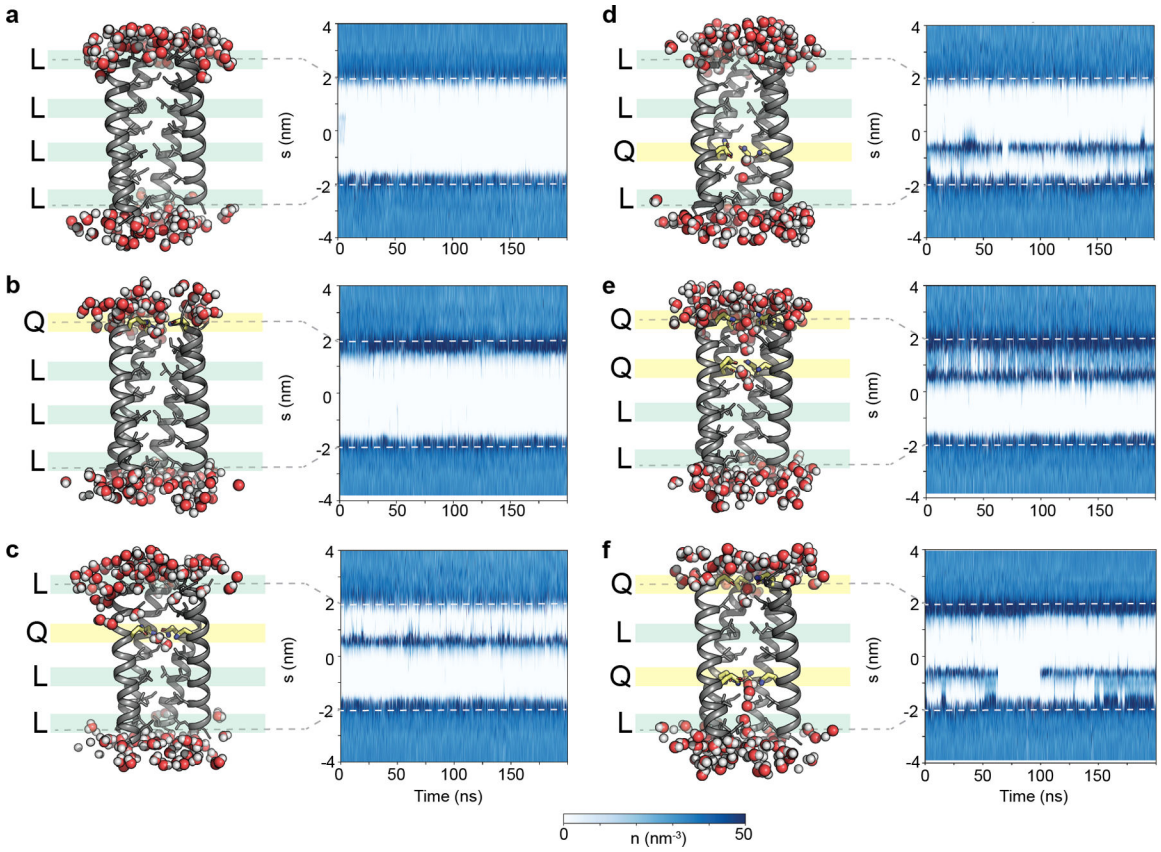



**Fig. 3. Permeation of water into the designed channels correlates with the position of the luminal Gln residues.**

MD simulations were analyzed using the Channel Annotation Package [44]. Plots of pore water density versus time reveal no water permeation into the hydrophobic pore in **a**, LLLL or **b**, QLLL. Pentamers with Gln in the second and third layers of the channel, including **c**, LQLL, **d**, LLQL, **e**, QQLL, and **f**, QLQL, have strong 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. Fifth helix in all figures removed for clarity.

