Single-chain heteropolymers transport protons selectively and rapidly

https://doi.org/10.1038/s41586-019-1881-0

Received: 31 December 2018

Accepted: 1 October 2019

Published online: 8 January 2020

Tao Jiang[1,11], Aaron Hall[1], Marco Eres[1,2], Zahra Hemmatian[3,12], Baofu Qiao[4], Yun Zhou[5], Zhiyuan Ruan[1], Andrew D. Couse[1,13], William T. Heller[6], Haiyan Huang[7], Monica Olvera de la Cruz[4,8], Marco Rolandi[3] & Ting Xu[1,2,9,10]*

Precise protein sequencing and folding are believed to generate the structure and chemical diversity of natural channels[1,2], both of which are essential to synthetically achieve proton transport performance comparable to that seen in natural systems. Geometrically defined channels have been fabricated using peptides, DNAs, carbon nanotubes, sequence-defined polymers and organic frameworks[3–13]. However, none of these channels rivals the performance observed in their natural counterparts. Here we show that without forming an atomically structured channel, four-monomer-based random heteropolymers (RHPs)[14] can mimic membrane proteins and exhibit selective proton transport across lipid bilayers at a rate similar to those of natural proton channels. Statistical control over the monomer distribution in an RHP leads to segmental heterogeneity in hydrophobicity, which facilitates the insertion of single RHPs into the lipid bilayers. It also results in bilayer-spanning segments containing polar monomers that promote the formation of hydrogen-bonded chains[15,16] for proton transport. Our study demonstrates the importance of the adaptability that is enabled by statistical similarity among RHP chains and of the modularity provided by the chemical diversity of monomers, to achieve uniform behaviour in heterogeneous systems. Our results also validate statistical randomness as an unexplored approach to realize protein-like behaviour at the single-polymer-chain level in a predictable manner.

The monomer selection was based on previously designed RHPs[14]. The two hydrophobic monomers, methyl methacrylate (MMA) and 2-ethylhexyl methacrylate (EHMA), promote RHP insertion into lipid bilayers. Oligo(ethylene glycol) methyl ether methacrylate (OEGMA; number average molecular weight, $M_n \approx 500$ Da) was included to tailor the overall hydrophilicity and promote hydrogen-bonded chain (HBC) formation[15,16] within the lipid bilayer for proton transport. 3-sulfopropyl methacrylate potassium salt (SPMA) was used to reduce the aggregation propensity of the RHPs.

Although statistically random on a whole-chain level, each RHP chain was composed of segments with a range of cumulative hydrophobicities. The RHP sequences were generated using Compositional Drift, a program based on the Mayo–Lewis model, which considers the relative concentrations and the reactivity ratios of each comonomer pair, and on global monomer conversion[17]. Figure 1a shows 20 representative sequences of the RHP (termed 'RHP1') with a feeding compositional ratio of MMA:OEGMA:EHMA:SPMA = 5:2.5:2:0.5. Although the RHP1 sequences are statistically random, the heterogeneity is obvious in the cumulative hydrophobicity among different local segments within and among RHP1 chains. The local hydrophobicity along an RHP1 chain can be evaluated on the basis of the average hydrophilic–lipophilic balance (HLB) value of a sliding window (Fig. 1b)[18]. An RHP chain can be broken down into three types of segments according to the cumulative HLB value: those that prefer water, the lipid bilayer and the interface (Fig. 1c). With the same ensemble composition and statistical monomer distribution, individual RHP chains feature different sequences. However, those types of segment are seen for all calculated RHP1 sequences. The insertion of RHP1 into lipid bilayers was confirmed by fluorescence imaging and differential scanning calorimetry (DSC) measurements. Figure 1d shows the confocal images of liposomes incubated with RHP1 end-labelled with Texas Red dye. After washing off the excess polymer, the fluorescence signal from RHP1 was detected near the lipid membranes and overlapped with the signal from a premixed membrane-specific dye (NBD-PE). The DSC results show the broadening of the lipid phase transition in the RHP1-containing liposomes, confirming RHP1 insertion, rather than only surface absorption (Extended Data Fig. 1a, b).

We ran all-atom explicit solvent molecular dynamics simulations to visualize the distributions of RHP1 within a lipid bilayer (Fig. 1e,

[1]Department of Materials Science and Engineering, University of California Berkeley, Berkeley, CA, USA. [2]Department of Chemistry, University of California Berkeley, Berkeley, CA, USA. [3]Department of Electrical and Computer Engineering, University of California Santa Cruz, Santa Cruz, CA, USA. [4]Department of Materials Science and Engineering, Northwestern University, Evanston, IL, USA. [5]Division of Biostatistics, University of California Berkeley, Berkeley, CA, USA. [6]Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. [7]Department of Statistics, University of California Berkeley, Berkeley, CA, USA. [8]Department of Chemistry, Northwestern University, Evanston, IL, USA. [9]Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [10]Kavli Energy NanoScience Institute, University of California Berkeley, Berkeley, CA, USA. [11]Present address: Department of Chemistry, The MOE Key Laboratory of Spectrochemical Analysis and Instrumentation, Xiamen University, Xiamen, China. [12]Present address: Quantapore Inc., San Francisco, CA, USA. [13]Present address: Department of Chemistry, Indiana University Bloomington, Bloomington, IN, USA. *e-mail: tingxu@berkeley.edu

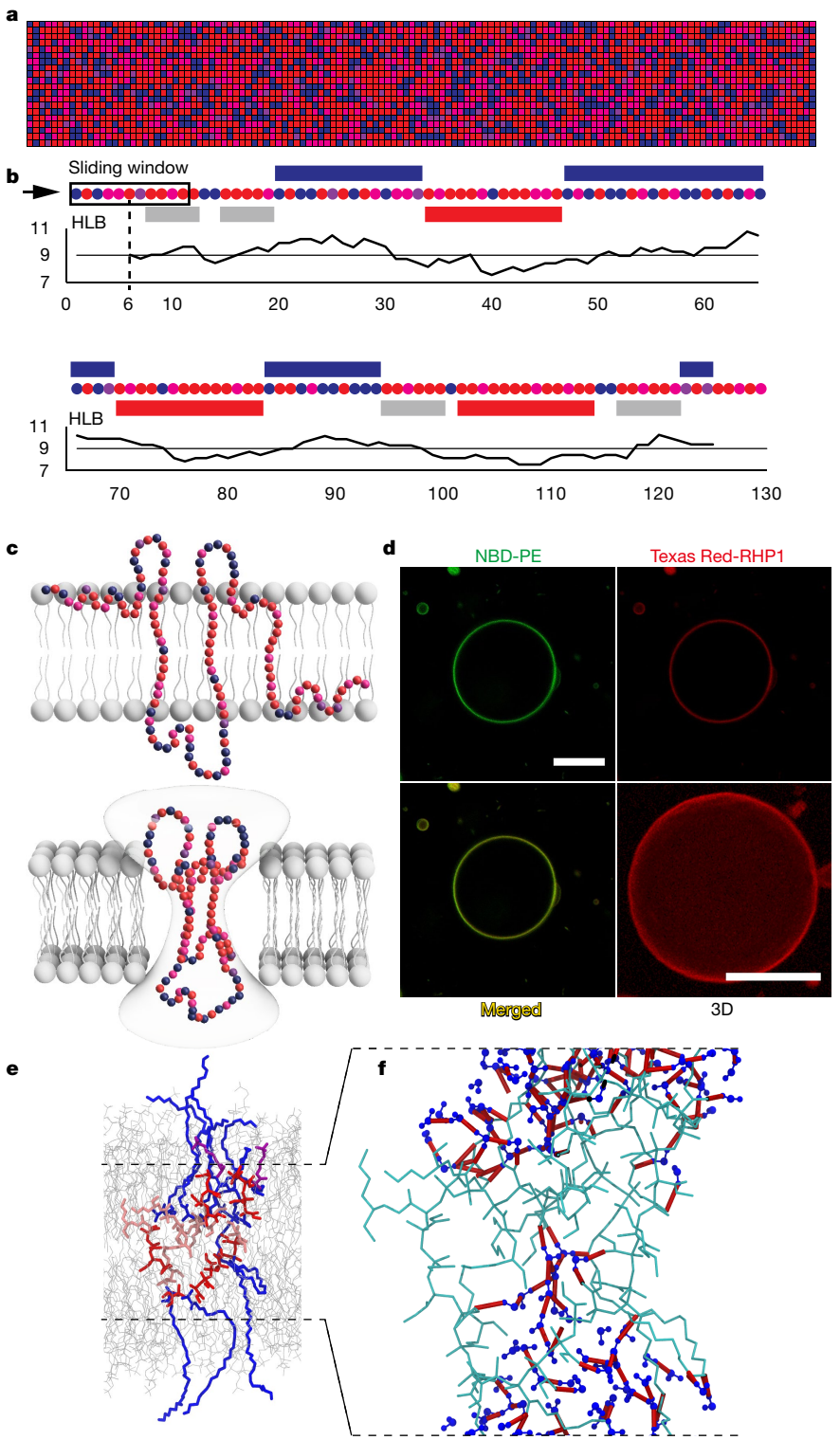


Fig. 1 | Segmental heterogeneity in RHP. a, Simulated sequences of 20 RHP1 chains (degree of polymerization, DP = 130). MMA, EHMA, OEGMA and SPMA are shown in red, pink, blue and purple, respectively. **b**, Hydrophilic (blue), long (red) and short (grey) hydrophobic segments in an RHP1 chain. Average HLB values for a sliding window of 11 monomers versus the monomer position are shown by the black line. **c**, Two-dimensional (2D) and three-dimensional (3D) illustrations of an RHP1 in a lipid bilayer. **d**, Confocal fluorescence images and 3D image reconstruction of an NBD-PE-labelled POPC liposome after incubation with Texas Red-labelled RHP1. Scale bar, 20 μm. **e**, Snapshot of the spatial distribution of an RHP1 chain (the first sequence in Extended Data Fig. 2a) in the lipid bilayer in the all-atom molecular dynamics simulation. **f**, Snapshot of the HBCs (red sticks) formed within a distance of 3.5 Å of the RHP1.

Extended Data Fig. 2). Most hydrophilic side chains are found to protrude into the water or near the membrane surfaces. Hydrophobic segments are distributed within the bilayer, anchoring a few OEG side chains to the nonpolar region of the bilayer. Those segments are rather static, and the OEG side chains in the bilayer move at slower rates than the ones in water (Supplementary Videos 1 and 2). The inserted RHP1 promotes the formation of rather dynamic HBCs among OEGs, methacrylate carbonyls and waters in the bilayer without adopting uniform three-dimensional structures (Fig. 1f, Supplementary Videos 3 and 4) and should act as 'stepping stones' to assist proton transport across the bilayers.

