High-pH structure of EmrE reveals the mechanism of proton-coupled substrate transport

Alexander A. Shcherbakov [1], Peyton J. Spreacker [2], Aurelio J. Dregni [1],
Katherine A. Henzler-Wildman [2] & Mei Hong [1✉]

The homo-dimeric bacterial membrane protein EmrE effluxes polyaromatic cationic substrates in a proton-coupled manner to cause multidrug resistance. We recently determined the structure of substrate-bound EmrE in phospholipid bilayers by measuring hundreds of protein-ligand H^N–F distances for a fluorinated substrate, 4-fluoro-tetraphenylphosphonium (F_4-TPP$^+$), using solid-state NMR. This structure was solved at low pH where one of the two proton-binding Glu14 residues is protonated. Here, to understand how substrate transport depends on pH, we determine the structure of the EmrE-TPP complex at high pH, where both Glu14 residues are deprotonated. The high-pH complex exhibits an elongated and hydrated binding pocket in which the substrate is similarly exposed to the two sides of the membrane. In contrast, the low-pH complex asymmetrically exposes the substrate to one side of the membrane. These pH-dependent EmrE conformations provide detailed insights into the alternating-access model, and suggest that the high-pH conformation may facilitate proton binding in the presence of the substrate, thus accelerating the conformational change of EmrE to export the substrate.

[1] Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139, USA. [2] Department of Biochemistry, University of Wisconsin at Madison, Madison, WI 53706, USA. ✉email: meihong@mit.edu

Multidrug resistance (MDR) is a major public health concern that can undermine years of drug development efforts and result in epidemics of drug-resistant infections[1]. One of the mechanisms by which cells can become resistant to therapeutics is via expression of transmembrane (TM) efflux pump proteins in the small multidrug resistance (SMR) family. These SMR transporters remove cytotoxins from the cytoplasm by coupling the uphill efflux process to the downhill influx of protons across the cytoplasmic membrane (Fig. 1a)[2]. Unlike selective active transport proteins that recognize and transport a single substrate, the SMR proteins efflux a variety of cytotoxic compounds with different shapes, sizes, and chemical properties[3,4]. The relatively small size of these transporters was originally thought to provide a minimal model system for studying secondary active transport[5]. However, due to the sensitivity of these proteins to their environment, their conformational plasticity, and lack of extracellular domains, high-resolution structural information had been limited for many years[6–10]. The best studied SMR transporter, EmrE, is involved in pH and osmotic stress responses, biofilm formation, and resistance to a variety of polyaromatic cations[11–14]. Biochemical and biophysical data have shown that the transport process of EmrE is highly complex. For example, in addition to acting as a proton-coupled antiporter, EmrE can also function as a proton-coupled symporter or uncoupled uniporter under different conditions[15–19]. Elucidating the mechanism of membrane transport by EmrE requires atomic-resolution structural information for multiple states of the protein, as well as dynamics information about the protein and the ligands throughout the transport cycle.

The 110-residue EmrE forms an antiparallel, asymmetric homodimer. Early electron microscopy (EM) data established that EmrE did not possess two-fold symmetry[9,20]. A subsequent 3.8 Å crystal structure showed that the two subunits are oriented in an antiparallel fashion, indicating that the protein has dual membrane topology[8]. This antiparallel conformation was later confirmed by NMR and single-molecule FRET data[21], mutagenesis[22–24], EPR data[25,26], and studies of homologous proteins[10,27]. Solution and solid-state NMR (ssNMR) data showed that the two subunits of the dimer are conformationally asymmetric, exhibiting two sets of chemical shifts, and the TM helices undergo major reorientations as they exchange between the inward- and outward-facing states[6,16,21,28,29]. Crosslinking the antiparallel dimer blocked alternating acces in vitro and ethidium resistance in vivo, demonstrating the functional importance of this structural transition[29]. The proton-binding residue of EmrE is E14, which exhibits pK_a values of 7.2 ± 0.1 and 8.4 ± 0.2 in the A and B subunits of the dimer in both lipid bilayers and bicelles[15,30]. Solution NMR experiments showed that the substrate, tetraphenylphosphonium (TPP$^+$), binds the protein asymmetrically, interacting primarily with one subunit and protecting that E14 from protonation, while E14 in the other subunit remains accessible to protonation with a pK_a of 6.8[15,17]. This proton-binding asymmetry
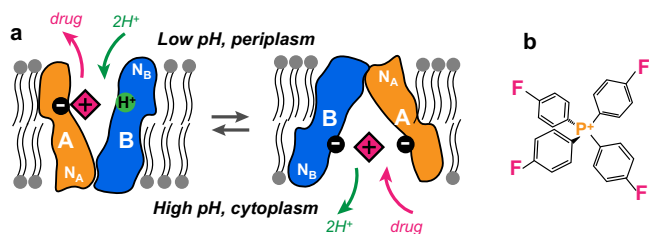
was confirmed by magic-angle-spinning (MAS) NMR data that show that the E14 sidechain carboxyl chemical shifts differ between the two subunits[31]. A 2.3 Å crystal structure of the EmrE homolog, Gdx, was determined in complex with monobodies[10]. This structure confirms the asymmetry of the antiparallel dimer structure and the asymmetric interaction of the substrate with the key glutamate residues in the pore. However, Gdx is a selective guanidinium efflux pump, not a multidrug transporter-like EmrE, and biophysical data is more limited for the Gdx transporter. To fully understand the molecular mechanism of EmrE transport, high-resolution structures are essential. This experimental structure information will enable an integrated analysis of the wealth of data on EmrE dynamics and function, allow assessment of the validity of molecular dynamics simulations[32,33] performed using the backbone-only EmrE crystal structure, and give insight into the similarities and differences between selective and non-selective transporters in the SMR family.

Recently, we discovered a single-point mutant, S64V-EmrE, that retains wild-type substrate-binding affinity but has slower rates of alternating access[7]. At the same time, we developed a ^{1}H–^{19}F REDOR NMR technique to measure distances to the ~2 nm range for structure determination[34–38]. Due to ^{1}H detection under fast MAS, this technique has high spectral sensitivity, thus increasing the throughput of the distance measurement. Exploiting these biochemical and spectroscopic advances, we determined an experimental structure of S64V-EmrE complexed to a fluorinated substrate, F$_4$-TPP$^+$ (Fig. 1b) in DMPC bilayers at pH 5.8[39]. At this acidic pH, one of the two E14 residues is protonated and neutral while the other residue remains deprotonated and anionic. By measuring ~200 protein–ligand distances as well as site-specific protein chemical shifts, we determined the structure of this acidic-pH complex (abbreviated as EmrE-TPP below) to an average pairwise backbone root-mean-square deviation (RMSD) of 1.6 Å. The structure was calculated by docking the ligand to the protein, followed by molecular dynamics simulations that equilibrate the protein in explicitly solvated lipid bilayers, all under experimentally measured distance and torsion angle constraints. This low-pH structure shows that the cationic substrate lies closer to E14A than E14B, consistent with the asymmetric pK_a's. The binding pocket is lined with numerous aromatic residues, including W63, Y60, F44, and Y40. These aromatic sidechains interact with the four ligand phenylene rings to stabilize the substrate, while still leaving sufficient space for the substrate to reorient. While this low-pH structure of EmrE gives a glimpse of the protein–substrate binding geometry, it is not sufficient for revealing the transport mechanism, because alternating access requires the protein to adopt multiple conformations throughout the transport cycle. Experimental data on multiple structural states is required to understand how drug binding and proton binding drive the conformational changes needed to transport the substrates across the membrane.

Here we determine the high-pH structure of the EmrE-TPP complex using the ^{1}H–^{19}F REDOR NMR experiment. We measure ~380 protein–ligand H^N–F distances, which combine with chemical-shift derived torsion angles to enable the calculation of the high-pH structure. We also investigate millisecond-timescale motion of F$_4$-TPP$^+$ at the binding site using 2D ^{19}F–^{19}F exchange NMR. These structural and dynamical results provide information about how changes in the protonation state of the protein drives structural transitions that enable EmrE to transport substrates in a proton-coupled manner.

Fig. 1 Schematic of the EmrE transport function. a Simplified mechanistic model of pH-dependent substrate transport by EmrE. The binding-site structure of the cationic substrate in the dimeric protein depends on the protonation state of the proton-binding residue E14. **b** Structure of the substrate F$_4$-TPP$^+$.

Results

Resonance assignment of the high-pH EmrE-TPP complex. To understand how proton binding and release change the structure of substrate-bound EmrE, we first ascertained the E14 pK_a of

S64V-EmrE in complex with F_4-TPP$^+$. We previously carried out extensive pH titrations of TPP$^+$-bound EmrE in bicelles using solution NMR and found that the WT transporter had a single pK$_a$ of 6.8 ± 0.1[17], while S64V-EmrE had a single pK$_a$ of 7.0 ± 0.1[7]. We verified that this pK$_a$ was not significantly different when S64V-EmrE is bound to F_4-TPP$^+$ by repeating this pH titration. Global analysis of multiple residues yields a pK$_a$ value of 6.9 ± 0.1 (Supplementary Fig. 1). We therefore prepared a sample of S64V-EmrE bound to F_4-TPP$^+$ in DMPC bilayers at pH 8.0 to determine the structure of the unprotonated complex, to compare with the protonated complex previously determined at pH 5.8 in DMPC bilayers.

We first assessed the conformational homogeneity of the high-pH EmrE-TPP complex using 2D ^{1}H–^{15}N correlation (hNH) spectra (Fig. 2a). The spectra show a typical ^{15}N linewidth of ~0.5 ppm, which is narrower than the low-pH sample[39], indicating that the protein is structurally more homogeneous at high pH than at low pH. The ^{15}N and ^{1}H chemical shifts at high pH differ moderately from the low-pH values, which preclude the transfer of the chemical shifts from the low-pH spectra. Thus, we measured eight 3D ^{1}H-detected correlation spectra (Fig. 2b, c and Supplementary Table 1) to independently assign the resonances of the high-pH complex. Among these experiments, the 3D hNcacoNH and HncacoHN spectra are particularly useful for sequence-specific assignment. Based on the peak connectivities, we obtained four backbone chemical shifts (Cα, CO, N^H, ^{H}N) for 77 residues in subunit A and 64 residues in subunit B. In addition, 64 residues in subunit A and 53 residues in subunit B have assigned Cβ chemical shifts. The resulting secondary chemical shifts confirm the presence of four TM helices in the protein (Supplementary Fig. 2). However, the high-pH EmrE exhibits significant chemical shift differences from the low-pH protein in the TM3 and TM1 helices of subunit B (Supplementary Fig. 3). These two TM helices contain the important functional residues W63 and E14, respectively, suggesting that the protein interacts with the substrate differently between high and low pH. Between subunits A and B, TM1 and TM3 helices have larger chemical shift differences compared to TM2 and TM4 helices, similarly indicating the importance of TM1 and TM3 helices for ligand binding (Supplementary Fig. 4).

Measurement of protein-ligand H^N–F distances. The assignment of the H^N and ^{15}N chemical shifts allowed us to measure and resolve protein-substrate H^N–F distances using the ^{1}H-detected and hNH-resolved ^{1}H–^{19}F REDOR experiment[37]. We measured the 2D REDOR spectra at three mixing times, 1.68, 2.53, and 3.78 ms. For each mixing time, a control 2D S_0 spectrum with ^{19}F pulses off and a dephased S spectrum with ^{19}F pulses on were measured. The former exhibits all backbone and sidechain H^N signals (Fig. 3a) whereas the latter shows weaker intensities for amide protons that experience significant ^{1}H–^{19}F dipolar couplings. The difference spectrum (ΔS) between S_0 and S thus manifests only the signals of amide protons that are near the fluorines. At 1.68 ms, we observed signals only from the nearest residues to TPP$^+$, such as E14, Y40, Y60, and W63, whereas at longer mixing times, signals from more remote residues such as G9, G57, and S75 are also detected. No difference signals are observed for residues C-terminal to the TM3 helix and for loop residues.