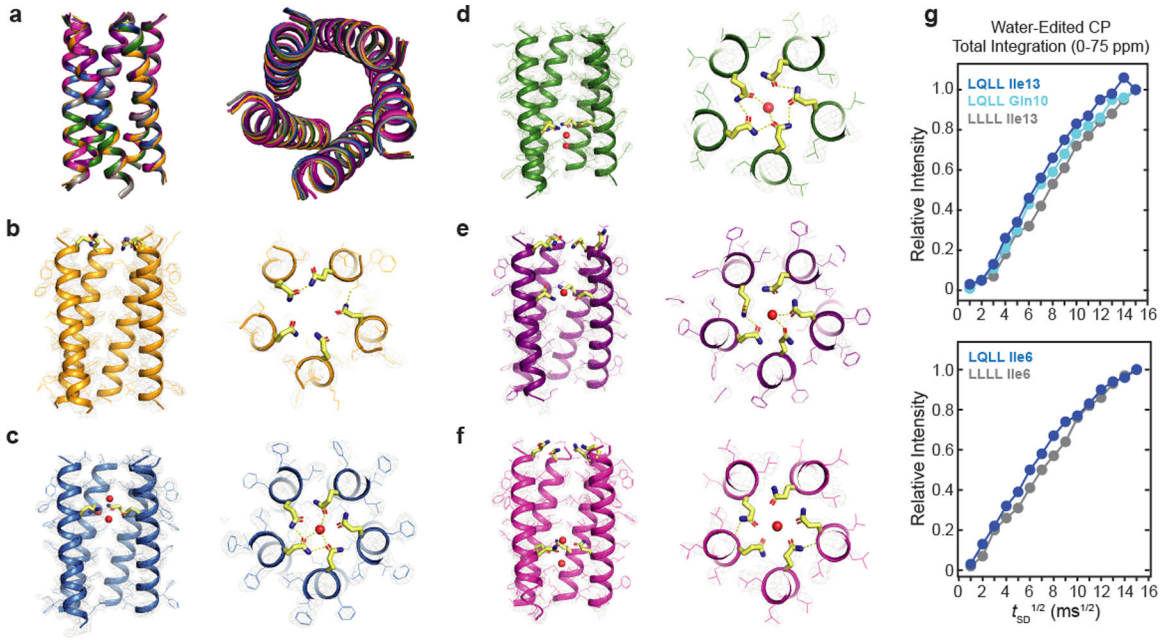



**Fig. 4. Crystal structures of the designed channels demonstrate the introduction of polar Gln residues mediates stable water pockets.**

The X-ray structures of the designed channels are within < 0.4 Å rmsd of the original design template, **a**, LLLL (pdb 6mct, gray), **b**, QLLL (pdb 7udy), **c**, LQLL (pdb 7udz), **d**, LLQL (pdb 7udv), **e**, QQLL (pdb 7udw), and **f**, QLQL (pdb 7udx). Fifth helix in all figures removed for clarity. **g,** Water buildup curves for uniformly labeled $^{13}$C, $^{15}$NIle13, Ile6, and Gln10 in LLLL and LQLL peptides. Both the Ile6 and Ile13 sites in the LQLL sample show faster water buildup than the corresponding sites in the LLLL sample.