A bioprotonic device (Fig. 2a) was used to probe proton transport through a supported lipid bilayer (SLB) composed of POPE and POPG (see Methods; POPE:POPG molar ratio, 3:1) on top of a Pd/PdH_x contact, which translates a proton current into an electronic current[19]. With −200 mV (versus Ag/AgCl) applied to the Pd contact, incorporating RHP1 (RHP1:lipid molar ratio, 0.0005) into the SLB increased the current of H^+ across the SLB with respect to the SLB alone. After H^+ permeated

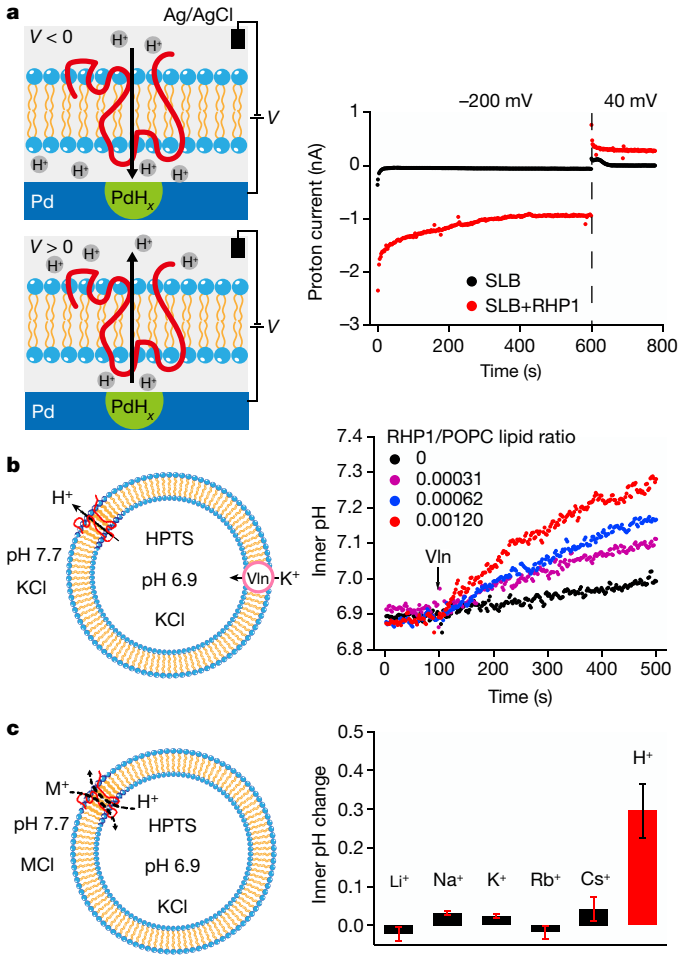


Fig. 2 | Transmembrane proton transport. a, Voltage-driven proton transport through RHP1 incorporated in an SLB on the Pd/PdH_x device. **b**, pH gradient-driven proton flux through RHP1 in the liposome. The arrow indicates the time when Vln was added. **c**, Inner pH changes to external alkali metal cations after an incubation time of 200 s without Vln (black bars). The red bar denotes the pH change with Vln and external K^+. Error bars are 1 s.d. ($n = 3$). The RHP1-to-lipid molar ratio is 0.0017.

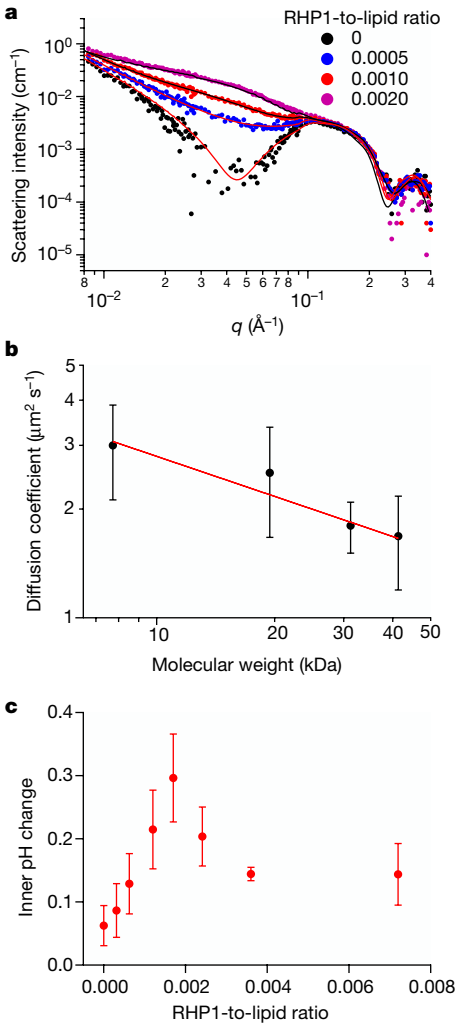


Fig. 3 | Interaction of RHP1 and the lipid bilayers. a, SANS profiles of d54-DMPC liposomes containing RHP1(~19 kDa) in 86% D_2O:H_2O. The solvent signal was subtracted. Solid lines denote a fit using a three-layer vesicle model for the liposome and a sphere model for the RHP1. **b**, Diffusion coefficients of the Texas Red-labelled RHP1s in the liposome membrane as a function of polymer molecular weight (7, 19, 30 and 40 kDa). Error bars are 1 s.d. ($n = 6$). **c**, Inner pH changes of the liposomes incubated with RHP1 over 200 s after adding Vln. Error bars are 1 s.d. ($n = 3$).

through the SLB, it was reduced at the Pd surface into H, which physisorbed onto the Pd surface to form PdH_x. When the voltage on the Pd contact was switched from −200 mV to +40 mV, the positive voltage oxidized H into H^+ and resulted in a positive proton current, i_{H+}, across the SLB. A higher positive i_{H+} was also detected for the RHP1 containing SLB. Using these data, we calculated that adding RHP1 increased the apparent membrane permeability of the SLB from 8.1×10^{-5} s^{-1} to 1.6×10^{-2} s^{-1}. The results confirm that RHP1 effectively promotes proton transport across the SLB.

The proton transport was further evaluated using a liposome-based fluorescence assay, in which the proton permeation was monitored using a pH-sensitive fluorophore pyranine (HPTS). The pH change was recorded from liposomes composed of the lipid POPC (see Methods; Fig. 2b) or the lipids POPE and POPG (POPE:POPG molar ratio, 3:1) (Extended Data Fig. 3a, b). Driven by a pH gradient of about 0.8, enhancement in the proton flux was observed by adding RHP1 to the liposome solutions, which was triggered by the K^+ ionophore valinomycin (Vln) that dissipates membrane potentials. To quantify the proton flux rate per RHP1 chain, we premixed the lipid and RHP1 before liposome preparation and went through freeze–thaw cycles to promote RHP1 insertion. We assumed that all RHP1 chains contribute to the proton transport and the calculated transport value represents the lower limit of the RHP1 performance. The stopped-flow fluorescence measurements show an initial proton flux per RHP1 of 8.1 ± 1.0 H^+ s^{-1} under these conditions, corresponding to a single-RHP1 conductance value of $(4.4 \pm 0.57) \times 10^{-17}$ S ($n = 3$, s.d.) (Extended Data Fig. 4a–d). This is comparable to the value reported for the proton-selective channel influenza virus M2, 2.9×10^{-17} S to 4.4×10^{-17} S, at pH 5.7 using a similar assay[20]. As a control, the conductance of the proton/cation channel gramicidin A (GramA) was measured to be $(4.0 \pm 0.43) \times 10^{-16}$ S ($n = 3$, s.d.), in the same magnitude as the reported values of 1.2×10^{-16} S to 2.1×10^{-16} S (refs. [21,22]). Slowed proton transport was observed after substituting D_2O for H_2O in this assay (Extended Data Fig. 4e). The isotope effect value of RHP1 (1.9) is larger than that of water pore channels (1.6), where protons are transported along water wires via the Grotthuss mechanism[13]. The large isotope effect agrees with the proposed transport pathway involving both RHP1 and water, rather than RHP1 forming a pore in the bilayer[13,23].

The insertion of RHP1 did not increase the water permeability of the bilayer, confirming the absence of distinct membrane leakage (Extended Data Fig. 4f). The direction of proton transport was varied to eliminate the snorkelling effect from the OEG side chains

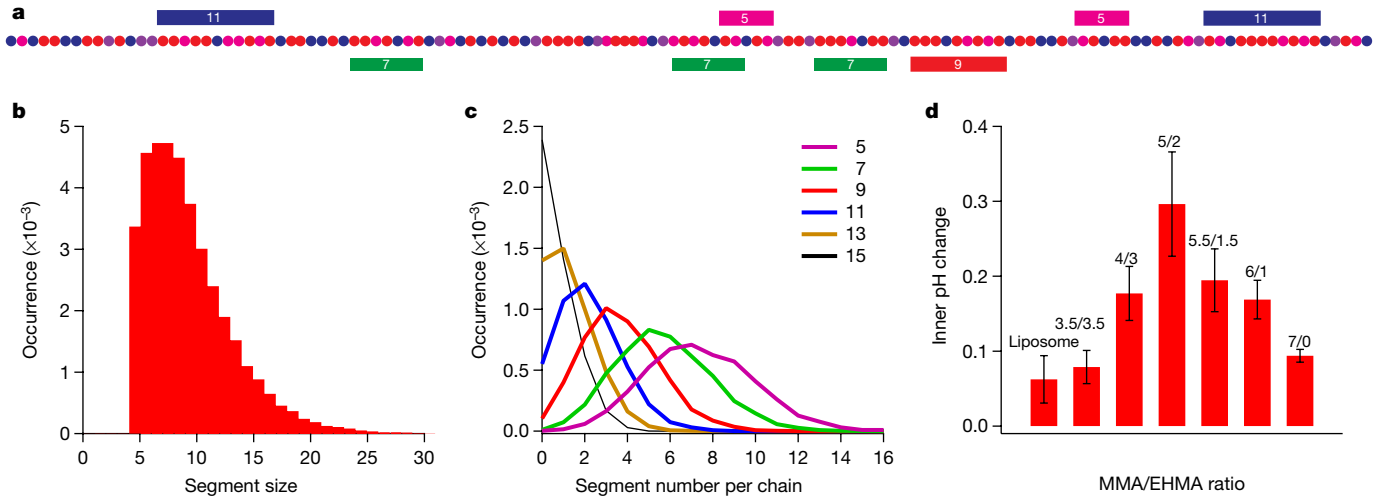


Fig. 4 | Sequence analyses of the RHPs. a, One set of hydrophobic segments containing one OEGMA in an RHP1 chain (DP = 130). **b**, Statistical size distribution of the hydrophobic segments containing one OEGMA in 4,500 simulated RHP1 sequences. **c**, Number of such segments per polymer chain, with the minimum segment size ranging from 5 to 15. Segments of even numbers are omitted for clarity. **d**, Transport performance of the RHP variants, evaluated using the liposome-based fluorescence assay (Fig. 2b). Inner pH changes of the RHP-containing liposomes (RHP:lipid ratio of 0.0017) over 200 s after adding Vln. Error bars are 1 s.d. ($n = 3$).

of bilayer-surface-absorbed RHP1 chains (Extended Data Fig. 3c). In addition, an RHP1 derivative (RHP1,100) containing an OEGMA with a longer side chain ($M_n \approx 1{,}100$ Da) was tested. Although it favours bilayer-surface binding, RHP1,100 did not promote proton permeation under the same conditions (Extended Data Figs. 1c, d, 3b).