Fitting the intensity ratios S/S_0 between the REDOR S_0 and S spectra allowed us to extract precise H^N–F distances (Fig. 3b and Supplementary Fig. 5). We observed significant dipolar dephasing for many residues. Among 116 resolved dipolar dephasing curves, 36 show REDOR dephasing that corresponds to distances of less than 9 Å. The shortest distance is found for W63B indole Nε, which is 3.8 Å from the nearest fluorine. In general, TM2 and TM3 residues have some of the closest contact with F_4-TPP$^+$. For example, the S43A H^N is 4.3 Å from the nearest ^{19}F. The distances in the high-pH complex differ substantially from the low-pH values. For example, A10B, S43B, and W63Bε are closer to the ligand fluorines by 1.2–2.5 Å while T18A and W63A backbone amides are further from the ligand fluorines by 0.2–0.3 Å.

Ligand docking and structure calculation. Using rigid-body docking (Supplementary Fig. 6a), we disambiguated the four-fold degeneracy of the fluorines and assigned the measured dipolar couplings to specific protein amide protons (Supplementary Tables 2 and 3). For weak REDOR dephasing that corresponds to distances longer than 10 Å, four constraints were created where the protein H^N atom must be at least 10 Å away from each of the four fluorines. In total, we obtained 387 protein–ligand distance restraints from this docking analysis (Supplementary Table 4). F_4-TPP$^+$ docks to a single location between the two subunits, analogous to the low-pH complex. But the ligand orientation relative to the protein differs substantially from the low-pH structure. At high pH, three of the four phenylene ring planes are approximately parallel to the bilayer normal. In comparison, at low pH, only one phenylene ring is tangential to the bilayer normal whereas three rings lie transverse to the bilayer normal (Fig. 4a, b). Using two lowest-violation docked structures, we refined the protein structure using molecular dynamics simulations under the protein-ligand distance constraints, 148 pairs of (ϕ, ψ) angles and 76 χ$_1$ torsion angles (Table 1 and Supplementary Table 5). The simulations equilibrated by 200 ns to a backbone RMSD of 2.85 ± 0.95 Å for the protein (Supplementary Fig. 6b) and 1.30 ± 0.64 Å for the F_4-TPP$^+$ phosphorous and its four directly bonded carbons (Table 1). Among the four TM helices, TM4 is the furthest away from the ligand: all resolved dipolar couplings correspond to distances of longer than 9 Å (Supplementary Fig. 5). This is consistent with the low-resolution crystal structure, which shows that TM1–TM3 form the substrate-binding pocket whereas the TM4 helices that control dimerization are away from the transport pore[8]. Since TM4 is not well constrained by the measured distances (Supplementary Fig. 6c), when we consider only the TM1–3 helices, the calculated structure has an improved backbone RMSD of 1.97 ± 0.67 Å.

Structural differences between the high-pH and low-pH complex. Interestingly, the high-pH EmrE-TPP complex exhibits numerous structural differences from the low-pH complex (Fig. 4 and Supplementary Fig. 7). First, the ligand is more symmetrically positioned between the two E14 residues at high pH: the distances between the F_4-TPP$^+$ phosphorus (P) and the two E14 Cδ carbons are 6.6 ± 0.7 Å to subunit A and 7.5 ± 0.5 Å to subunit B. In comparison, at low pH the P atom is 1.9 Å closer to the negatively charged E14A than to the neutral E14B (Table 2 and Fig. 4e, f). Thus, the protonation states of the two E14 residues directly impact the substrate position. Second, the inter-subunit proximities and relative orientations of the TM helices have changed between the high-pH and low-pH complexes. At high pH, the two E14-bearing TM1 helices are further away from each other, while the two W63-bearing TM3 helices are more parallel to each other (Fig. 4a, b and Supplementary Fig. 7) compared to the low-pH complex. The E14A CA–E14B CA distance is 16.9 ± 0.6 Å at high pH and shortens to 15.7 ± 0.8 Å at low pH. Overall, the TM helices become more parallel to each other at high pH, creating an elongated binding cavity in which F_4-TPP$^+$ is oriented with three
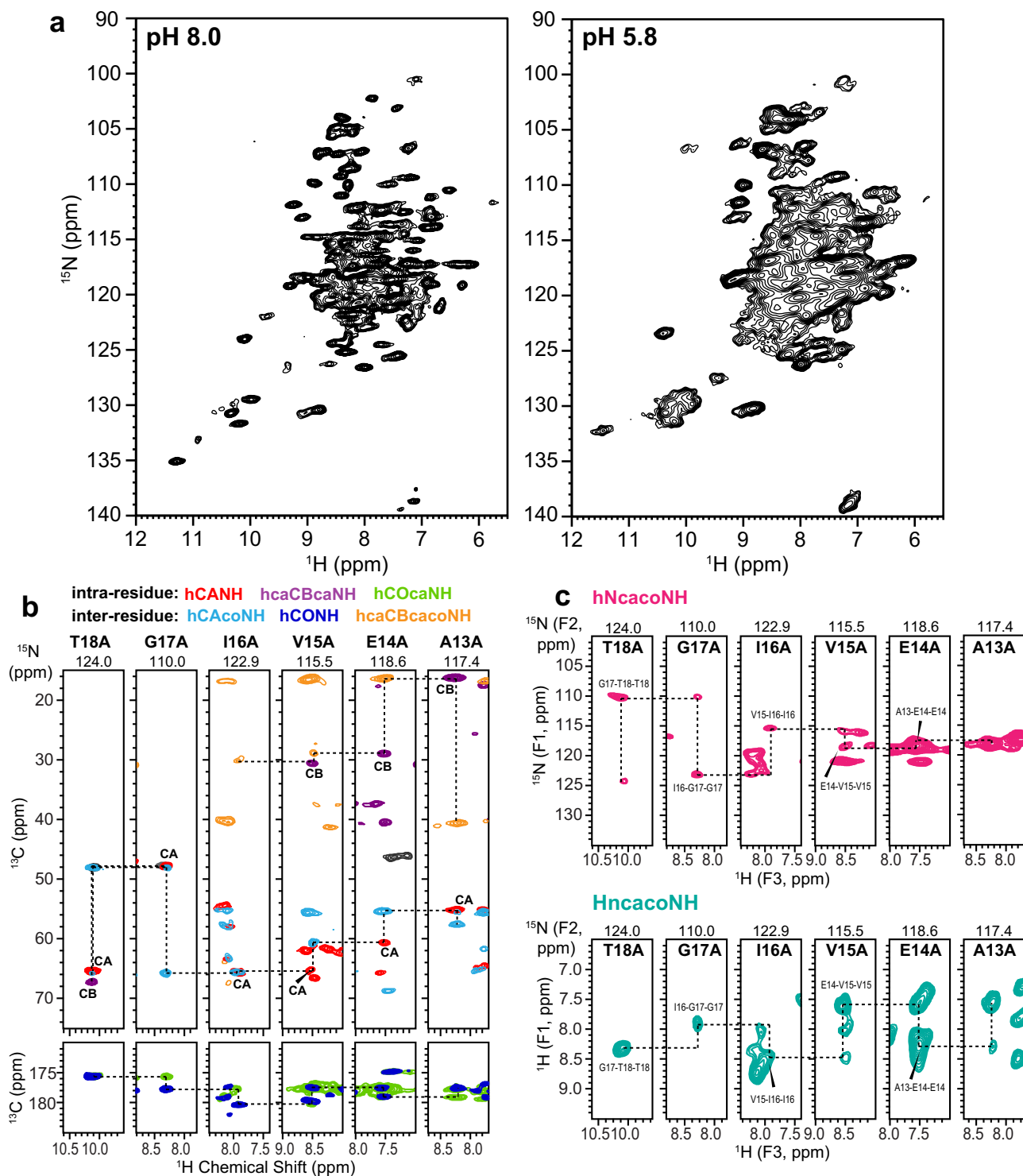
Fig. 2 2D and 3D ¹H-detected MAS correlation spectra of TPP-bound EmrE in DMPC bilayers for resonance assignment. a 2D hNH spectra of the protein at pH 8.0 compared to pH 5.8. The spectral linewidths are narrower at high pH, indicating higher structural homogeneity. **b** Six ¹H–¹³C–¹⁵N 3D correlation spectra to obtain intra-residue and inter-residue peak connectivities. The strips for residues T18 to A13 in subunit A are shown. The aliphatic ¹³C strips overlay four 3D spectra whereas the carbonyl ¹³C strips overlay two 3D spectra. Negative intensities in the hcaCBcaNH and hcaCBcacoNH spectra are shown in gray. Glycine residues show a negative CA peak rather than a positive CB peak. **c** Inter-residue 3D NNH and HNH correlation spectra. All spectra were measured under 55 kHz MAS on a 600 MHz NMR spectrometer.

out of four phenylene rings tangential to the bilayer normal. In comparison, at low pH the TM helices are oriented at very different angles with respect to each other. In particular, the N-terminal end of TM3A approaches the C-terminal end of TM3B (Supplementary Fig. 7b), which closes off water access on