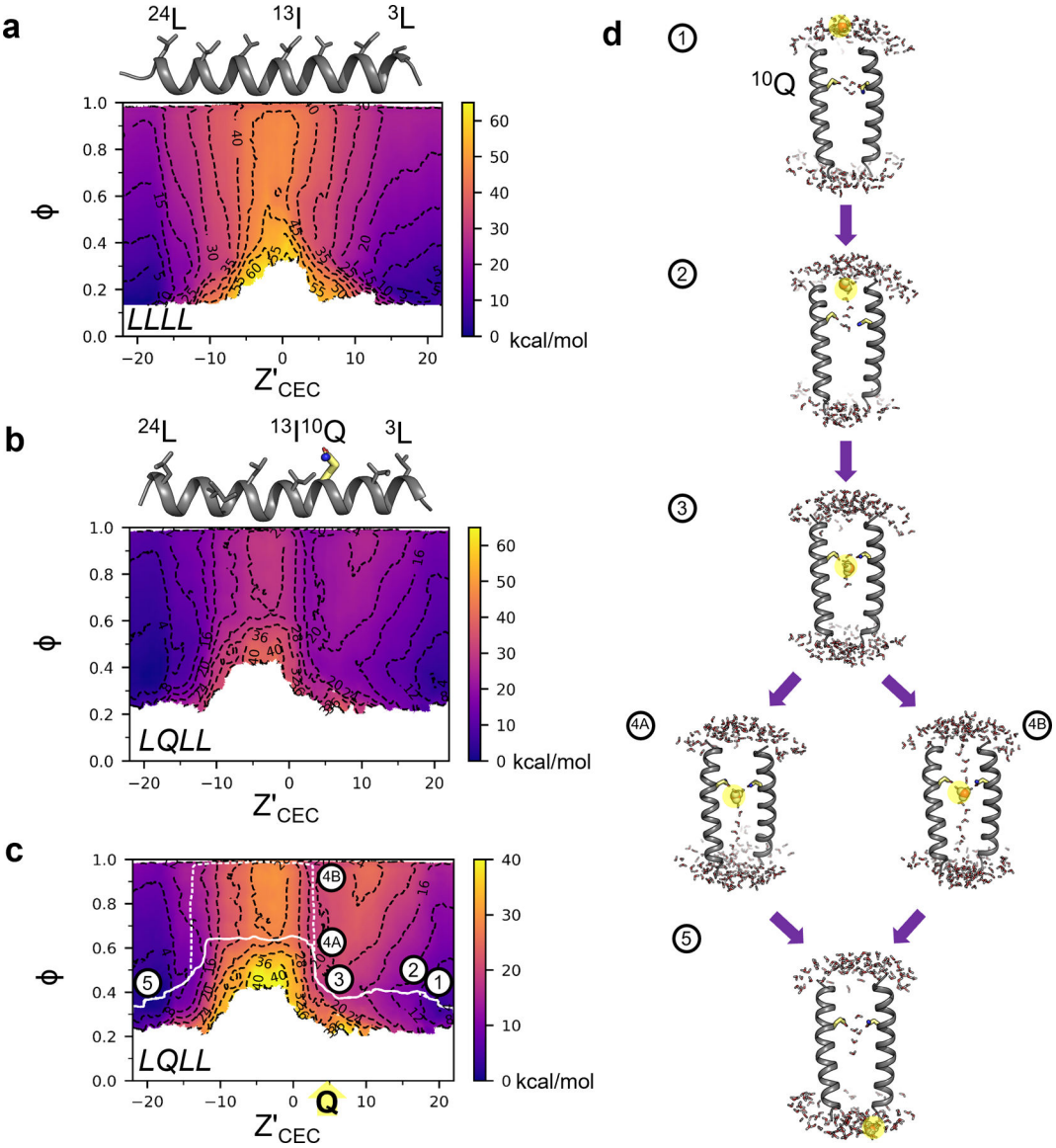



**Fig. 5. MS-RMD predicts that introduction of the PLS enables the formation of proton-conducive transient water wires.**

**a**, 2D-PMF of LLLL shows high barrier when the proton is at $Z'_{CEC}$ = 0 Å, or the center of mass of the channel at the Ile13 alpha-carbons. $Z'_{CEC}$ in **a-c** are in units of Å. **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. **c**, The two lowest mean free energy paths (MFEPs, white solid and dashed lines), derived from string theory (see Methods) through the LQLL channel. Note the scale change on the color bar in 5b versus 5c. **d**, Snapshots along the two pathways for LQLL (from panel 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.