Small-angle neutron scattering (SANS) was used to characterize the RHP1 size and oligomeric state in a lipid bilayer (Fig. 3a, Extended Data Fig. 5). Owing to the large scattering length density (SLD) contrast against RHP1, d54-DMPC (see Methods) was used to differentiate the signal of RHP1 from that of the lipid background while still providing RHP1 proton transport activity (Extended Data Fig. 3d). The SANS measurements provided an average RHP1 (about 19 kDa) radius of 2.2 nm. The physical volume occupied by one dry RHP1 chain was calculated to be about 32 nm^3, assuming a polymer density of 1.0 g ml^{-1}. Using a sphere model for single RHP1 chains, the SANS results lead to an average particle size of about 45 nm^3. Thus, the RHP1 chains formed a unimolecular assembly under the experimental conditions used. Fluorescence recovery after photobleaching (FRAP) measurements were performed on the liposome surface using RHP1s with molecular weights ranging from ~7 kDa to ~40 kDa (Fig. 3b). Despite large fluctuations in the measured diffusion rate (D), the correlation between D and the RHP1 molecular weights further excludes extensive polymer aggregation in the lipid bilayer. Additionally, in the proton transport study, promoted proton transport was observed with RHP1-to-lipid ratios increasing from 0 to 0.0017 (RHP1 concentration, 0–48 nM) (Fig. 3c). The trend of the pH increase is reversed at higher RHP1 concentration, probably owing to the formation of RHP1 oligomers that were excluded from the lipid bilayer. Together, these results confirmed that the monomeric form of RHP1 is the active species for the polymer concentration window reported here.

Diffusion of alkali metal cations across a lipid bilayer requires a pore size larger than the size of the dehydrated cations, whereas protons can either diffuse through pores in the form of water–proton complexes or be transported without pores by hopping along HBCs[15,16]. The ionic selectivity of RHP1-based transport against other alkali metal cations (including Li^+, Na^+, K^+, Rb^+ and Cs^+) was studied using an external cation exchange approach[10] (Fig. 2c). In the presence of both proton and cation gradients across the bilayer, no distinct co-transport of protons and alkali metal ions was observed, which was manifested by minimal pH changes in the interior of the liposomes. Such insensitivity indicates that RHP1-based transport is highly selective towards protons. This is consistent with the proposed HBC-based transport mechanism in the transmembrane RHP1 segments and the unimolecular nature of RHP1 within the bilayer.

An important hallmark of RHPs is their sequence heterogeneity. Even with the same ensemble composition and statistical monomer distribution, individual RHP chains feature different sequences. However, the statistical population and distribution of segments according to cumulative hydrophobicity should remain the same. Despite the heterogeneity, RHP1 shows consistent transport performance. This may be attributed to the enhanced adaptability that is unavailable in more defined systems, and may be advantageous for achieving efficient molecular transport. Analyses were performed to gain insight into the interplay between differences in monomer sequences and similarities in the statistical structure of segments (that is, similar statistical parameters, which characterize the heterogeneity of segments) and how these may affect the RHP1 arrangement in lipid insertion and proton transport. Segments rich in MMA and EHMA can be overall hydrophobic while still including a few OEGMA monomers, and sufficiently long to span across lipid bilayers. Similarly, there exist segments rich in hydrophilic monomers that include a few hydrophobic units, which generally prefer the aqueous phase. Both are chemically heterogeneous but with distinctive cumulative hydrophobicities. Within each RHP1 chain, there are also abundant segments that cannot be easily assigned as overall hydrophobic or hydrophilic. Rather, their cumulative chemical features vary substantially depending on the choice of segment length and ending position. Nevertheless, they appear to be generally amphiphilic and will probably have a short residence time in either phase. Their rapid kinetics of insertion and desorption from bilayers, instead of a static state, may also facilitate proton transport.

We hypothesize that the inclusion of a small amount of OEGMA in the lipid bilayer core is a key requirement for the observed proton transport. Analysing the hydrophobic segments including OEGMA can be done with various criteria over the block length, number of OEGMAs included and flanking monomer sequence. As one example, Fig. 4a highlights the hydrophobic segments containing one OEGMA in a representative RHP1 chain. An in-house-developed program was used to evaluate the length distribution of such segments in a total of 4,500 RHP1 (degree of polymerization, DP = 130) sequences generated by Compositional Drift, considering compositional dispersity

(Extended Data Fig. 6a). As seen in Fig. 4b, RHP1 contains the segments that fulfil the length requirement to span the lipid bilayers. The core of a typical lipid bilayer (~2.7 nm for a POPC bilayer[24]) can theoretically accommodate 11 MMA-based units in their fully extended form. Within a certain length range, shorter and longer segments should also be capable of integrating into the bilayer. RHP1 segments and surrounding lipids can adjust their chain conformation and/or bilayer thickness, respectively, to adapt to hydrophobic mismatch. Further analysis, shown in Fig. 4c, suggests that there are a limited number of such segments per RHP1 chain (DP = 130). For example, an average of three segments per polymer chain was counted with a segmental length of ≥9, and two segments per chain with length ≥11. RHP1 length dispersity was also considered in the segment analyses. The statistical distribution of the segment size is less affected by the polymer length variation once the RHP1 is sufficiently long (DP ≥ 40) (Extended Data Fig. 6b, c), although the segment numbers per chain have a positive correlation with the RHP1 chain length (Extended Data Fig. 6d). We also used the probabilistic framework of the hidden Markov model (HMM) to predict the transmembrane segments of RHP1 (Extended Data Fig. 7). The HMM investigation agrees with the aforementioned sequence analysis, indicating that RHP1 chains are rich in transmembrane segments and include those that allow OEGMAs to embed in the bilayer core to facilitate proton transport. Such hydrophobic segments are limited in their population for the molecular weight range studied here, thus forming a narrow crossing path for proton transport.

According to the information gained, segmental heterogeneity afforded by four distinct monomers appears to be a key design feature that is necessary to capture several critical features of natural channels with regards to proton transport. Experimentally, none of the three-monomer-based RHPs shows efficient transport activity (Extended Data Fig. 3b). It is reasonable to believe that a statistical design approach should also lead to other RHPs with proton transport properties. We synthesized a series of RHP variants, varying the MMA:EHMA ratio from 3.5:3.5 to 7:0. The overall fraction of these two hydrophobic monomers remains the same in order to maintain similar segmental distributions. The RHPs within the compositional range of 4:3 to 6:1 MMA:EHMA could promote proton transport, albeit with different performances (Fig. 4d). RHP1 (MMA:EHMA = 5:2) elicits the best proton permeation, whereas either increasing or decreasing the portion of EHMA reduces the overall transport performance. Shifting the MMA:EHMA ratio should alter the capability of polymer insertion into lipid bilayers by changing the overall polymer hydrophobicity/hydrophilicity and the balance between polymer–polymer and polymer–lipid interactions. Decreasing the MMA:EHMA ratio increases the occurrence of EHMA monomers distributed adjacent to each other. With a bulky ethyl hexyl side-chain, local abundance of EHMA would also sterically hinder the interplay between the polymer carbonyl and water/OEGMA, affecting the proton transport efficiency (Extended Data Fig. 6e, 6f).

The present study confirms that exploring and manipulating the segmental heterogeneity in RHPs at the single-chain level is feasible and effective to mimic protein-like behaviour. Considering the vast number of monomers available and recent advances in polymer chemistry, the possibility to generate RHP-based functional materials[14,25–27] and eventually marry the synthetic and biological fields are almost unlimited.

Online content

Any methods, additional references, Nature Research 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/s41586-019-1881-0.

1. Stouffer, A. L. et al. Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* **451**, 596–599 (2008); corrigendum **452**, 380 (2008).
2. Schnell, J. R. & Chou, J. J. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **451**, 591–595 (2008).
3. Ghadiri, M. R., Granja, J. R. & Buehler, L. K. Artificial transmembrane ion channels from self-assembling peptide nanotubes. *Nature* **369**, 301–304 (1994).
4. Xu, T. et al. Subnanometer porous thinfilms by the co-assembly of nanotube subunits and block copolymers. *ACS Nano* **5**, 1376–1384 (2011).
5. Joh, N. H. et al. De novo design of a transmembrane Zn^{2+}-transporting four-helix bundle. *Science* **346**, 1520–1524 (2014).
6. Langecker, M. et al. Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **338**, 932–936 (2012).
7. Hinds, B. J. et al. Aligned multiwalled carbon nanotube membranes. *Science* **303**, 62–65 (2004).
8. Geng, J. et al. Stochastic transport through carbon nanotubes in lipid bilayers and live cell membranes. *Nature* **514**, 612–615 (2014).
9. Percec, V. et al. Self-assembly of amphiphilic dendritic dipeptides into helical pores. *Nature* **430**, 764–768 (2004).
10. Weiss, L. A., Sakai, N., Ghebremariam, B., Ni, C. Y. & Matile, S. Rigid rod-shaped polyols: functional nonpeptide models for transmembrane proton channels. *J. Am. Chem. Soc.* **119**, 12142–12149 (1997).
11. Vial, F., Oukhaled, A. G., Auvray, L. & Tribet, C. Long-living channels of well defined radius opened in lipid bilayers by polydisperse, hydrophobically-modified polyacrylic acids. *Soft Matter* **3**, 75–78 (2007).
12. Jung, M., Kim, H., Baek, K. & Kim, K. Synthetic ion channel based on metal–organic polyhedra. *Angew. Chem.* **47**, 5755–5757 (2008).
13. Si, W. et al. Selective artificial transmembrane channels for protons by formation of water wires. *Angew. Chem.* **50**, 12564–12568 (2011).
14. Panganiban, B. et al. Random heteropolymers preserve protein function in foreign environments. *Science* **359**, 1239–1243 (2018).
15. Nagle, J. F. & Morowitz, H. J. Molecular mechanisms for proton transport in membranes. *Proc. Natl Acad. Sci. USA* **75**, 298–302 (1978).
16. Decoursey, T. E. Voltage-gated proton channels and other proton transfer pathways. *Physiol. Rev.* **83**, 475–579 (2003).
17. Smith, A. A. A., Hall, A., Wu, V. & Xu, T. Practical prediction of heteropolymer composition and drift. *ACS Macro Lett.* **8**, 36–40 (2019).
18. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
19. Hemmatian, Z. et al. Electronic control of H^+ current in a bioprotonic device with Gramicidin A and Alamethicin. *Nat Commun.* **7**, 12981 (2016)
20. Lin, T. I. & Schroeder, C. Definitive assignment of proton selectivity and attoampere unitary current to the M2 ion channel protein of influenza A virus. *J. Virol.* **75**, 3647–3656 (2001).
21. Krishnamoorthy, G. Temperature jump as a new technique to study the kinetics of fast transport of protons across membranes. *Biochemistry* **25**, 6666–6671 (1986).
22. 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).
23. DeCoursey, T. E. & Cherny, V. V. Deuterium isotope effects on permeation and gating of proton channels in rat alveolar epithelium. *J. Gen. Physiol.* **109**, 415–434 (1997).
24. Kučerka, N., Tristram-Nagle, S. & Nagle, J. F. Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains. *J. Membr. Biol.* **208**, 193–202 (2006).
25. Golumbfskie, A. J., Pande, V. S. & Chakraborty, A. K. Simulation of biomimetic recognition between polymers and surfaces. *Proc. Natl Acad. Sci. USA* **96**, 11707–11712 (1999).
26. Mowery, B. P. et al. Mimicry of antimicrobial host-defense peptides by random copolymers. *J. Am. Chem. Soc.* **129**, 15474–15476 (2007).
27. Tribet, C., Audebert, R. & Popot, J. L. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc. Natl Acad. Sci. USA* **93**, 15047–15050 (1996).