one side of the helical bundle. This closed-on-one-side configuration pushes the ligand towards the opposite end of the helical bundle, where the C-terminal end of TM3A is now splayed open from the N-terminal end of TM3B. The resulting shallow and open binding site at low pH exposes the ligand (Supplementary
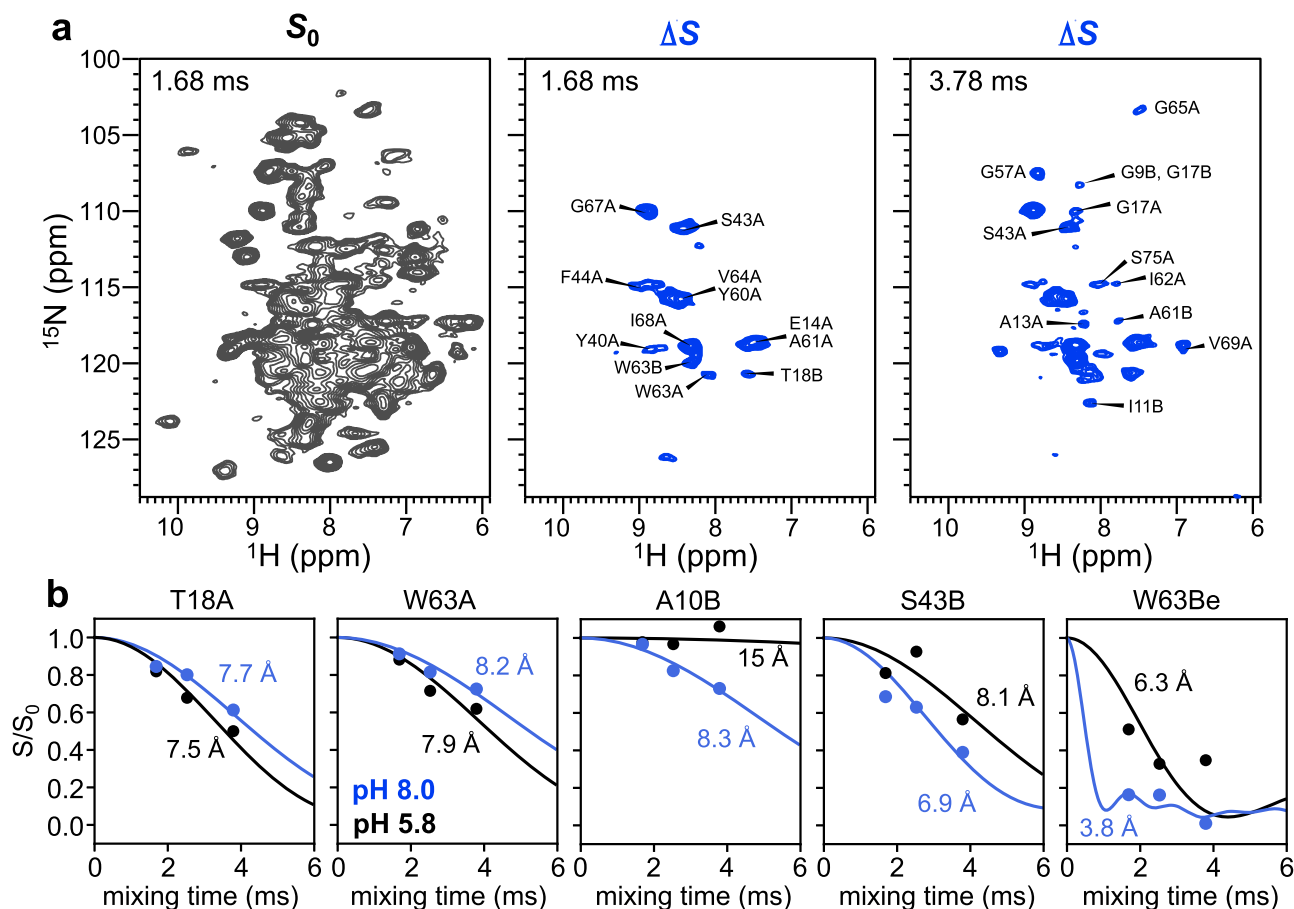
Fig. 3 ^{1}H–^{19}F distance measurements between EmrE H^N and F$_4$-TPP$^+$ using 2D-hNH resolved ^{1}H–^{19}F REDOR. a Representative S_0 (black) and ΔS (blue) spectra for two mixing times, 1.68 and 3.78 ms. Assignment is shown for peaks in the ΔS spectrum. As the mixing time increases, more difference peaks are observed, corresponding to H^N sites that are further from the fluorinated substrate. **b** Representative ^{1}H–^{19}F REDOR S/S_0 dephasing curves with best-fit simulations. ^{1}H–^{19}F distances at pH 8.0 (blue) differ from the pH 5.8 data (black). The indole H^N of W63B has much shorter distances to the substrate fluorines at pH 8.0 than at pH 5.8. REDOR dephasing values are provided as a Source Data file.

Fig. 7b), allowing three of the four phenylene rings to lie transverse to the bilayer normal.

Compared to the low-pH complex, the high-pH EmrE-TPP complex has a longer and more symmetric binding cavity that is not fully closed on either side. Because the binding pocket is populated by protein sidechains, we next investigated which binding site is more spacious by measuring substrate dynamics using ^{19}F NMR. This is a more direct and functional probe of the binding-site volume compared to computing the volume based on the structure. One-dimensional ^{19}F direct-polarization (DP) NMR spectra of the substrate at high pH resolve three peaks with intensity ratios of 1:1:2 (Fig. 5a), indicating that the four fluorines of the ligand reside in distinct chemical and conformational environments. The most downfield peak (peak 4) has a narrow linewidth of 0.8 ppm (~450 Hz), indicating that one of the fluorines resides in a well-ordered structural environment. The presence of one narrow ^{19}F peak is also observed at low pH, but this narrow peak has the most upfield chemical shift. In both cases, ^{1}H–^{19}F cross-polarization (CP) spectra preferentially enhanced the intensity of the narrow peak relative to the other ^{19}F signals, indicating that this fluorine is the most immobilized. Compared to the low-pH sample, the high-pH ^{19}F chemical shifts of F$_4$-TPP$^+$ shifted by about 6 ppm downfield, indicating that the binding-site aromatic residues interact with the substrate very differently at high pH.

Dynamics and hydration of the ligand at the binding site. To probe millisecond-timescale motions of the ligand in the binding pocket, we measured 2D ^{19}F–^{19}F exchange spectra of F$_4$-TPP$^+$ at 285 K using mixing times of 0.1 ms to 80 ms (Fig. 5b). Exchange cross peaks are readily observed by ~20 ms at this temperature, but are absent at 265 K[39], indicating that exchange on this timescale is due to motion rather than ^{19}F spin diffusion. The diagonal intensity decays and cross-peak intensity buildup occur with rates of 165–318 s^{-1} (Fig. 5c and Supplementary Fig. 8). These rates are about two-fold faster than the low-pH rates (75–103 s^{-1}), indicating that the ligand is more dynamic in the high-pH complex. These increased dynamics at high pH agree well with the measured thermodynamic parameters for ligand binding. EmrE is a proton-coupled transporter, and we have previously shown that protons are released from both E14 and H110 upon TPP$^+$-binding[18]. ITC experiments were performed using multiple buffers with different ionization enthalpies to determine the number of protons released. These experiments were performed at four different pH values. This data was analyzed to extract the enthalpy and entropy of binding independent of the buffer contribution by extrapolating to $\Delta H_{\text{ionization}} = 0$ (Supplementary Fig. 9). The resulting thermodynamic parameters are shown in Table 3. This data shows that the well-established increase in binding affinity at high pH is driven by an increasingly favorable entropic contribution, while the enthalpy of binding becomes less favorable.
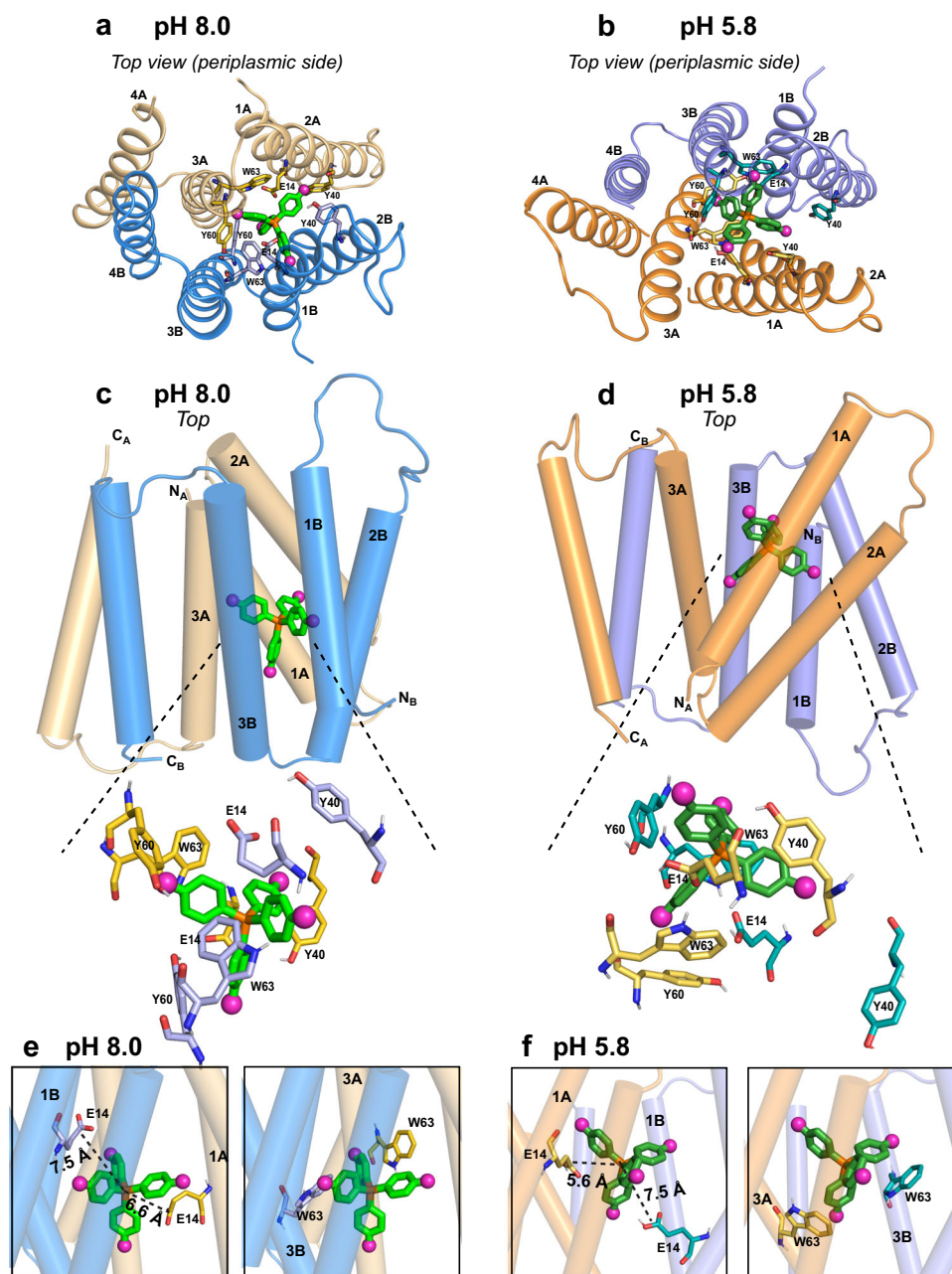
Fig. 4 Structure of the EmrE-TPP complex in DMPC bilayers at high pH (PDB: 7SFQ) and its comparison with the previously determined low-pH structure (PDB: 7JK8). a Top view of the high-pH EmrE-TPP complex structure, seen from the periplasmic side. The crucial binding-site residues E14, Y40, Y60, and W63 are shown as sticks. **b** Top view of the low-pH EmrE-TPP complex for comparison. Note that conformer A (beige) in the high-pH complex has changed to conformer B (purple) in the low-pH complex. This conformational change switches the designation of the two subunits between the high and low pH complexes. **c** Side view of the high-pH structure. **d** Side view of the low-pH structure. The ligand is buried deep in the high-pH complex but is more exposed to the top side in the low-pH complex. The arrangement of the TM helices differs noticeably between the two structures. **e** TPP$^+$ position relative to E14 and W63 in the high-pH structure. The ligand center P atom is similarly distanced from the two E14 residues. **f** TPP position relative to E14 and W63 in the low-pH structure. The ligand is ~2 Å closer to E14A than E14B. .

The faster ligand reorientations in the high-pH complex suggest that the binding pocket should be more hydrated compared to the low-pH complex. To test this hypothesis, we measured water-edited 2D hNH spectra of the protein using water-to-protein ^{1}H polarization transfer times of 30 and 325 ms (Fig. 6). The short-mixing-time spectra selectively exhibit well-hydrated residues. After correcting for ^{1}H R_1 relaxation (Supplementary Fig. 10), we find that the high-pH complex exhibits higher water-transferred intensities than the low-pH protein. In particular, the TM1B and TM3B helices in subunit B

are much more water accessible at high pH than at low pH. Moreover, the extent of hydration is more comparable between the two subunits at high pH compared to the low-pH complex. This increased similarity of the hydration extent at high pH is exemplified by the G26 pair, and is consistent with the smaller chemical shift asymmetry of the protein at high pH. We attribute these observations to the more parallel orientations of the TM helices in the high-pH complex, which reduce the difference in the degree of opening between the two ends of the helical bundle. In comparison, the low-pH EmrE-TPP complex shows much

Table 1 NMR and refinement statistics for F_4-TPP$^+$ bound S64V-EmrE structure in DMPC bilayers at pH 8.0.