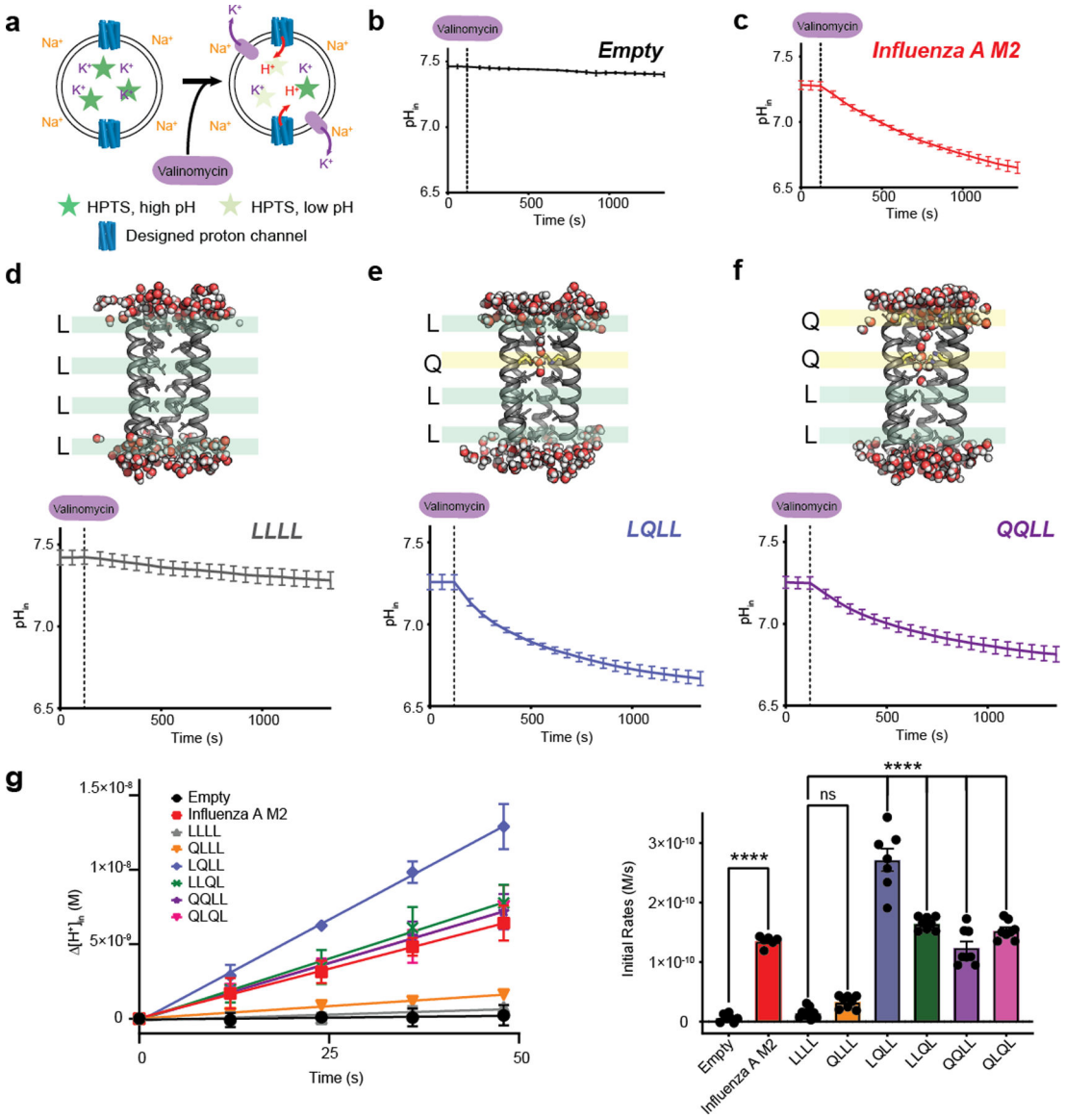



**Fig. 6. Designed channels selectively move protons across the membrane.**

**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, like 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 significant decrease in the $pH_{in}$ over time as protons move into the vesicle up a concentration gradient. Snapshots (fifth helix removed for clarity), and proton flux assays for **d**, LLLL, **e**, LQLL, and **f**, QQLL indicates that addition of polar Gln near the middle of the channel enables water permeation events into the pore which facilitates proton-selective transport. **g**, Change in the $H^+$ concentration of representative samples upon addition of valinomycin at t=0. For control samples, Influenza A M2 showed

significant ($p < 0.0001$) initial rates relative to Empty. QLLL initial rates are not significant ($p = 0.3947$) relative to LLLL. 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** presented as mean values +/− SD for $n > 6$ independent experiments. Data in **g** shown as mean values +/− SD (left) and +/− SEM (right) bars. Data analyzed using unpaired t-test for Empty vs. Influenza A M2 and ordinary one-way ANOVA for multiple comparisons (Dunn's test) of LLLL vs. designed channels (see Supplementary Tables S4 and S5 for complete analysis).