Methods

Materials

Chemicals were purchased from Sigma Aldrich and Fisher Scientific unless otherwise noted. Phospholipids were purchased from Avanti Polar Lipids Inc., including 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1′-rac-glycerol) (sodium salt) (POPG), 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC), 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) and 1,2-dimyristoyl-d54-sn-glycero-3-phosphocholine (d54-DMPC). Water was Milli-Q-grade water (18.2 MΩ cm).

RHP syntheses

RHPs were synthesized as described previously[14]. Azobisisobutyronitrile was recrystallized from ethanol. Inhibitors were removed using cryodistillation (methyl methacrylate, 2-ethylhexyl methacrylate) or by passing over a column of neutral alumina (ethylene glycol methyl ether methacrylate ($M_n \approx 500$ g mol^{-1}). 3-sulfopropyl methacrylate potassium salt (98%), ethyl-2(phenylcarbanothioylthio)-2-phenylacetate (98%), trioxane and solvents were used as received. Polymerization solutions were prepared and transferred to 20-ml glass ampules. These solutions were subjected to four freeze–pump–thaw cycles and then flame-sealed at ~50 mtorr. Polymerizations were carried out at 80 °C for typically 6–12 h. The ampules were then quenched in liquid nitrogen and opened. Polymers were precipitated by dropwise addition to stirring pentane and then transferred to 3,500 molecular-weight cut-off centrifuge filters (Amicon), washed five times using water and lyophilized.

^{1}H nuclear magnetic resonance (NMR) spectra were obtained with either a Bruker Avance 400 spectrometer (400 MHz) using a 5-mm *Z*-gradient broadband observe probe, a Bruker Avance AV 500 spectrometer (500 MHz) using a *Z*-gradient triple broadband inverse detection probe, or a Bruker Avance AV 600 spectrometer (600 MHz) using a *Z*-gradient triple broadband inverse detection probe. Monomer conversion was measured on crude reaction mixtures in DMSO-d_6, using trioxane as an internal standard. The RHPs were characterized using gel permeation chromatography with an Agilent 1260 Infinity series instrument equipped with two Agilent PolyPore columns (300 mm × 7.5 mm), calibrated using poly(ethylene glycol) standards. Dimethylformamide with 0.05 M LiBr was used as eluent at 0.7 ml min^{-1} at 50 °C. Analyte samples at 2 mg ml^{-1} were filtered through 0.2-μm polytetrafluoroethylene membranes (VWR) before injection (20 μl). The physiochemical properties of the synthesized RHPs are listed in Extended Data Table 1.

Preparation of large unilamellar vesicles

2.5 mg of POPC or a mixture of POPE and POPG (3:1 molar ratio) was dissolved in chloroform and subjected to solvent removal to make a lipid film, which was dried in vacuum overnight. The lipids were rehydrated by a 0.5 ml inner buffer (10 mM HEPES-KOH, pH 6.9, 100 mM KCl) at 37 °C for 1 h. 0.5 mM pyranine was included in the inner buffer in the pyranine assay. The resulting suspension was subject to six cycles of freeze–thawing using a dry ice–ethanol bath and a 37 °C water bath. Large unilamellar vesicles (LUVs) were generated by extruding 11 times through a 400-nm polycarbonate membrane followed by a 200-nm polycarbonate membrane. Free dye was removed using a Sephadex G-25 desalting column (GE healthcare). LUV formation was confirmed using dynamic light scattering and transmission electron microscopy. The LUVs composed of POPE+POPG and of POPC only had diameters of 182 ± 16 nm and 189 ± 9 nm, respectively. The lipids were quantified using the Stewart method[28].

Preparation of giant unilamellar vesicles

Giant unilamellar vesicles (GUVs) for microscopy and FRAP investigations were prepared using a reported agarose-based hydration method[29]. 300 μl of 1% melted agarose solution (low melt; gelling temperature, $T_g \approx 26$ °C; melting temperature, $T_m \leq 65$ °C; electroendosmosis, EEO ≤ 0.12) was added to a cover glass (22 mm × 22 mm) and dried on a hotplate. 80 μg POPC in chloroform was evenly spread on the agarose film. The chloroform was removed under vacuum. The film was covered by a 1-ml HEPES buffer (10 mM HEPES-KOH, pH 6.9) to induce GUV growth for 0.5 h. Liposomes were released by gentle pipetting and immobilized into a thin layer of 0.5% agarose on a glass slide for microscopy investigations.

Interactions between RHPs and liposomes

Confocal fluorescence imaging. RHPs were end-modified with a Texas Red fluorescent dye via thiol-maleimide chemistry. GUVs doped with NBD-PE (100:1 molar ratio) were incubated with the Texas Red-labelled RHP1s (2,000:1 molar ratio) for 0.5 h. Unbound polymers were rinsed off right before imaging. The imaging was performed on a Zeiss LSM 880 with a 63× (numerical aperture, NA, 1.4) oil or 20× (NA, 1.0) water-immersion objectives and one Airy unit. Lasers at 488 nm and 594 nm were employed to excite the NBD and Texas Red, respectively. Fluorescence signals were collected in the 500–550 nm and 605–655 nm bands.

DSC investigation. DSC thermograms probing the lipid phase transition were acquired using a VP-MicroCal calorimeter (GE). The samples and buffers were equilibrated at 5 °C for 30 min and heated from 5 °C to 40 °C at a rate of 1 °C min^{-1}. The heating and cooling process was repeated for several cycles. The buffer or the RHP-containing buffers were used as the background. The full-width at half-maximum of the peak was obtained with baseline correction using the Origin software (Microcal).

FRAP measurement. FRAP experiments were conducted using Texas Red-labelled RHP1 containing-GUVs. Five prebleach images were taken at an attenuated laser intensity (0.5%, 561 nm). Subsequently, photobleaching was performed through a circular region of interest of nominal radius 3 μm and 20 bleaching iterations using the 514-nm and 561-nm laser lines at maximum intensity. After bleaching, recovery was monitored from the subsequent 95 frames with the excitation switched back to the attenuated intensity. A background scan and bleaching under the same conditions were conducted for FRAP curve correction. The recovery curves were fitted and the diffusion coefficient was obtained as reported previously[30].

Molecular dynamics simulation

All-atom explicit solvent molecular dynamics simulation. The simulations were performed using GROMACS (5.0.7)[31] with the CHARMM 36m force field[32,33] and the recommended TIP3P water model, the structure of which was constrained using the SETTLE algorithm[34,35]. RHP1 sequences (DP = 40) were created by varying the random seed with fixed component ratios. The initially extended RHP1 chains were first equilibrated in vacuum for 0.1 ns until they collapsed. Each RHP1 was then embedded at the centre of a bilayer with 120 POPC lipids. 7,200 water molecules were then added with approximately 0.1 M KCl. The initial structure (Extended Data Fig. 2d) was created using Packmol[36]. This process mimics the experimental procedures used to promote RHP1 insertion by premixing RHP1 and lipid and the freeze–thawing process used to create transient pores on the membranes.

The potential energy was minimized, followed by equilibration using the NPT ensemble (constant particle number, pressure and temperature). To relax the lipid bilayer, the *Z*-dimensional coordinates of all of the POPC atoms were constrained for 0.1 ns using a force constant of 1,000 kJ mol^{-1} nm^{-2}, followed by an exclusive constraint of the POPC phosphorous atoms for 0.1 ns. The equilibration simulation was performed at 310.15 K for 300 ns. In the production simulation at

298.15 K, the Nose–Hoover algorithm was employed to separately couple the temperatures of the polymer–POPC complex and the others. Semiisotropic pressure coupling ($P_{XY} = P_Z = 1$ bar; time constant, 5.0 ps; compressibility 4.5×10^{-5} bar^{-1}) was applied via the Parrinello–Rahman algorithm[37]. The neighbour search was performed up to 1.2 nm. The short-range Coulomb interactions were calculated up to 1.2 nm using the particle mesh Ewald method[38,39] for the long-range electrostatic interactions. The Lennard–Jones 12–6 potential was switched off from 0.8 nm to 1.2 nm with the force switching method. These parameters have been recommended for the accurate reproduction of the original CHARMM simulations on lipid membrane systems[40]. A simulation integration time step of 2 fs was employed, with all the hydrogen-involved covalent bond lengths constrained. Each production simulation was performed for 500 ns, with the final configuration presented in Extended Data Fig. 2d. The convergence of the simulations was confirmed by the area per lipid and the polymer–lipid interactions (Extended Data Fig. 2e). The HBCs were visualized using the package VMD[41].

Nine parallel simulations were performed with nine different polymer chains (Extended Data Fig. 2a). Five chains were found to span the bilayer roughly symmetrically (Fig. 1e, Extended Data Fig. 2b). The other four chains were found near the membrane–water interfaces. We believe that multiple factors influence the insertion of RHP1, including the overall monomer distributions along the RHP1 chain, the chain length and the transient RHP1 conformations. The sequence analysis showed that the average number of hydrophobic segments per RHP1 chain increases with longer chain lengths. For relatively short RHP1 chains (DP = 40; green line in Extended Data Fig. 6d), it is not surprising that not all of the nine different chains span through the lipid bilayer in these simulations.