	Monomer A	Monomer B
NMR distance and dihedral constraints		
Dipolar coupling measurements	63	52
Distance constraints	205	182
Total dihedral-angle restraints		
ϕ	83	65
ψ	83	65
χ_1	43	33
Structure statistics		
Violations (mean ± s.d.)		
Distance constraints (Å)	0.004 ± 0.042	0.017 ± 0.112
Max. distance-constraint violation (Å)	0.85	1.62
ϕ Dihedral-angle constraints (°)	0.302 ± 1.815	0.545 ± 3.454
ψ Dihedral-angle constraints (°)	0.744 ± 4.289	0.508 ± 3.905
Max. ϕ dihedral-angle violation (°)	17.2	31.9
Max. ψ dihedral-angle violation (°)	53.8	60.1
Average pairwise r.m.s.d (Å)[a]		
Protein heavy atom	3.43 ± 1.06	
Protein backbone	2.85 ± 0.95	
Protein transmembrane heavy atom	2.66 ± 0.86	
Protein transmembrane backbone	2.23 ± 0.83	
Ligand heavy	1.84 ± 0.61	
Ligand center[b]	1.30 ± 0.64	

[a]Pairwise RMSD was calculated among 10 lowest-violation structures between the two independent MD runs after the refinement had equilibrated.
[b]Ligand center is operationally defined as phosphorus and the four directly bonded carbon atoms of F_4-TPP$^+$.

Table 2 Protein-substrate distances extracted from the NMR-refined structural models.

	pH 8.0		pH 5.8	
	Monomer A	Monomer B	Monomer A	Monomer B
P–E14 Cδ	6.6 ± 0.7 Å	7.5 ± 0.5 Å	5.6 ± 0.3 Å	7.5 ± 1.0 Å
P–Y40 Oζ	7.6 ± 1.8 Å	11.2 ± 1.2 Å	6.8 ± 0.5 Å	16.7 ± 0.3 Å
P–Y60 Oζ	9.0 ± 1.2 Å	6.9 ± 1.1 Å	9.8 ± 0.7 Å	5.9 ± 0.4 Å
P–W63 Nε	5.6 ± 0.5 Å	5.7 ± 0.9 Å	6.0 ± 0.4 Å	5.6 ± 0.3 Å
Min. F[a]–E14 Cδ	4.7 ± 0.8 Å	5.2 ± 0.6 Å	4.6 ± 0.5 Å	6.5 ± 0.7 Å
Min. F[a]–Y40 Oζ	6.0 ± 1.0 Å	9.5 ± 0.8 Å	6.2 ± 0.6 Å	12.2 ± 0.4 Å
Min. F[a]–Y60 Oζ	6.1 ± 0.8 Å	5.7 ± 0.6 Å	6.9 ± 0.4 Å	5.6 ± 0.5 Å
Min. F[a]–W63 Nε	4.8 ± 0.3 Å	4.9 ± 0.7 Å	5.8 ± 0.3 Å	5.7 ± 0.4 Å

The average distances and standard deviations are from the ensemble of 10 minimum constraint-violating structures in the final 240 ns of MD trial 1 and 160 ns of MD trial 2. P refers to TPP phosphorus atom while protein atom is denoted by standard IUPAC nomenclature.
[a]Distances of the nearest fluorine to specific protein atoms.

higher hydration for subunit A than subunit B, in good agreement with the larger conformational asymmetry of the two subunits. These spectral observations are borne out by the MD equilibrated structural ensemble, as the structure calculation explicitly solvated the protein-ligand complex in lipid bilayers. Figure 6e shows membrane-embedded water molecules whose oxygen atoms lie within 15 Å of any of the ligand atoms. Strikingly, the ligand-binding site is much more hydrated in the high-pH complex: a total of 69 water molecules are found in the ligand-binding pocket, and these water molecules approach the ligand from both sides of the lipid bilayer. In comparison, only 23 water molecules are found in the ligand-binding pocket in the low-pH complex; moreover, they approach the ligand only from one side, the putative periplasmic side, of the lipid bilayer.

Discussion

The high-pH structure the EmrE-TPP complex is consistent with the available biochemical data and suggests a mechanism for how proton binding drives release of high-affinity substrates from this promiscuous transporter. Extensive mutagenesis of residues throughout EmrE have demonstrated that E14 in TM1 is critical for binding both drug-like substrates and protons for the transport activity[40–42]. Residues in TM2 and TM3 are important for substrate binding and substrate specificity[4,43–45]. TM3 also contains a putative hinge region that is important for controlling the rate of alternating access of the transporter. TM4 is outside the binding pocket and contains the dimerization motif that stabilizes this highly dynamic homodimer[24,32,46]. The measured substrate-protein distances shown here and in the previous study are the shortest for residues in TM1–3 at both low and high pH, consistent with the substrate-binding site inferred from mutagenesis and the overall structural organization of the transporter observed in low-resolution EM[9] and crystal structures[8]. The asymmetry of chemical shifts between subunits A and B is the largest in TM3 (Supplementary Fig. 4). This is consistent with the proposal that the hinge and the difference in local conformation of this helix[21] between the two subunits control the asymmetric structure of the EmrE dimer and the rate of alternating access.

The fact that these significant structural changes result exclusively from a pH change is remarkable. Deprotonation of the two
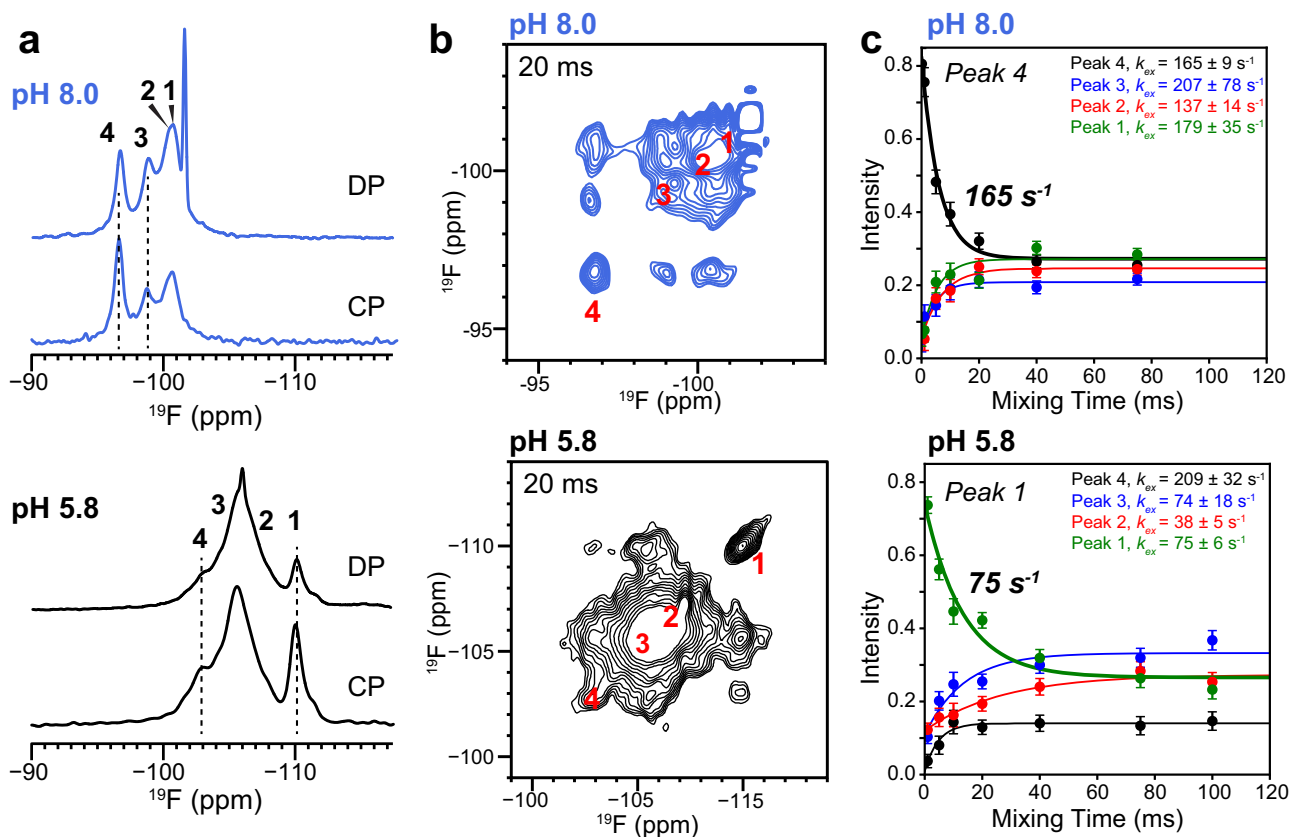
Fig. 5 ^{19}F NMR spectra of F$_4$-TPP$^+$ to probe substrate structure and dynamics. a 1D ^{19}F NMR spectra of F$_4$-TPP$^+$ bound to EmrE at pH 8.0 and pH 5.8. Multiple ^{19}F chemical shifts are resolved, indicating heterogeneous structural environments of the four fluorines. **b** 2D ^{19}F–^{19}F correlation spectra of EmrE-bound F$_4$-TPP$^+$ at pH 8.0 (blue) and pH 5.8 (black), measured with a mixing time of 20 ms. **c** ^{19}F exchange buildup and decay curves for the resolved peak 4 at pH 8.0 and the resolved peak 1 at pH 5.8. Data are presented as mean values +/− 2σ. Error of intensity values was propagated from spectral signal to noise, while fitting parameter errors were estimated by Monte Carlo methods. These ^{19}F spectra were measured under 38 kHz MAS at a sample temperature of ~285 K. Cross-peak intensities are provided as a Source Data file.

Table 3 pH-dependent binding parameters determined by ITC.

| pH	ΔG (kJ/mol)	ΔH (kJ/mol)	$-T\Delta S$ (kJ/mol)
5.5	−25.3 ± 0.3	−40.0 ± 0.6	15 ± 0.7
6.5	−34.9 ± 0.4	−31.5 ± 1.7	−3.4 ± 1.7
7.5	−41.1 ± 0.8	−19.1 ± 0.6	−21.0 ± 1.0
8.5	−46.0 ± 0.6	−20.7 ± 2.7	−25.3 ± 2.8

Original data for TPP$^+$ binding to WT EmrE was reported in[18]. Here we present the average pH-dependent binding affinity, enthalpy, and entropy for substrate–protein interaction, independent of any contribution from buffer ionization.

E14 residues not only symmetrizes the TPP$^+$ position but also increases the separation between the two TM1 helices. The E14 Cδ-Cδ distance increased from 11.8 ± 1.3 Å at low pH to 13.3 ± 0.8 Å at high pH. This subtle increase of the E14–E14 separation, and the ensuing change in the binding pocket geometry (Supplementary Fig. 7), translates to a noticeable effect on the drug dynamics, speeding up ligand reorientation rates two-fold compared to the low-pH complex. The motion is likely tetrahedral jumps. 2D ^{19}F–^{19}F exchange data suggest that this tetrahedral jump is incomplete in geometry: while peak 4 displays equilibrated exchange intensities of ~0.25 by 80 ms, the other three sites do not fully equilibrate, suggesting that these fluorines are impeded by the protein sidechains. The substrate motion is more complete in the high-pH complex than in the low-pH complex, whose ^{19}F exchange intensities are less

equilibrated (Supplementary Fig. 8d). The differences between the structures and drug dynamics of the low- and high-pH states indicate that ligand dynamics depend both on the size of the binding pocket and on sidechain obstructions.