Proton transport on LUVs

Transport on LUVs. With excitation at both 460 nm and 415 nm, the ratio of pyranine emission (I) at 510 nm (I_{ex460}/I_{ex415}) was monitored to probe the inner pH change over time. The tested fluorescence ratios of pyranine in the lumen of the LUVs were linearly proportional to pH values in the range 6.4–7.4. Purified LUV solutions were diluted ten times into the 1-ml sample with an outside buffer (10 mM HEPES-KOH pH 7.7, 100 mM KCl). Proper amounts of the RHP solutions at 0.1 or 1.0 mg ml^{-1} were added into the LUV solutions with continuous stirring. After incubation for 5 min, I_{460}/I_{415} was monitored over 500 s. 3 µl of 10 µM Vln in ethanol was added at 100 s to trigger the proton flux. To evaluate the ion selectivity of RHP1, purified solutions of RHP1-containing LUV were diluted ten times in 1 ml of outside buffers containing one type of alkali metal cation (10 mM HEPES-KOH pH 7.7, 100 mM MCl, M = Li, Na, K, Rb, Cs). Without Vln, changes in the inner pH were recorded over 200 s. The fluorescent measurements were performed on a Perkin Elmer KS55 fluorescence spectrometer and a micro-cuvette under constant stirring.

Transport investigation on stopped-flow fluorimeter

Proton transport rate. RHP1 and the lipids were premixed in chloroform before LUV preparation. Purified RHP1-containing LUV solutions were diluted 33 times with the inner solution and mixed with 30 nM Vln. 15 µl of the RHP1–Vln–LUV mixture in the inner solution was rapidly exposed to 15 µl of the outside solution using the AutoSF 120 stopped-flow fluorimeter (KinTek) at a flow rate of 8 ml min^{-1}. With excitation at 460 nm, the emission of the encapsulated pyranine was monitored as a function of time using a 520 nm/20 nm fluorescence filter. To evaluate the isotope effect of the proton transport, a similar proton transport experiment was performed by substituting D_2O for H_2O, including in the LUV preparation and purification processes. A deuterated base was used to titrate the pD. The pD value was measured using a glass pH electrode and corrected by adding 0.4 to the pH readout. The isotope effect value was obtained from the ratio of the conductance in the H_2O assay to that in the D_2O-based assay. The initial pH change rate was fitted to obtain the conductance of the channels or RHP1 and the permeability of membranes, using equations reported previously[22].

Water transport. LUVs were prepared in a similar way as for the proton transport assay, using 20 mM HEPES-KOH, pH 7.5, 100 mM KCl containing 10 mM carboxyfluorescein. The LUVs were purified using the same buffer without the dye. Using stopped-flow fluorescence at 7 °C, the channel- or RHP1-containing LUVs in this inner solution were rapidly exposed to an outside solution (20 mM HEPES-KOH, pH 7.5, 100 mM KCl, 100 mM sucrose) and subjected to a fluorescence measurement with excitation at 490 nm and the 520 nm/20 nm filter. The relative volume was linearly related to the relative fluorescence at the present conditions. The fluorescence quenching curves for the shrinking liposomes were fitted with a single-exponential function. The osmotic water permeability was calculated as described previously[42,43].

RHP conformation in liposomes

SANS investigation. SANS measurements were conducted at the GP-SANS instrument of the High Flux Isotope Reactor at Oak Ridge National Laboratory[44] and at the NG7 instrument of the National Institute of Standards and Technology. The sample-to-detector distances were set to generate a range of momentum transfer, q, from 0.004 to 0.6 Å^{-1}, where $q = 4\pi\sin\theta/\lambda$, where 2θ denotes the scattering angle from the incident beam. The wavelength λ is 5 Å for NG7 and 4.75 Å for GP-SANS. Small unilamellar vesicles (SUVs) composed of d54-DMPC were prepared as described above, using 86 vol% D_2O in H_2O and a 100-nm membrane. The SLD of the 86% D_2O theoretically matches the average SLD of d54-DMPC. Although 100% contrast match cannot be achieved owing to the deuterated lipid tail and H-based lipid head, using 86% D_2O dampens the SANS signals from the liposome. The SANS measurements were performed at 37 °C in 2-mm-path-length quartz cells using solutions of the SUVs. The 2D scattering results were azimuthally averaged to produce one-dimensional SANS curves. The data were modelled using a three-shell vesicle model for the RHP1-free SUVs and a sphere model for the free RHP1. A sum of these two models was used to fit the SANS profiles of the SUV–RHP1 samples. All data fitting was performed using the SASView program (version 4.2.0, http://www.sasview.org/).

Sequence analyses

HLB. The HLB value was used to evaluate the solubility of monomer side-chains through group contribution theory. Using the equation $\mathrm{HLB} = 7 + \sum_i n_i \mathrm{HLB}_i$, where n_i is the number of the ith chemical group in the molecule with corresponding value HLB_i. The HLB value for each monomer side chain was estimated as: HLB(MMA) = 8.45, HLB(EHMA) = 5.12, HLB(OEGMA) = 11.4 and HLB(SPMA) = 18.5. A lower HLB value denotes higher hydrophobicity and a higher value means greater hydrophilicity. A Python program was created to continuously calculate the average segmental HLB values for a window sliding from the alpha to the omega ends of the simulated RHP1 chains. The window advanced by one monomer each time. We used a span containing odd numbers of monomers, and assigned the average HLB value of that span to its middle monomer. For instance, for the window of 11 monomers shown in Fig. 1b, the average HLB value of the span from monomer 1 to monomer 11 was calculated and plotted at monomer position 6. The next HLB value, corresponding to the span from monomer 2 to monomer 12, was plotted at monomer position 7.

Hydrophobic segments. A Python program was used to analyse the length and number of hydrophobic segments containing one OEGMA in the RHPs. The hydrophobic embedded OEGMA was set to be two or more monomers away from the ends of the segments. The number of uninterrupted hydrophobic monomers on both sides of an OEGMA was counted.

Segment prediction

HMM prediction. The HMM has a successful history in describing latent statistical structures and incorporating various constraints, despite its limitation in duration modelling. Given an observable output sequence of length T (for example, $y = (y_1, y_2, \ldots, y_T)$), the basic goal for an HMM[45] is to infer the corresponding hidden state sequence of the same length, $q = (q_1, q_2, \ldots, q_T)$. In our study, y_i denotes the monomers and q_i represents the states defined above: state 1, state 2 and state 3. The HMM parameters include:

1. The initial probabilities of each state, $\Delta = \{\pi_i\}$, where $\pi_i = P(q_1 = i)$ and i is the index for the three different hidden states (that is, $i = 1, 2, 3$).

2. The transition probability matrix $A = \{a_{ij}\}$, where $a_{ij} = P(q_{t+1} = j \,|\, q_t = i)$, that is, the probability of next step being in state j ($j = 1, 2, 3$) if currently it is in state i.

3. The emission probability matrix $B = \{b_{ij}\}$ is defined as $b_{ij} = P(y_t = j | q_t = i)$, where j is the index for the four different types of monomers (that is, $j = 1, 2, 3, 4$).

Then the joint probability of states (that is, $q = (q_1, q_2, \ldots, q_T)$) and observed outputs (that is, $y = (y_1, y_2, \ldots, y_T)$) can be written as:

$$P(y, q) = \pi_{q_1} b_{q_1 y_1} \left(\prod_{t=2}^{T-1} b_{q_t y_t} a_{q_t q_{t+1}} \right) b_{q_T y_T}$$

The Baum–Welch algorithm[46] is typically used to find the unknown parameters Δ, A and B, and the Viterbi algorithm[47] is used to determine the optimal state path for each observed sequence.

The unobserved states represent the three possible regions on RHP1s that prefer water, the lipid–water interface, or the core of lipid bilayers, and states 1, 2 and 3 denote the corresponding hidden states. We implemented a third-order HMM in which (i) the transitions between states were modelled by a third-order Markov chain and (ii) third-order emission probabilities were used to characterize the monomer preferences in different RHP1 regions and to model the dependencies between neighbouring monomers—that is, the probability of generating a monomer at a position would depend not only on the hidden state at that position, but also on the states at the previous two positions. Additionally, we incorporated the following rules in the HMM to better approximate reality and facilitate parameter inference (Extended Data Fig. 7b):

1. Despite the ambiguous criteria for defining a state-3 segment, we began with segments full of hydrophobic monomers and hydrophobic segments containing one OEGMA.

2. To differentiate between OEGMAs included in a state-3 segment and those adjacent to such segments, the OEGMA included was set to be at least two monomers away from the ends of the segments.

3. Owing to the high polarity, SPMA was fixed to appear only in state 1 in the current setup.

4. Hydrophobic segments (states 2 and 3) end with hydrophilic monomers (OEGMA and/or SPMA). Those boundary monomers, if not included in an extended state-3 segment (that is, the OEGMA within a long hydrophobic segment), were considered to be in state 1. We note that when a length constraint for state 3 has to be incorporated (as shown in Extended Data Fig. 7b), the design of the HMM would be more challenging and the HMM might not perfectly represent the observed data distribution.

Proton transport on SLBs

Transport on an SLB. SUVs were prepared in a similar way to the LUVs, except for using a 100-nm extrusion membrane. 150 µl of 2.5 mg ml^{-1} SUVs was deposited on a clean Pd contact and incubated overnight with mild agitation to facilitate formation of the SLB. Unfused SUVs were washed off using the liposome buffer. The SLB formation was verified by liquid atomic force microscopy characterization (Extended Data Fig. 8a). The fresh SLB was covered with 150 µl of 0.05 mg ml^{-1} RHP1 solution for 30 min. The RHP1–SLB solution was subjected to a liposome buffer rinse to remove free polymers. Current–voltage and cyclic voltammetry measurements were performed on a home-built system (Extended Data Fig. 8b), and the membrane permeability was fitted using a model described previously[19].

Data availability

The data supporting the findings of this study, as well as descriptions of the methodologies used in the code, are available within the article and the Extended Data items. For reproduction purposes, the raw data used to generate the figures and input scripts used to produce the simulations are available from the Dryad Digital Repository (https://doi.org/10.6078/D1VX0B).

Code availability

All custom scripts (sequence analyses and HMM) are available from the corresponding author upon request.