Comparison of the structures of the EmrE-TPP complex at low and high pH is most informative in understanding how protonation of one E14 residue can drive release of the TPP$^+$ substrate. It is well established that the apparent affinity of EmrE for drug-like substrate is weaker at low pH than at high pH[42], and this is due to a faster substrate off-rate at low pH. This pH-dependent substrate affinity was originally attributed to a simple competition between TPP$^+$ and proton for binding to E14. However, we now know that EmrE can bind proton and TPP$^+$ simultaneously at low pH[17,39], and the two E14 residues are sufficiently far apart from each other that electrostatic coupling is minimal and proton binding and release occur relatively independently of either residue[30]. Examining the structural changes of the EmrE–TPP complex from high pH to low pH immediately suggests why TPP$^+$ affinity is lower in the protonated state. At high pH, TPP$^+$ is buried deep within the helical bundle, positioned nearly symmetrically between the two TM1 helices within an elongated binding cavity (Fig. 4 and Supplementary Fig. 7). In contrast, at low pH, the transporter structure is more asymmetric and clearly closed on one side and open on the other. TPP$^+$ is positioned much closer to the open end of the binding cavity, primed for release. The weakening of substrate affinity upon proton binding is likely important for speeding up the release of tight-binding substrates so that this promiscuous transporter can rapidly efflux
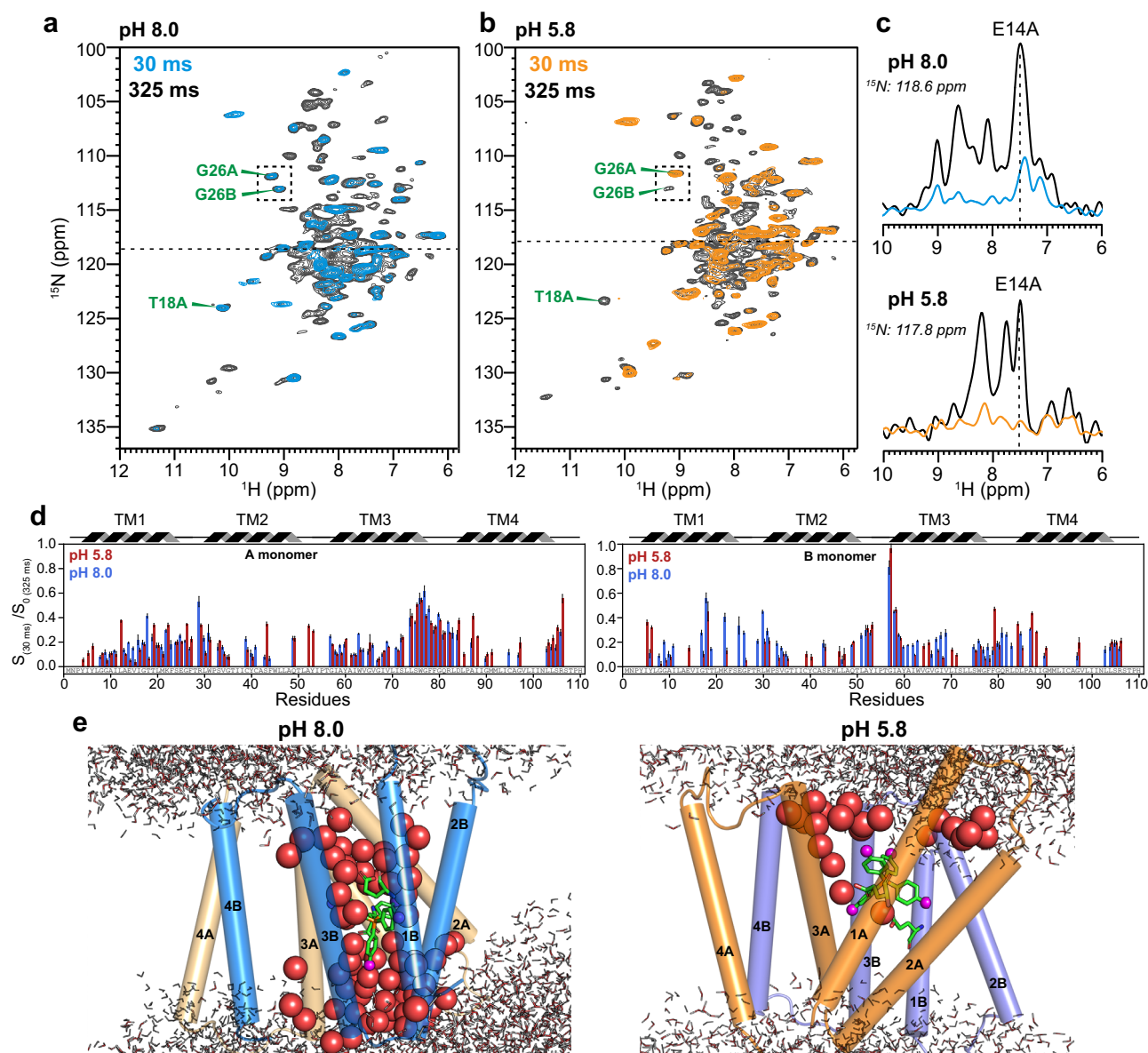
Fig. 6 Water-edited 2D hNH spectra of TPP$^+$-bound EmrE. a pH 8.0 spectra. **b** pH 5.8 spectra. **c** 1D ^{15}N cross-sections of E14A extracted from the water-edited 2D hNH spectra. The pH 8.0 sample shows higher water-transferred intensities than the pH 5.8 sample. **d** Intensity ratios S/S_0 between the 30 and 325 ms spectra of the water-edited 2D spectra to compare the water accessibilities between the high and low pH proteins. The intensities have been corrected for ^{1}H R_1 relaxation. **e** MD equilibrated membrane-bound structures of the EmrE-TPP$^+$ complex at pH 8.0 and pH 5.8. Water oxygens within 15 Å of any of the ligand atoms and that are within the lipid bilayer are shown as red spheres. Membrane-surface water molecules are shown as thin lines, and water molecules >15 Å away from the protein are shown in gray. The number of water molecules is 69 in the high-pH complex but only 23 in the low-pH complex, consistent with the higher water-transferred intensities of the protein at high pH. Moreover, water hydrates the ligand from both sides of the membrane in the high-pH complex, but only accesses the ligand from the periplasmic side in the low-pH complex. The water-edited intensities are provided as a Source Data file.

a broad range of substrates with widely varying affinities[17,19]. The pH-dependent thermodynamic parameters (Table 3) show that higher-affinity drug binding by this promiscuous transporter is driven by increasingly favorable binding entropy. This is consistent with the enhanced ligand dynamics and hydration observed here in the high-pH structure. There is not sufficient ITC data or in vitro transport data available to assess whether the relative promiscuity of EmrE changes appreciably with pH. That remains an open question for further study.

The native environment for EmrE is the inner membrane of *E. coli*. This is an asymmetric environment unlike the in vitro conditions used for structure determination. Usually, the periplasmic pH is lower than the cytoplasmic pH[47]. Given the significant structural changes of the protein we determined here as a function of E14 protonation state in a symmetric membrane, it is reasonable to assume that the protonation state of E14 regulates the overall structural change of the protein. We can then infer what the structure might look like in the presence of a transmembrane pH gradient. Cytoplasmic pH in *E. coli* is generally 7.4–7.8[48], above the pK$_a$ of E14 in the F$_4$-TPP$^+$-bound transporter. Thus, the open-in conformation of EmrE is expected to have predominantly deprotonated E14 residues and more closely resemble the high-pH structure (Fig. 4a, c, e). In contrast, the periplasmic pH closely correlates with the pH of the external

environment[47]. Thus, when *E. coli* are in an acidic environment like the human gut, the periplasmic pH will be well below the pK_a of E14 and the structure of EmrE should more closely resemble the low pH structure (Fig. 4b, d, f).

It is interesting to note that the high-pH structure of the EmrE-TPP complex is not fully closed on either side of the membrane. This is borne out by the water-accessibility data (Fig. 6a–d) and the MD simulations (Fig. 6e), which show larger water accessibility of the ligand-binding cavity at high pH than at low pH. Moreover, water molecules approach the ligand from both sides of the membrane at high pH, but only access the ligand from one side of the membrane at low pH. The dual-open conformation at high pH suggests that after a toxic substrate has bound to the protein from the cytoplasm and EmrE adopts a conformation similar to Supplementary Fig. 7a, there is sufficient space for protons to enter and bind E14 from the periplasmic side. This could then trigger a conformational change to a state similar to the low-pH structure, with the transporter open to the periplasmic side of the membrane. In this outward-facing state, the substrate is bound peripherally to the central transport pore, and is thus primed for release in accord with the enhanced off rate in the drug- and proton-bound state.

Methods

S64V-EmrE expression and purification. S64V-EmrE was expressed and purified as previously described[7] using the same procedure as for WT EmrE[49]. ^{2}H,^{13}C,^{15}N (CDN)-labeled S64V-EmrE was expressed in ^{2}H$_2$O media containing 2.5 g/L U–^{2}H,^{13}C glucose, 1 g/L ^{15}NH$_4$Cl, and 0.5 g/L ^{2}H,^{13}C,^{15}N-labled ISOGRO. Lysis and purification were performed as previously reported[50,51] using Ni-NTA affinity column followed by thrombin cleavage of the His-tag and size exclusion chromatography on a S200 column in buffer containing 50 mM MES, 20 mM NaCl, 10 mM decyl-maltoside, 5 mM BME, pH 7.0.

Solution NMR spectroscopy and pK_a analysis. Solution NMR data were collected on samples with 1.0 mM ^{2}H, ^{15}N S64V-EmrE in DMPC/DHPC bicelles ($q = 0.33$, with a protein to DMPC molar ratio of 1:50). The buffer contained 20 mM NaCl, 50 mM sodium acetate, 50 mM MOPS, 50 mM MES, 50 mM bicine, 2 mM TCEP, 0.01% DSS, and 10% D$_2$O. About 10 mg F$_4$-TPP$^+$ was added to this protein bicelle solution and incubated at 45 °C overnight to saturate binding, then the excess drug was removed through microcentrifugation. Spectra were measured at 45 °C on a Bruker 750 MHz spectrometer (Avance) equipped with a TCI cryoprobe. Spectra from 4 different pH conditions were processed and analyzed using NMRPipe[52] and CcpNmr Analysis[53]. Chemical shift changes for ^{1}H and ^{15}N were separately globally fit to a single pK_a value using the following equation (Eq. 1):

$$\delta = \frac{\delta_H 10^{-pH} + \delta_D 10^{-pK_a}}{10^{-pH} + 10^{-pK_a}} \tag{1}$$

where δ_H and δ_D are the chemical shifts of the protonated and deprotonated states of each residue, respectively. Six residues in close proximity to E14 with large chemical shift changes with pH were analyzed. For a single pK_a, the modified Henserson–Hasselbach equation describes the chemical shift for each residue (δ) as a function of pH[15,54,55].

Reconstitution and preparation of solid-state NMR samples. CDN-labeled S64V-EmrE was reconstituted into d_{54}-DMPC (Avanti Polar Lipids) liposomes at a protein monomer to lipid molar ratio (P: L) of 1: 25. DMPC was resuspended in 50 mM MES, 20 mM NaCl, pH 8.0 buffer at 20 mg/mL. The lipid mixture was incubated at 45 °C for 1 h to hydrate, then bath-sonicated for 1 min before addition of 0.5% octyl-glucoside (OG) followed by 30 s bath sonication. The lipid mixture was incubated at 45 °C for an additional 15 min before mixing with purified S64V-EmrE solution. After 20 min incubation at room temperature (RT), Amberlite (Supelco) was added (3 × 30 mg Amberlite per mg total detergent) to remove the detergent. The amberlite was removed after 16–24 h by simple filtration. Liposomes were collected by ultracentrifugation (165,000 × *g*, 6 °C, 2 h) and resuspended in a small volume (~20 mg/mL lipid concentration) of buffer. To ensure complete detergent removal, the sample was dialyzed against 1 L of the same buffer (50 mM MES, 20 mM NaCl, pH 8.0) with buffer change every 24 h over a 72 h period. The sample was then incubated with excess solid F$_4$-TPP$^+$ at RT with end-over-end rocking for at least 16 h. Excess F$_4$-TPP$^+$ was removed using microcentrifugation (7,500 × *g*, 5 min). Proteoliposomes were then pelleted at 100,000 × *g*, 4 °C, 2 h in an ultracentrifuge. Proteoliposomes were dried to ~40% hydration by mass in a desiccator. Samples were centrifuged into 1.9 and 1.3 mm MAS rotors. Two rotors were packed: a 1.9 mm sample of CDN-EmrE sample containing ~3.5 mg protein

in ~15 mg proteoliposomes, and a 1.3 mm sample of CDN-EmrE containing ~0.9 mg protein in ~3.6 mg proteoliposomes.

Solid-state NMR experiments. All MAS SSNMR experiments were conducted on a 600 MHz Bruker AVANCE II spectrometer. Chemical shift assignment and protein hydration experiments were carried out under 55 kHz MAS on a 1.3 mm HCN probe, whereas 2D ^{1}H–^{15}N resolved ^{1}H–^{19}F REDOR experiments and 2D ^{19}F–^{19}F exchange experiments were conducted under 38 kHz MAS on a 1.9 mm HFX probe. Sample temperature was controlled by matching the chemical shift of water in the proteoliposomes between samples and probes. The effective sample temperatures were estimated from the water ^{1}H chemical shifts using the equation T_{eff} (K) = 96.9(7.83 − δ_{H_2O}) where δ_{H_2O} is the observed water chemical shift[56]. By keeping the water ^{1}H chemical shift at 4.89 ppm using appropriate bearing temperatures, we maintained a constant sample temperature of 285 K for all experiments. Detailed experimental conditions are provided in Supplementary Table 6.

Eight ^{1}H-detected 3D MAS correlation experiments were used to assign the chemical shifts of pH 8.0 EmrE. Supplementary Fig. 11 shows the pulse sequences of those experiments that were not included in our low-pH EmrE study[39] and the pulse sequences of the 2D water-edited experiments. The hCANH, hCOcaNH, and hcaCBcaNH experiments allow intra-residue CA, CO, and CB assignment, while the hCAcoNH, hCONH, and hcaCBcacoNH experiments allow sequential residue assignment. An amide-to-amide 3D correlation was achieved using the HncacoNH (Supplementary Fig. 11c) and hNcacoNH (Supplementary Fig. 11d) experiments, which allow inter-residue H$_{i-1}$–N$_i$H$_i$ and N$_{i-1}$–N$_i$H$_i$ assignment, respectively, to further disambiguate the backbone walk[57]. The coherence transfer steps and resonance assignment connectivities of these eight ^{1}H-detected experiments are summarized in Supplementary Table 1. ^{15}N–^{13}C correlation experiments used specific CP for polarization transfer[58]. Under 55 kHz MAS, the double-quantum (DQ) matching condition of 30 kHz ^{13}C and 25 kHz ^{15}N, or 35 kHz ^{15}N and 20 kHz ^{13}C, was used to achieve selective ^{15}N–^{13}C polarization transfer. ^{13}CA–^{13}CO polarization transfer was achieved using the DQ DREAM sequence under 55 kHz MAS. Due to off-resonance effects, a short 1.3–1.6 μs trim pulse (marked as a θ pulse) was used after the spin lock to rotate the magnetization to the *XY* plane[59]. For the two CB-NH correlation experiments (Fig. S11a, b), out-and-back CA-CB-CA INEPT transfer was used, using the Q3 shaped pulse for 180° pulses to invert aliphatic coherences while not inverting CO. A 6.5 ms INEPT mixing period was used for both the creation and reconversion of the antiphase magnetization, for a total of 13 ms of transfer time. Low-power ^{1}H decoupling was employed for all SSNMR experiments with either CW irradiation or the WALTZ-16 scheme[60]. These ^{1}H-detected MAS NMR experiments employed MISSISSPPI to suppress the water signal, using 150–200 ms of 15 kHz irradiation[61].

Water-edited NMR experiments were carried out with 2D hNH detection by inserting a water-selective echo and water-to-protein ^{1}H spin diffusion mixing period following proton excitation and before CP to ^{15}N (Supplementary Fig. 11e). A 3.5 ms Gaussian pulse with 5% truncation and 400 points for the shape was used to selectively refocus water coherence within a 3.6 ms echo, during which all protein coherence is destroyed by T$_2$ relaxation. Due to fast MAS suppressing spin diffusion, long mixing times of 325 ms for the equilibrated S_0 spectrum and 30 ms for the edited S spectra were needed, between which significant T$_1$ relaxation occurs. To account for relaxation between the edited and equilibrated spectra, site-resolved T$_1$ measurements were carried out through saturation recovery using the pulse sequence in Supplementary Fig. 11f. After the pre-scan delay d1, we inserted a MISSISSIPPI dephasing block to saturate all ^{1}H magnetization in the sample. Following this saturation, a variable delay allows T$_1$ relaxation to occur before ^{1}H excitation. The experiments were run in a constant-time fashion, where the combined pre-scan delay and τ_{relax} were set to 3, 4, or 5 s depending on the relaxation time τ_{relax} used. ^{1}H–^{19}F REDOR measurements and 2D ^{19}F–^{19}F exchange experiments were run as described previously[39] to allow for direct comparison of the data between the two samples of distinct pH.

^{19}F chemical shifts were externally referenced to the −122.1 ppm signal of 5F-tryptophan on the CF$_3$Cl scale[62]. ^{13}C, and ^{15}N chemical shifts were internally referenced to the DSS-referenced chemical shifts of the solution-state values. The G90A site was previously shown to have little chemical shift perturbation with pH[15], and was therefore chosen as the referencing site; the ^{1}H and ^{15}N values were set to 8.5 ppm and 107.4 ppm, respectively, and the ^{13}C reference was set from the hCANH peak for this residue at 47.6 ppm.

Chemical shift assignment. The SSNMR spectra were processed in the Bruker Topspin 3.5 software package with zero-filling, apodization, and Fourier-transformation (FT). Spectra acquired in blocks for signal averaging were added in the time domain prior to FT using a custom Python script that utilized the NMRGlue and NumPy Python packages[63,64]. Chemical shift assignment and 3D spectral plotting were performed in NMRFAM-Sparky[65]. 1D and 2D correlation spectra were plotted using Bruker Topspin's XWINPLOT. Comparisons of chemical shifts between protein monomers and samples were computed in Python and plotted with Matplotlib[66]. The asymmetry of protein monomer chemical shifts was

calculated using composite H^N and N^H chemical shifts according to Eq. 2:

$$\Delta\omega_{NH} = \sqrt{\frac{1}{2}[\Delta\omega_H^2 + (0.10 * \Delta\omega_N)^2]} \qquad (2)$$

Here $\Delta\omega_{H/N}$ is the ^1H and ^{15}N chemical shift difference between monomers A and B. The composite CA and CO chemical shift difference was calculated similarly but without the factor of 0.10 for the gyromagnetic ratio difference. Protein ϕ, ψ, and χ_1 torsion angles were predicted using the TALOS-N software[67] based on the measured non-H chemical shifts. A deuterium isotope correction was applied to the $C\alpha$ and $C\beta$ chemical shifts.

2D ^{19}F–^{19}F exchange analysis. We quantified the exchange rates between four peaks in the 2D ^{19}F–^{19}F correlation spectra at both high and low pH. Peak volumes were integrated in Topspin using the same peak areas across all mixing times, and were normalized with respect to the sum of the integrated intensities of all peaks in each row. Thus, at short mixing times where most intensities reside on the diagonal, the three cross peaks should have normalized intensities near 0, while at sufficiently long mixing times, the diagonal and three cross peaks should equilibrate to similar intensities of ~0.25. In practice, due to spectral overlap, the measured intensities deviate from the ideal equilibrium values. The normalized intensities were fit to a single-exponential decay (Eq. 3) and single-exponential buildup (Eq. 4) equations for the diagonal and cross peaks, respectively:

$$I_{Diagonal}(t_{mix}|Y_0, P, k) = (Y_0 - P) * e^{-k*t_{mix}} + P \qquad (3)$$

$$I_{CrossPeak}(t_{mix}|Y_0, P, k) = (P - Y_0) * (1 - e^{-k*t_{mix}}) + Y_0 \qquad (4)$$

Here Y_0 is the initial intensity, P is the plateau value, and k is the exponential buildup or decay rate, and t_{mix} is the mixing time. Fitting was performed in the SciPy optimization module of python[68]. Errors are 2σ for individual points, and were estimated from the signal-to-noise ratios (SNR) of the spectral peaks. The SNR for a reference peak was measured using Sparky's built-in routine using 1000 random points to calculate SNR, and the SNR for the remaining peaks was estimated from this value by scaling by the intensity ratio of the reference peak to the remaining peaks. The errors were propagated using Eq. 5[36].

$$\epsilon_i = 2 * \sigma = 2 * I_i * \sqrt{\left(\frac{1}{SNR_i}\right)^2 + \left(\frac{1}{SNR_{norm}}\right)^2} \qquad (5)$$

where I_i is the normalized intensity of the i-th peak, SNR_i is the SNR of the i-th peak, and SNR_{norm} is the relative SNR of the normalization factor for the row. The errors for the fitting parameters were determined using Monte Carlo analysis as shown before for $T_{1\rho}$ experiments[69]. Briefly, we simulated 1000 additional datasets for each decay and buildup curve by adding a random value from a Gaussian distribution centered at $\mu = 0$, $\sigma = 0.3$ (chosen to give values mostly within ± 1) multiplied by the error for each point, ϵ_I. In this way, we created 1000 new datasets with the points moving randomly within the error bar for each point, but with a Gaussian distribution around the actual measured value. These 1000 datasets were then fit to the same buildup or decay curves to obtain the fitting parameters for each simulated dataset. The standard deviation of the fit parameters in the Monte Carlo datasets was then used as the fitting parameter error reported.

^1H–^{19}F distance extraction. 2D hNH-resolved H^N–^{19}F REDOR distance restraints were determined as described before[36,37,39]. Briefly, we integrated peak volumes in the 2D ^1H–^{19}F REDOR-edited hNH S_0 and S spectra to obtain the intensity ratios S/S_0 for all mixing times for each protein H^N. Using the SIMPSON software package, we simulated the two-spin REDOR dephasing curves for distances of 3.0–15.0 Å in 0.1 Å increments, including the magnitude and asymmetry parameters of the ^{19}F CSA, but with default orientation[70]. Finite-pulse effects were explicitly included in the SIMPSON program. RF inhomogeneity was accounted for by simulating for pulse flip angles of 180°–145° in 5° increments, weighted by a half-Gaussian function centered at 180° and a standard deviation of 15°[36,37]. The REPULSION320 scheme with 32 gamma angles was used for powder averaging[71]. The best-fit ^1H–^{19}F distance was extracted by minimizing the RMSD between the simulated and measured S/S_0 values. The uncertainty in the best-fit distance was set by the same RMSD threshold of 0.2 as the previous study[39]. Distances below this RMSD value were considered significant. In cases where little to no dephasing was observed, we set the distance upper uncertainty to 40 Å, the approximate longest possible distance in the dimer (Tables S2, S4). For residues whose signals overlap in the 2D hNH spectrum, the lower-limit distance uncertainty was increased.