28. Stewart, J. C. M. Colorimetric determination of phospholipids with ammonium ferrothiocyanate. *Anal. Biochem*. **104**, 10–14 (1980).
29. Lira, R. B., Steinkuhler, J., Knorr, R. L., Dimova, R. & Riske, K. A. Posing for a picture: vesicle immobilization in agarose gel. *Sci. Rep*. **6**, 25254 (2016).
30. Manneville, J. B. et al. COPI coat assembly occurs on liquid-disordered domains and the associated membrane deformations are limited by membrane tension. *Proc. Natl Acad. Sci. USA* **105**, 16946–16951 (2008).
31. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput*. **4**, 435–447 (2008).
32. Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
33. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
34. MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
35. Miyamoto, S. & Kollman, P. A. Settle: an analytical version of the shake and rattle algorithm for rigid water models. *J. Comput. Chem*. **13**, 952–962 (1992).
36. Martínez, L., Andrade, R., Birgin, E. G. & Martinez, J. M. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem*. **30**, 2157–2164 (2009).
37. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys*. **52**, 7182–7190 (1981).
38. Darden, T., York, D. & Pedersen, L. Particle mesh ewald: an *N*·log(*N*) method for Ewald sums in large systems. *J. Chem. Phys*. **98**, 10089–10092 (1993).
39. Essmann, U. et al. A smooth particle mesh Ewald method. *J. Chem. Phys*. **103**, 8577–8593 (1995).
40. 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).
41. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph. Model*. **14**, 33–38 (1996).
42. Zeidel, M. L., Ambudkar, S. V., Smith, B. L. & Agre, P. Reconstitution of functional water channels in liposomes containing purified red-cell CHIP28 protein. *Biochemistry* **31**, 7436–7440 (1992).
43. Shen, Y. X. et al. Highly permeable artificial water channels that can self-assemble into two-dimensional arrays. *Proc. Natl Acad. Sci. USA* **112**, 9810–9815 (2015).
44. Heller, W. T. et al. The suite of small-angle neutron scattering instruments at Oak Ridge National Laboratory. *J. Appl. Cryst*. **51**, 242–248 (2018).
45. Rabiner, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **77**, 257–286 (1989).
46. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. A* **39**, 1–22 (1977).
47. Forney, C. D. The Viterbi algorithm. *Proc. IEEE* **61**, 268–278 (1973).

Acknowledgements This work was supported by the US Department of Defense, Army Research Office, under contract W911NF-13-1-0232 and the National Science Foundation under contract DMR-183696. M.O.d.l.C. and B.Q. acknowledge support through grant DE-FG02-08ER46539 from the Department of Energy (DOE) Basic Energy Science Office and the Center for Computation and Theory of Soft Materials, as well as computational support by the Sherman Fairchild Foundation. Z. H. and M. R. acknowledge support from the Air Force Office of Sponsored Research Award FA9550-15-1-0273. A part of this research used resources at the High Flux Isotope Laboratory, a DOE Office of Science User Facility operated by the Oak Ridge National Laboratory. We acknowledge the support of the National Institute of Standards and Technology, US Department of Commerce, in providing the neutron research facilities used in this work. We thank L. He and Y. Liu for help in the SANS studies. Scattering studies at the Advanced Light Source and RHP characterization at the Molecular Foundry were supported by the Office of Science, Office of Basic Energy

Sciences of the US DOE under contract DE-AC02-05CH11231. We thank A. A. A. Smith for help in polymer syntheses; Y. W. Qian for characterizing the liposome sizes; A. Martin and E. M. López-Alfonzo for help in stopped-flow experiments. We thank CoC-NMR for help with RHP characterization; instruments at CoC-NMR are supported in part by NIH S10OD024998.

Author contributions T.J. made the first experimental observation of RHP-based proton transport. T.X. directed the project development. T.J. and T.X. designed the experiments and data analysis. T.J. performed imaging, FRAP and transport studies on liposomes. A.H. synthesized and characterized RHPs. M.E. and Z.R. performed the RHP segmental analysis. Z.H. and M.R. performed transport studies on SLBs. B.Q. and M.O.d.l.C. performed all-atom simulations and data analysis. Y.Z. and H.H performed HMM studies. T.J. and A.D.C. prepared and characterized liposomes. T.J. and W.T.H. collected and analysed SANS data with T.X. All authors participated in the writing of the manuscript.

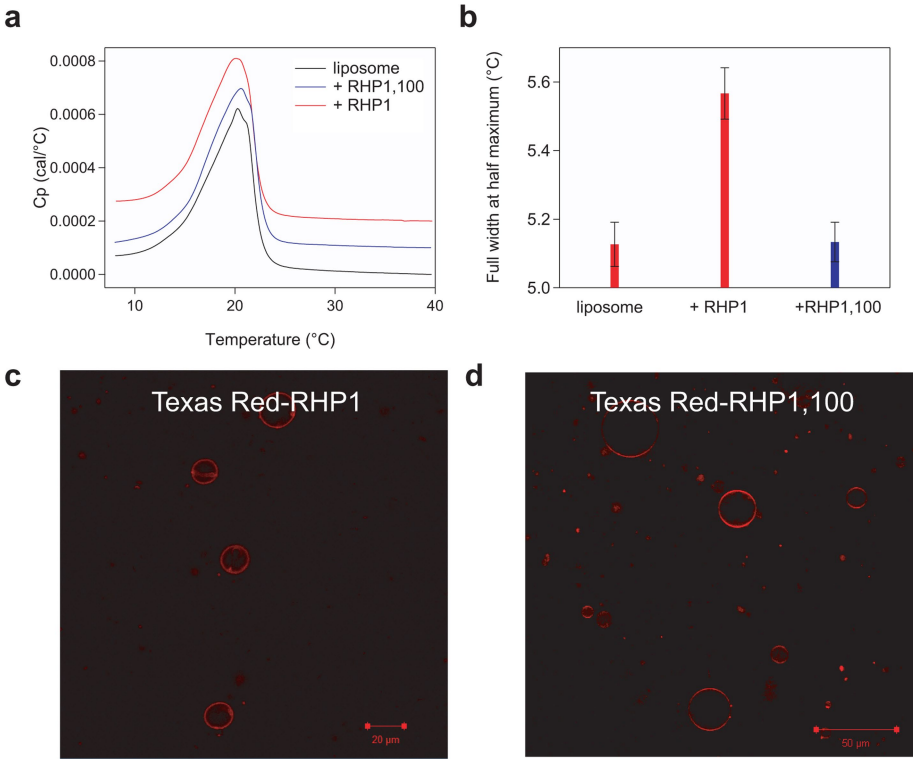
Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1881-0.

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

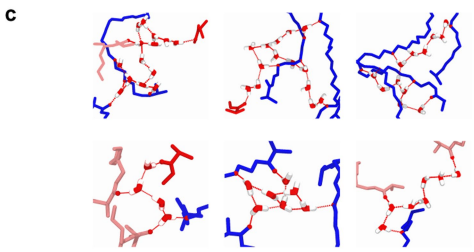
Peer review information *Nature* thanks Junli Hou, Kenichi Kuroda and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

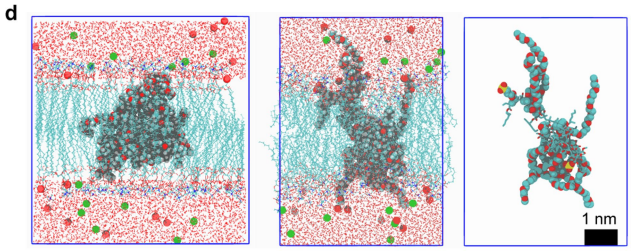


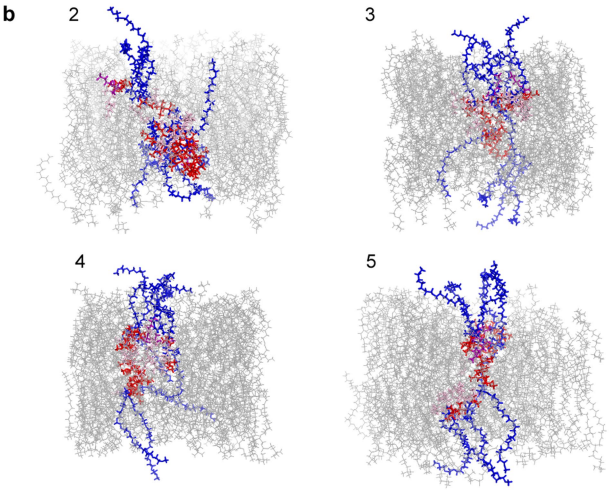
Extended Data Fig. 1 | Interactions between RHPs and liposomes. a, DSC profiles of LUVs and RHP1- and RHP1,100-containing LUVs (POPE:POPG molar ratio, 3:1). LUVs were grown with RHP1 or RHP1,100 with a lipid-to-RHP molar ratio of 50:1 and a lipid concentration of 1 mg mL^{-1}. The background signals of the buffer and the polymers were subtracted. No phase transition was observed for RHP itself. The plots are shifted vertically for clarity by +0.0001 for RHP1,100 and +0.0002 for RHP1. **b**, Full-width at half-maximum of the lipid phase transition peaks shown in **a**. Error bars are 1 s.d. ($n = 3$). **c**, **d**, Representative confocal fluorescence image of POPC GUVs incubated with Texas Red-labelled RHP1 (**c**) and Texas Red-labelled RHP1,100 (**d**).

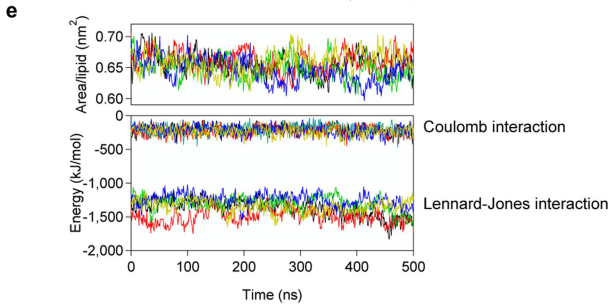
a

RHP1 chains	monomer sequences
1	11114114211231121111212131111323323233 22
2	31412212331123322213121213311111111121 14
3	13121234123233143112112111311121311112 22
4	13231222313113311111221312111113121212 44
5	31111111211212123422114111123121123233 33
6	41332213233212222121111131211311111114 311
7	31321231221121322242321143311311111111 111
8	12113111232311134121321222112331121141 13
9	11121112211311113122311112212312312343 34