Analysis of water-edited spectra under fast MAS. Protein hydration was investigated using a water-edited 2D experiment where water ^1H polarization was selectively excited and transferred to the protein and detected in 2D hNH spectra (Supplementary Fig. 10e). The hydration intensities were analyzed using a modified procedure from the previously reported approach[72–74] to account for the effects of fast MAS. Because fast MAS suppresses spin diffusion, it was necessary to use longer ^1H mixing times for both the equilibrated S_0 spectrum (325 ms) and the edited S spectrum (30 ms). The edited spectra were signal averaged with 3.5–4.5 times as many scans as the equilibrated spectra to obtain sufficient SNR. Site-resolved hydration intensities were calculated by dividing the integrated peak

volumes of the 30 ms S spectrum by the corresponding peak volumes in the 325 ms S_0 spectrum. However, at 325 ms mixing, ^1H T_1 relaxation is non-negligible, causing the S/S_0 values to be larger than 1 for some residues. To correct for this T_1 relaxation, we measured 2D hNH resolved ^1H saturation-recovery spectra (Supplementary Fig. 10) to obtain site-specific ^1H R_1 rates at both pH values. The ^1H R_1 rates were extracted from the integrated peak volumes using the same integration areas as for the water-edited spectra. The intensities of each H^N site were normalized to the maximum intensity for that site. Error bars for each point were propagated from the SNR of the spectra according to Eq. 6:

$$\epsilon_i = 2 * \sigma = 2 * I_i * \sqrt{\left(\frac{1}{SNR_i}\right)^2 + \left(\frac{1}{SNR_{norm}}\right)^2} \qquad (6)$$

The saturation recovery curves were fit to a single-exponential buildup function (Eq. 7) to obtain the R_1 values:

$$I(t_{relax}|P, R_1) = P * (1 - e^{-R_1*t_{relax}}) \qquad (7)$$

The R_1 uncertainty, σ_{R1}, was estimated using the same Monte Carlo method described above for the ^{19}F exchange analysis. With the site-specific R_1 rates known, the hydration S/S_0 values were corrected for relaxation between the two mixing times (325 and 30 ms) according to Eq. 8:

$$H\left(\frac{S}{S_0}, R_1\right) = \frac{S}{S_0 * e^{R_1(t_2 - t_1)}} = \frac{S}{S_0} * e^{R_1(t_1 - t_2)} \qquad (8)$$

The error of the corrected S/S_0 value, H, is a result of the errors in the S, S_0, and R_1 values and was calculated according to Gauss' error propagation according to Eqs. 9 and 10:

$$\sigma_{S/S_0} = \frac{S}{S_0} * \sqrt{\left(\frac{1}{SNR_S}\right)^2 + \left(\frac{1}{SNR_{S_0}}\right)^2} \qquad (9)$$

$$\sigma_H = \sqrt{e^{2R_1(t_1-t_2)} * \sigma_{S/S_0}^2 + \left[\frac{S}{S_0}(t_1 - t_2)e^{R_1(t_1-t_2)}\right]^2 * \sigma_{R_1}^2} \qquad (10)$$

Using this method, the water-edited intensities can be compared fairly for different residues with different R_1 rates and for two pH conditions.

Structure calculation of EmrE with bound F_4-TPP$^+$ at high pH. The high-pH EmrE-TPP structure was calculated in a two-stage process similar to that detailed recently[39]. F_4-TPP$^+$ (PDB: VCJ) was first docked into the pH 5.8 protein structure (PDB: 7JK8) using ^1H–^{19}F distance constraints measured at pH 8.0. This docking orients the ligand and disambiguates the four-fold degenerate ^1H–^{19}F constraints. The protein was then subject to all-atom molecular dynamics refinement in explicitly solvated DMPC bilayers. In both stages, the two E14 residues were modeled in the deprotonated state. Docking was performed using the HADDOCK v2.4 webserver[75,76]. The "active residues" list was set to a list of 18 protein residues between the A and B subunits which are known to with the ligand from prior biochemical data[77] and our REDOR data. The new ^1H–^{19}F REDOR distances were used as "unambiguous" constraints which are always enforced, with ranges determined by the RMSD cutoff of 0.2 (Supplementary Fig. 5). At this stage, the ^{19}F atoms were input as four-fold ambiguous. Both ambiguous (active-residue defined) and unambiguous (distance measurement defined) constraints used default HADDOCK energy weighting values of 10.0, 10.0, 50.0, and 50.0 for the hot, cool1, cool2, and cool3 stages of the docking simulations. Docking was performed in DMSO and started with 1000 structures, from which 200 were outputted after refinement. These 200 structures were analyzed against the four-fold ambiguous ^1H–^{19}F distance constraints using an integrated Python-Pymol script[63,68] to select the structures that best agree with the experimental data (Supplementary Table 2). We scored the structures by the lowest sum-total of violations and the least number of violations to create two separate ensembles (Supplementary Fig. 6a and Supplementary Table 3). The lowest violation by each criterium was used to structurally disambiguate the ^1H–^{19}F pairs to create 387 distance constraints from the 116 dipolar coupling measurements[39] (Supplementary Table 4). The majority (92) of the measured dipolar couplings are weak and correspond to long distances that are four-fold degenerate. About 18 dipolar couplings are strong and can be assigned to unambiguous H^N–F pairs based on docking. Each of the two lowest violation structures was used for further refinement by MD simulations in GROMACS.

The docked high-pH EmrE-TPP complexes were aligned to the membrane normal using the OPM webserver[78] and were inserted into explicitly hydrated DMPC bilayers with the CHARMM-GUI[79,80] membrane builder tool[81,82]. The bilayer included 224 DMPC molecules, with 114 in one leaflet and 110 in the second, and was hydrated with TIP3 water molecules[83]. The ligand forcefield was parameterized from the coordinates, and the simulation was conducted in GROMACS[84] on NMRBox virtual servers[85]. The simulation was conducted at 310 K with CHARMM36 force fields[86–89] including the WYF parameter for cation–π interactions[90]. The simulation was restrained by the protein–ligand ^1H–^{19}F distances and TALOS predicted (ϕ, ψ) and sidechain χ_1 torsion angles (Supplementary Fig. 2b). The restoring force for time-averaged interatomic distance restraints used was the piecewise linear force described in the online GROMACS documentation, where if $\bar{r}_{ij}$ is the time averaged distance between atoms i and j, the

force would be proportional to $\bar{r}_{ij} - r_0$ below r_0, zero between r_0 and r_1, proportional to $\bar{r}_{ij} - r_1$ between r_1 and r_2, and proportional to $r_2 - r_1$ above r_2. Simple time-averaged distance restraints using conservative weighting were used. The time averaging constant was set to 5 ps, and the distance restraint force constant was set to 1000 kJ/(mol*nm^2). Dihedral restraints were also implemented with an energy weighting of 1000. The simulation started with a 5000-step minimization with position restraints, followed by 1.875 ns of equilibration over which the position restraints were progressively weakened. The production stage of the simulation was subject only to experimental constraints and was carried out in 2 fs steps for 360 and 440 ns in two runs (Supplementary Fig. 6b). Structural ensembles from the two trajectories were assembled by taking structures in 5 ns increments in the equilibrated portion of the simulation (200–400 ns for run 1, 200–360 ns for run 2) for a total of 74 structures. These structures were subjected to 5000 steps of energy minimization with position and torsion angle restraints to remove improper dihedral angles introduced by fast-timescale fluctuations. These 74 structures were scored against the original four-fold ambiguous distance constraints to select 10 structures that best agreed with the experimental data according to the sum total of distance violations (Supplementary Table 5). By chance, five of the best structures came from the first run (mean violation magnitude 0.4 ± 0.1 Å, mean number of violations = 4.4 ± 0.5), while another five came from the second run (mean violation magnitude 0.4 ± 0.2 Å, mean number of violations = 4.2 ± 0.8). Within each run, the ensembles are well ordered, with a mean backbone RMSD of 1.6 ± 0.3 Å for the 5 structures of run 1, and 2.3 ± 0.5 Å for the 5 structures of run 2. Between the two sub-ensembles, the variability was higher, at 2.9 ± 0.5 Å backbone RMSD; most of the differences are localized to the loop regions and TM4 (Supplementary Fig. 6c) where the MD simulation has few constraints: the mean backbone RMSD between the two runs for the transmembrane helices 1–3 was 2.2 ± 0.3 Å, lower than the variability within the sub-ensemble of run 2. The water accessibility of the binding site was examined in the fully hydrated lowest-violation complexes of the bilayer-protein-ligand system after final MD energy minimization. Water molecules whose oxygen atom is within 15 Å of any ligand atom and whose oxygen z-coordinate also lies within the top and bottom planes of the protein TM helical bundles are selected. For the pH 8.0 complex, the top and bottom z-planes were set to be between the F23B and G77A Cα atoms. For the pH 5.8 complex, the top and bottom z-planes were set to be between the z-coordinates of the Y53A and T56B Cα atoms.

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

Data availability

Solid-state NMR chemical shifts and distance restraints have been deposited in the Biological Magnetic Resonance Bank (BMRB) with ID number 30957. The structural coordinates have been deposited in the Protein Data Bank with accession code 7SFQ. Source data are provided with this paper.

Code availability

Python code for ^{1}H–^{19}F REDOR analysis, structurally-based H–F pair assignment, Gromacs simulations, ^{19}F–^{19}F exchange analysis, and water-edited ^{1}H–^{15}N analysis are available upon request to meihong@mit.edu.