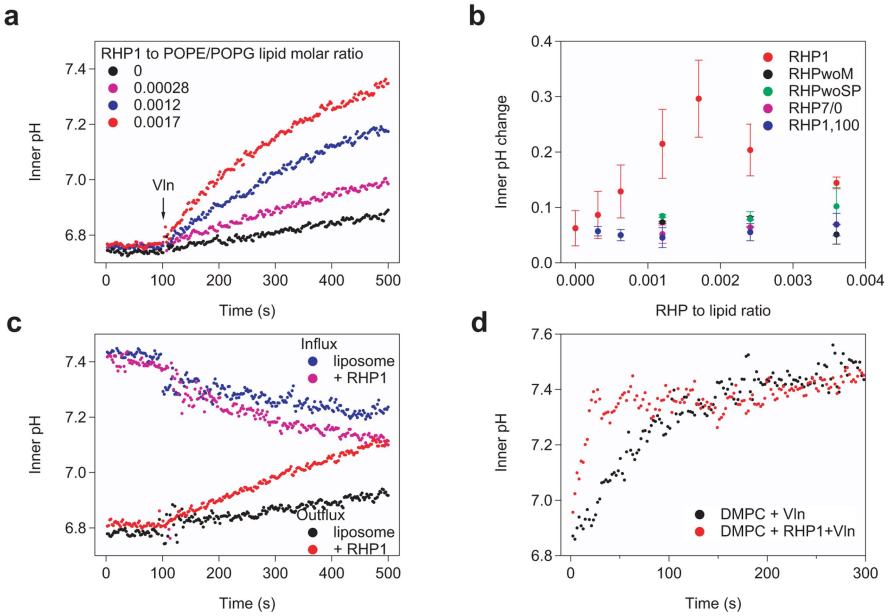




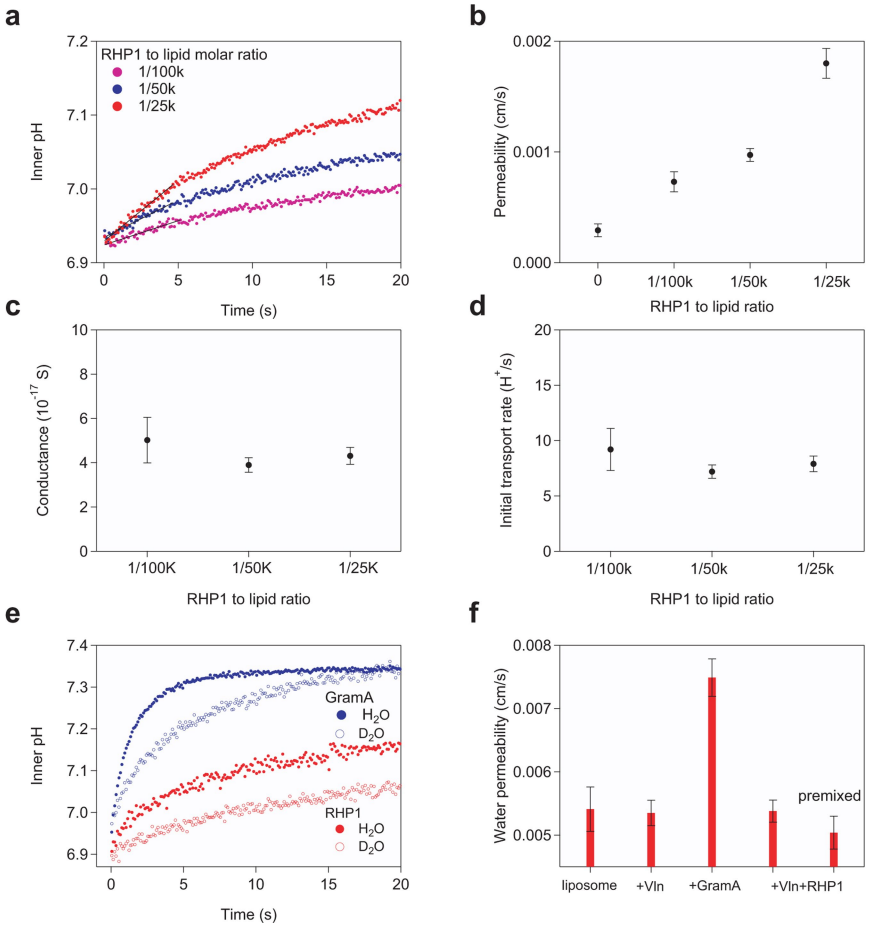




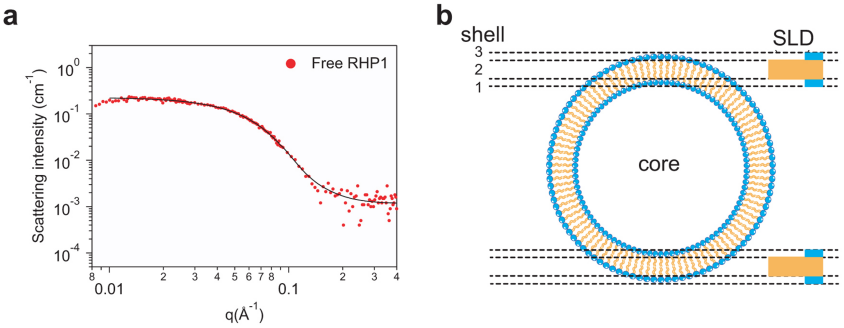
Extended Data Fig. 2 | All-atom molecular dynamics simulations. **a**, Sequences of the nine RHP1 chains used in the simulations. **b**, Snapshot of the 2nd–5th RHP1 chains in the POPC lipid bilayer in the simulations. MMA, EHMA, OEGMA and SPMA are coloured red, pink, blue and purple, respectively. Lipids are shown in grey colour. **c**, Representative HBCs in the transmembrane regions of RHP1s. The oxygen and hydrogen atoms of water molecules are coloured red and white, respectively. Hydrogen bonds are shown with red dashed lines. Lipids are omitted for clarity. **d**, Snapshots of the initial (left) and final (right) RHP1 conformations in the simulations. The RHP1 chain is highlighted, along with K^+ (orange) and Cl^- (green) ions. **e**, Convergence of the simulations. Top, area per lipid as a function of the simulation time. Bottom, Coulomb and Lennard–Jones interactions between POPC lipids and the 1st–5th RHP1s as a function of simulation time. Different colours denote the results from five parallel molecular dynamics simulations.



Extended Data Fig. 3 | Transmembrane proton transport. a, pH gradient-driven proton flux through RHP1 in the LUVs of POPE/POPG (molar ratio 3:1). **b**, Inner pH changes in POPC LUVs incubated with RHP1 and RHP variants over 200 s after adding Vln. 'RHPwoM', 'RHPwoSP' and 'RHP7/0' represent the RHP without MMA, the RHP without SPMA and the RHP without EHMA, respectively. **c**, pH gradient-driven proton flux in the influx and outflux directions. RHP1 was added to the solutions of preformed LUVs. **d**, pH gradient-driven proton flux through RHP1 in DMPC SUVs at 37 °C. RHP1-containing SUVs were prepared as indicated in the Methods section 'SANS investigation' using H_2O-based buffers, with an RHP1:DMPC ratio of 0.001.



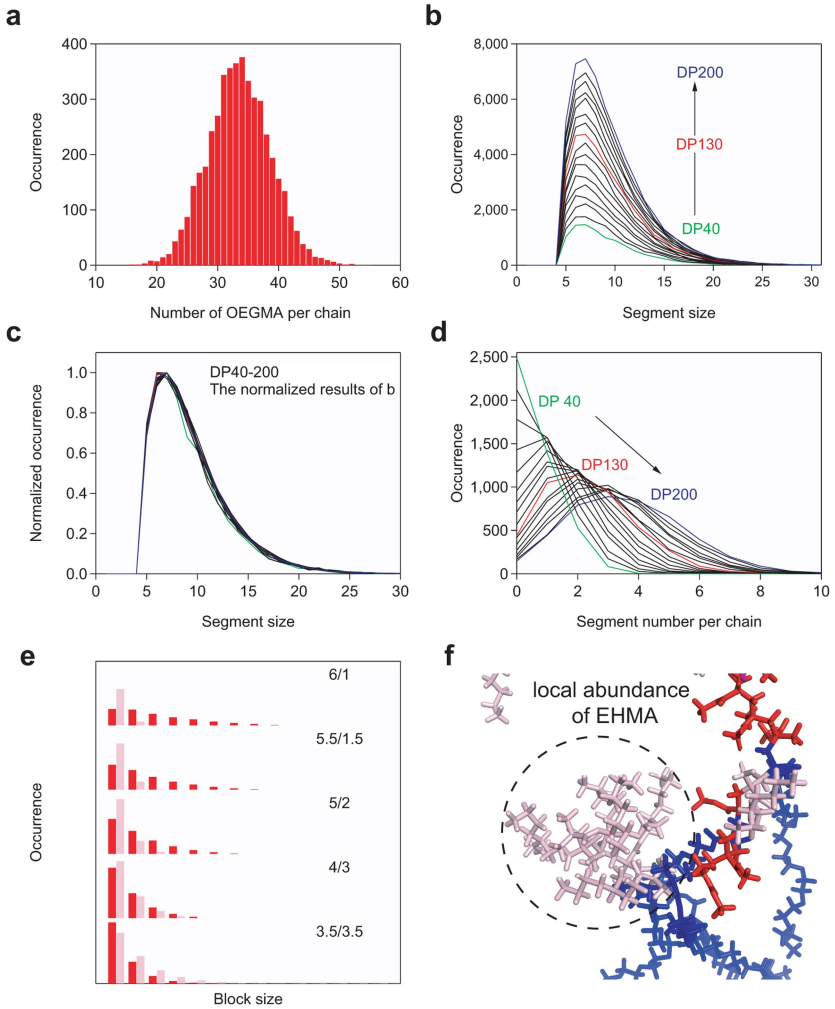
Extended Data Fig. 4 | Proton and water transport rates. a, Representative stopped-flow fluorescence traces of RHP1-containing POPC LUVs after rapid exposure to a buffer of a higher pH value. The initial and final equilibrated pH values are 6.85 and 7.35, respectively. Vln was added before the measurements. **b**, Calculated permeability values of RHP1-containing LUVs at various RHP1-to-lipid ratios. Error bars are 1 s.d. ($n = 3$). **c**, **d**, Conductance (**c**) and initial proton transport rate (**d**) per RHP1 chain at various RHP1-to-lipid ratios. Error bars are 1 s.d. ($n = 3$). **e**, Stopped-flow fluorescence traces of RHP1-containing LUVs after rapid exposure to a buffer of a higher pH or pD value. The initial and final equilibrated pH values are 6.85 and 7.35, respectively, for the H_2O-based assay. The initial and final pD values are 7.25 and 7.75, respectively, for the D_2O-based assay. Vln was added to the RHP1-containing LUV solutions, except for the GramA-containing LUVs. **f**, Osmotic water permeability of Vln, GramA and RHP1+Vln in POPC LUVs at 7 °C. Vln (K^+ channel) and GramA (proton/cations/water channel) are the negative and positive controls, respectively. GramA, Vln and RHP1 were added to the solutions of preformed LUVs. 'Premixed' denotes premixing of RHP1 and the lipid before the preparation of LUVs. Error bars are 1 s.d. ($n = 7$).



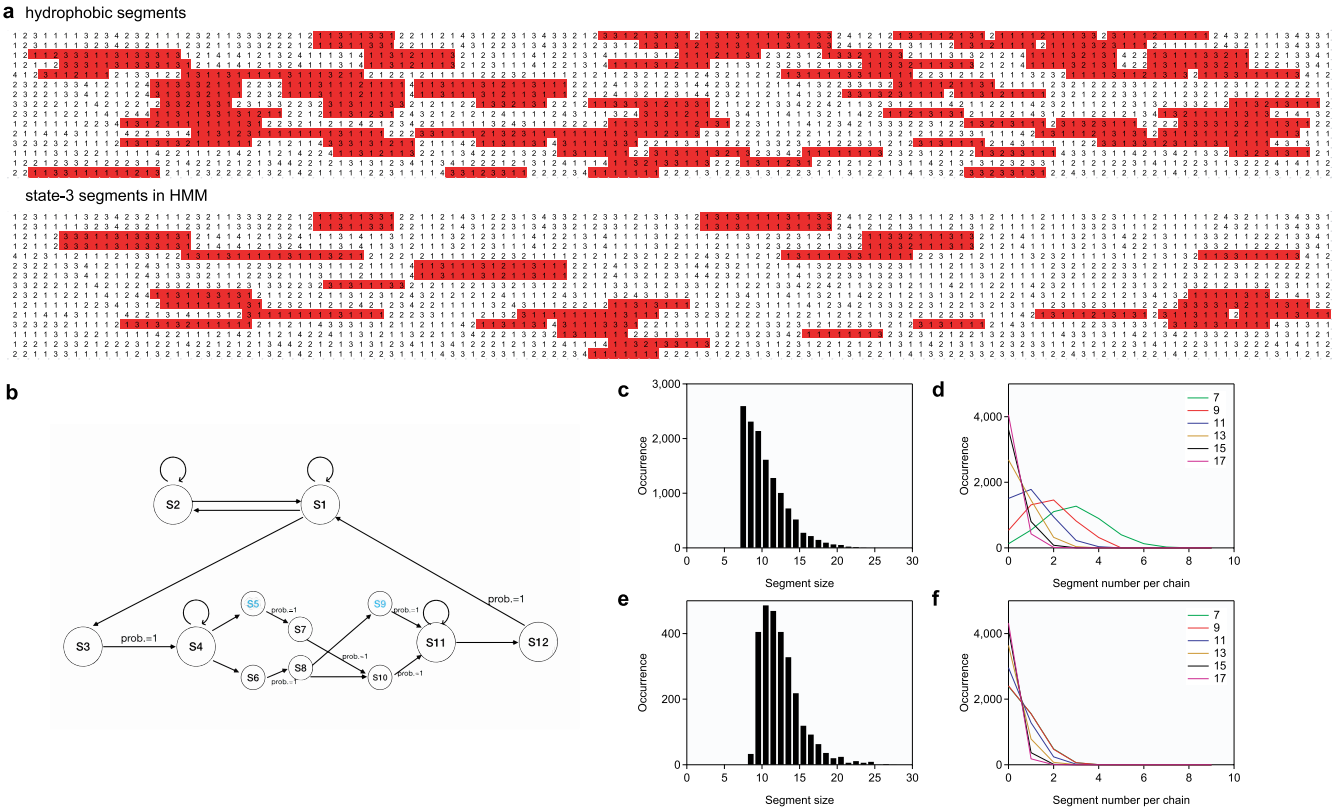
c

	liposome only	RHP1 only	RHP1/lipid 0.0005	0.001	0.002
liposome core radius (Å)	500f		500f	500f	500f
shell 1 thickness (Å)	9.32		9.92	12.1	10.5
shell 2 thickness (Å)	29.3		28.3	27.4	28.4
shell 3 thickness (Å)	7.24		6.35	3.85	5.70
RHP1 radius (Å)		23.3	17.4	26.8	22.5
RMS deviation of RHP1 radius (Å)		9.79	8.87	14.2	11.0

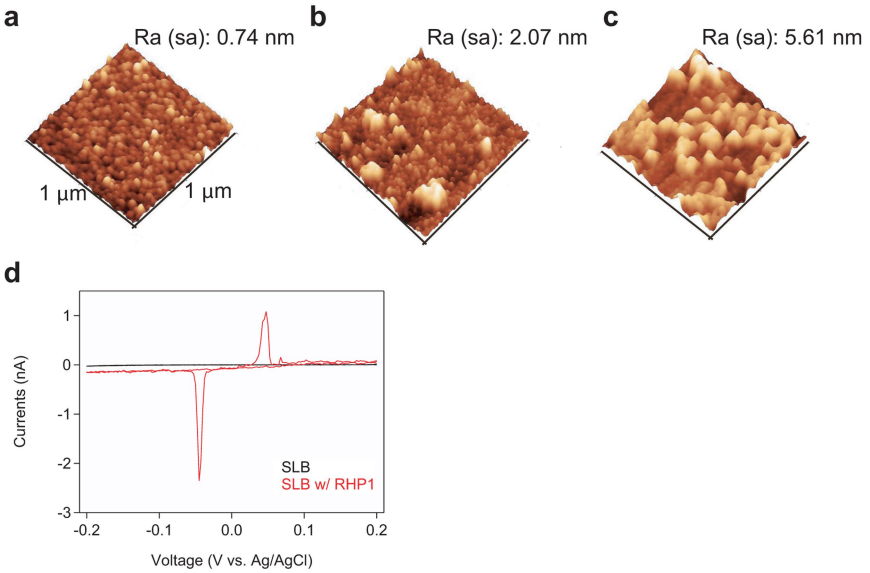
Extended Data Fig. 5 | SANS investigations. a, SANS profiles of free RHP1 in 86% D_2O. The solid line denotes a fit using a polydisperse sphere model for RHP1. The signal of the solvent has been subtracted in the presented curves. **b**, Schematic illustration of the three-shell vesicle model for the SUV, showing the regions of the liposome core (core), inner bilayer headgroup (shell 1), lipid bilayer core (shell 2) and outer bilayer headgroup (shell 3). **c**, Structural parameters obtained by fitting the SANS data with the three-shell vesicle model for the liposome and the polydisperse sphere model for RHP1. Parameters that were fixed during the fitting are denoted with 'f'; RMS represents root mean square. SLDs that were fixed were determined from the known sample composition, including when RHP1 was present in the bilayer, because the concentration was too low to impact the SLD value of shells significantly. The SLD values used for the liposome core, for shells 1, 2, 3 and for the solvent are 5.41, 1.88, 6.84, 1.88 and 5.41×10^{-6} Å^{-2}, respectively. The SLD value for RHP1 is 1.4×10^{-6} Å^{-2} in the sphere model.



Extended Data Fig. 6 | Sequence analyses of RHPs. a, Statistical distribution of number of OEGMA monomers per RHP1 chains in the 4,500 simulated RHP1 sequences (DP = 130). This distribution reflects the compositional dispersity in RHP1 chains. **b**, Statistical size distribution of hydrophobic segments with one OEGMA in RHP1s of various chain lengths ranging from DP = 40 to 200. The number of chains (4,500) was kept constant for each DP analysis. **c**, Normalized profiles of the size distribution in **b**. **d**, Number of hydrophobic segments per chain at various chain lengths. The counted hydrophobic segments here contain one OEGMA with segment size ≥11. **e**, Statistical size distributions of blocks of EHMA (pink) and MMA (red) within hydrophobic segments containing one OEGMA in RHP variants. Results are shown for hydrophobic segments with segment size ≥11. **f**, Schematic illustration of five EHMA monomers distributed at the end of an inserted RHP1 chain in a molecular dynamics simulation. EHMA, MMA and OEGMA are coloured pink, red and blue, respectively. Lipids are omitted for clarity.



Extended Data Fig. 7 | HMM prediction. a, Comparison between state-3 segments (size ≥7; top) assigned using the method described in Methods section 'Hydrophobic segments' and those predicted by HMM (bottom). '1', '2', '3' and '4' denote MMA, OEGMA, EHMA and SPMA, respectively. **b**, Transition diagram with a state-3 segment of size ≥7. States 1 and 2 are denoted as S1 and S2, respectively. State 3 here is broken into 10 states to comply with the rules outlined in Methods section 'HMM prediction'. S5 and S9 are the special states in a state-3 segment that must emit OEGMA with probability 1 (denoted as 'prob. = 1'). **c**, Statistical size distribution of the state-3 segments predicted by HMM in 4,500 RHP1 sequences. **d**, Number of the predicted state-3 segments per polymer chain. **e**, Statistical size distribution of predicted state-3 segments containing one OEGMA. **f**, Number of predicted state-3 segments containing one OEGMA per polymer chain. 4,375 out of 4,500 RHP1 chains contain at least one state-3 segment.



Extended Data Fig. 8 | Proton transport on SLBs. a, Liquid atomic force microscopy images of a Pd surface covered with SLBs, showing a surface roughness of about 0.74 nm. **b**, Image for SLB+RHP1 (RHP1-to-lipid molar ratio, 0.001), showing a roughness of about 2.07 nm. **c**, Surface covered with SLB+RHP1 (RHP1-to-lipid molar ratio, 0.01), showing a roughness of about 5.61 nm. **d**, Current–voltage sweep of the protonic device. The Pd/SLB contact has a current of <0.01 nA across the entire voltage range, confirming the high polarization resistance of the SLB (black trace). With RHP1 incorporated, an oxidation peak is observed at about 48 mV, indicating H^+ flow across the SLB; a reduction peak is observed at about −45 mV, indicating conduction of H across the SLB.

Extended Data Table 1 | Physical-chemical properties of RHPs

	MMA	EHMA	OEGMA	SPMA	Mn(kg/mol)*	Mn(kg/mol)†	Đ†
RHP1	50 (50.8)	20 (21.5)	25 (24.1)	5 (3.7)	30.6	29.0	1.3
RHP70+0	70 (72.1)	0 (0)	25 (23.1)	5 (4.8)	18.0	20.2	1.4
RHP60+10	60 (61.9)	10 (10.0)	25 (23.5)	5 (4.6)	18.5	21.0	1.3
RHP55+15	55 (51.9)	15 (12.2)	25 (33.3)	5 (2.5)	28.0	33.6	1.6
RHP40+30	40 (40.4)	30 (30.3)	25 (24.3)	5 (5.0)	26.5	22.0	1.3
RHP35+35	35 (36.1)	35 (36.4)	25 (23.8)	5 (3.8)	40.0	51.1	1.3
RHPwoSP	53	21	26	0	29.2	44.4	1.3
RHPwoM	0	40	50	10	35.6	45.9	1.2
RHP1-7k	50	20	25	5	7.0	16.3	1.2
RHP1-19k	50	20	25	5	18.7	18.8	1.4
RHP1-40k	50	20	25	5	40.6	40.2	1.3
RHP1,100	50	20	25	5		57.8	1.2

*Determined by NMR.

†Measured by gel permeation chromatography, using PEG standards for calibration.

‡Numbers in parentheses are the monomer compositions determined by ^{13}C NMR (150 MHz).