References

1. Nikaido, H. Multidrug resistance in bacteria. *Annu. Rev. Biochem.* **78**, 119–146 (2009).
2. Higgins, C. F. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* **446**, 749 (2007).
3. Yerushalmi, H., Lebendiker, M. & Schuldiner, S. EmrE, an *Escherichia coli* 12-kDa multidrug transporter, exchanges toxic cations and H+ and is soluble in organic solvents. *J. Biol. Chem.* **270**, 6856–6863 (1995).
4. Saleh M., Bay D. C. & Turner R. J. Few conserved amino acids in the small multidrug resistance transporter EmrE influence drug polyselectivity. *Antimicrob. Agents Chemother.* **62**, e00461–18 (2018).
5. Schuldiner, S. et al. Small is mighty: EmrE, a multidrug transporter as an experimental paradigm. *N. Physiol. Sci.* **16**, 130–134 (2001).
6. Cho, M.-K., Gayen, A., Banigan, J. R., Leninger, M. & Traaseth, N. J. Intrinsic conformational plasticity of native EmrE provides a pathway for multidrug resistance. *J. Am. Chem. Soc.* **136**, 8072–8080 (2014).
7. Wu, C. et al. Identification of an alternating-access dynamics mutant of EmrE with impaired transport. *J. Mol. Biol.* **431**, 2777–2789 (2019).
8. Chen, Y.-J. et al. X-ray structure of EmrE supports dual topology model. *Proc. Natl Acad. Sci. USA* **104**, 18999 (2007).
9. Ubarretxena-Belandia, I., Baldwin, J. M., Schuldiner, S. & Tate, C. G. Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J.* **22**, 6175–6181 (2003).
10. Kermani, A. A. et al. The structural basis of promiscuity in small multidrug resistance transporters. *Nat. Commun.* **11**, 6064 (2020).
11. Sulavik, M. C. et al. Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug Efflux pump genes. *Antimicrob. Agents Chemother.* **45**, 1126–1136 (2001).
12. Nishino, K., Nikaido, E. & Yamaguchi, A. Regulation and physiological function of multidrug efflux pumps in *Escherichia coli* and Salmonella. *Biochim. Biophys. Acta* **1794**, 834–843 (2009).
13. Matsumura, K., Furukawa, S., Ogihara, H. & Morinaga, Y. Roles of multidrug efflux pumps on the biofilm formation of *Escherichia coli* K-12. *Biocontrol Sci.* **16**, 69–72 (2011).
14. Bay, D. C., Stremick, C. A., Slipski, C. J. & Turner, R. J. Secondary multidrug efflux pump mutants alter *Escherichia coli* biofilm growth in the presence of cationic antimicrobial compounds. *Res. Microbiol.* **168**, 208–221 (2017).
15. Morrison, E. A., Robinson, A. E., Liu, Y. & Henzler-Wildman, K. A. Asymmetric protonation of EmrE. *J. Gen. Physiol.* **146**, 445–461 (2015).
16. Gayen, A., Leninger, M. & Traaseth, N. J. Protonation of a glutamate residue modulates the dynamics of the drug transporter EmrE. *Nat. Chem. Biol.* **12**, 141 (2016).
17. Robinson, A. E., Thomas, N. E., Morrison, E. A., Balthazor, B. M. & Henzler-Wildman, K. A. New free-exchange model of EmrE transport. *Proc. Natl Acad. Sci. USA* **114**, E10083 (2017).
18. Thomas, N. E. et al. The C terminus of the bacterial multidrug transporter EmrE couples drug binding to proton release. *J. Biol. Chem.* **293**, 19137–19147 (2018).
19. Hussey G. A., Thomas N. E. & Henzler-Wildman K. A. Highly coupled transport can be achieved in free-exchange transport models. *J. Gen. Physiol.* **152**, e201912437 (2020).
20. Tate, C. G., Ubarretxena-Belandia, I. & Baldwin, J. M. Conformational changes in the multidrug transporter EmrE associated with substrate binding. *J. Mol. Biol.* **332**, 229–242 (2003).
21. Morrison, E. A. et al. Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature* **481**, 45 (2011).
22. Rapp, M., Granseth, E., Seppälä, S. & von Heijne, G. Identification and evolution of dual-topology membrane proteins. *Nat. Struct. Mol. Biol.* **13**, 112–116 (2006).
23. Lloris-Garcerá, P. P. et al. Antiparallel dimers of the small multidrug resistance protein EmrE are more stable than parallel dimers. *J. Biol. Chem.* **287**, 26052–26059 (2012).
24. Lloris-Garcerá, P. et al. In vivo Trp scanning of the small multidrug resistance protein EmrE confirms 3D structure models. *J. Mol. Biol.* **425**, 4642–4651 (2013).
25. Amadi, S. T., Koteiche, H. A., Mishra, S. & McHaourab, H. S. Structure, dynamics, and substrate-induced conformational changes of the multidrug transporter EmrE in liposomes. *J. Biol. Chem.* **285**, 26710–26718 (2010).
26. Dastvan, R., Fischer, A. W., Mishra, S., Meiler, J. & McHaourab, H. S. Protonation-dependent conformational dynamics of the multidrug transporter EmrE. *Proc. Natl Acad. Sci. USA* **113**, 1220–1225 (2016).
27. Kermani, A. A., Macdonald, C. B., Gundepudi, R. & Stockbridge, R. B. Guanidinium export is the primal function of SMR family transporters. *Proc. Natl Acad. Sci. USA* **115**, 3060–3065 (2018).
28. Gayen, A., Banigan, J. R. & Traaseth, N. J. Ligand-induced conformational changes of the multidrug resistance transporter EmrE probed by oriented solid-state NMR spectroscopy. *Angew. Chem. Int. Ed.* **52**, 10321–10324 (2013).
29. Dutta, S., Morrison, E. A. & Henzler-Wildman, K. A. Blocking dynamics of the SMR transporter EmrE impairs efflux activity. *Biophys. J.* **107**, 613–620 (2014).
30. Li J., Sae Her A. & Traaseth N. J. Asymmetric protonation of glutamate residues drives a preferred transport pathway in EmrE. *Proc. Natl Acad. Sci. USA* **118**, e2110790118 (2021).
31. Lehner, I. et al. The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric. *J. Biol. Chem.* **283**, 3281–3288 (2008).
32. Ovchinnikov, V., Stone, T. A., Deber, C. M. & Karplus, M. Structure of the EmrE multidrug transporter and its use for inhibitor peptide design. *Proc. Natl Acad. Sci. USA* **115**, E7932 (2018).
33. Vermaas, J. V., Rempe, S. B. & Tajkhorshid, E. Electrostatic lock in the transport cycle of the multidrug resistance transporter EmrE. *Proc. Natl Acad. Sci. USA* **115**, E7502 (2018).
34. Roos, M., Mandala, V. S. & Hong, M. Determination of long-range distances by fast magic-angle-spinning radiofrequency-driven 19F–19F dipolar recoupling NMR. *J. Phys. Chem. B* **122**, 9302–9313 (2018).
35. Roos, M., Wang, T., Shcherbakov, A. A. & Hong, M. Fast magic-angle-spinning F-19 spin exchange NMR for determining nanometer distances in proteins and pharmaceutical compounds. *J. Phys. Chem. B* **122**, 2900–2911 (2018).
36. Shcherbakov, A. A. & Hong, M. Rapid measurement of long-range distances in proteins by multidimensional ^{13}C–^{19}F REDOR NMR under fast magic-angle spinning. *J. Biomol. NMR* **71**, 31–43 (2018).

37. Shcherbakov, A. A., Mandala, V. S. & Hong, M. High-sensitivity detection of nanometer 1H–19F distances for protein structure determination by 1H-detected fast MAS NMR. *J. Phys. Chem. B* **123**, 4387–4391 (2019).

38. Shcherbakov, A. A., Roos, M., Kwon, B. & Hong, M. Two-dimensional 19F–13C correlation NMR for 19F resonance assignment of fluorinated proteins. *J. Biomol. NMR* **74**, 193–204 (2020).

39. Shcherbakov, A. A. et al. Structure and dynamics of the drug-bound bacterial transporter EmrE in lipid bilayers. *Nat. Commun.* **12**, 172 (2021).

40. Yerushalmi, H. & Schuldiner, S. An essential glutamyl residue in EmrE, a multidrug antiporter from Escherichia coli. *J. Biol. Chem.* **275**, 5264–5269 (2000).

41. Yerushalmi, H., Mordoch, S. S. & Schuldiner, S. A single carboxyl mutant of the multidrug transporter EmrE is fully functional. *J. Biol. Chem.* **276**, 12744–12748 (2001).

42. Muth, T. R. & Schuldiner, S. A membrane-embedded glutamate is required for ligand binding to the multidrug transporter EmrE. *EMBO J.* **19**, 234–240 (2000).

43. Brill, S., Falk, O. S. & Schuldiner, S. Transforming a drug/H+ antiporter into a polyamine importer by a single mutation. *Proc. Natl Acad. Sci. USA* **109**, 16894–16899 (2012).

44. Brill, S., Sade-Falk, O., Elbaz-Alon, Y. & Schuldiner, S. Specificity determinants in small multidrug transporters. *J. Mol. Biol.* **427**, 468–477 (2015).

45. Wang, J., Rath, A. & Deber, C. M. Functional response of the small multidrug resistance protein EmrE to mutations in transmembrane helix 2. *FEBS Lett.* **588**, 3720–3725 (2014).

46. Elbaz, Y., Salomon, T. & Schuldiner, S. Identification of a glycine motif required for packing in EmrE, a multidrug transporter from Escherichia coli. *J. Biol. Chem.* **283**, 12276–12283 (2008).

47. Wilks, J. C. & Slonczewski, J. L. pH of the cytoplasm and periplasm of Escherichia coli: Rapid measurement by green fluorescent protein fluorimetry. *J. Bacteriol.* **189**, 5601–5607 (2007).

48. Slonczewski, J. L., Rosen, B. P., Alger, J. R. & Macnab, R. M. pH homeostasis in Escherichia coli: Measurement by 31P nuclear magnetic resonance of methylphosphonate and phosphate. *Proc. Natl. Sci. USA* **78**, 6271–6275 (1981).

49. Morrison, E. A. & Henzler-Wildman, K. A. Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile. *Biochim. Biophys. Acta* **1818**, 814–820 (2012).

50. Hong, M. & Schmidt-Rohr, K. Magic-angle-spinning NMR techniques for measuring long-range distances in biological macromolecules. *Acc. Chem. Res.* **46**, 2154–2163 (2013).

51. Morrison, E. A. & Henzler-Wildman, K. A. Transported substrate determines exchange rate in the multidrug resistance transporter EmrE. *J. Biol. Chem.* **289**, 6825–6836 (2014).

52. Delaglio, F. et al. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

53. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins: Struct., Funct., Bioinform.* **59**, 687–696 (2005).

54. Shrager, R. I., Cohen, J. S., Heller, S. R., Sachs, D. H. & Schechter, A. N. Mathematical models for interacting groups in nuclear magnetic resonance titration curves. *Biochemistry* **11**, 541–547 (1972).

55. McIntosh, L. P. et al. Dissecting electrostatic interactions in Bacillus circulans xylanase through NMR-monitored pH titrations. *J. Biomol. NMR* **51**, 5–19 (2011).

56. Cavanagh J., Fairbrother W. J., Palmer A. G. III & Skelton N. J. *Protein NMR Spectroscopy: Principles and Practice* (Academic Press, 1996).

57. Vasa, S. K., Singh, H., Rovó, P. & Linser, R. Dynamics and interactions of a 29 kDa human enzyme studied by solid-state NMR. *J. Phys. Chem. Lett.* **9**, 1307–1311 (2018).

58. Baldus, M., Petkova, A. T., Herzfeld, J. & Griffin, R. G. Cross polarization in the tilted frame: Assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **95**, 1197–1207 (1998).

59. Verel, R., Ernst, M. & Meier, B. H. Adiabatic dipolar recoupling in solid-state NMR: The DREAM scheme. *J. Magn. Reson.* **150**, 81–99 (2001).

60. Shaka, A. J., Keeler, J., Frenkiel, T. & Freeman, R. An improved sequence for broadband decoupling: WALTZ-16. *J. Magn. Reson.* **52**, 335–338 (1983).

61. Zhou, D. H. & Rienstra, C. M. High-performance solvent suppression for proton detected solid-state NMR. *J. Magn. Reson.* **192**, 167–172 (2008).

62. Dürr, H. N., Grage, S. L., Witter, R. & Ulrich, A. S. Solid state 19F NMR parameters of fluorine-labeled amino acids. Part I: Aromatic substituents. *J. Magn. Reson.* **191**, 7–15 (2008).

63. Walt, Svd, Colbert, S. C. & Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).

64. Helmus, J. J. & Jaroniec, C. P. Nmrglue: An open source Python package for the analysis of multidimensional NMR data. *J. Biomol. NMR* **55**, 355–367 (2013).

65. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2014).

66. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

67. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).

68. Virtanen, P. et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

69. Öster, C., Kosol, S. & Lewandowski, J. R. Quantifying microsecond exchange in large protein complexes with accelerated relaxation dispersion experiments in the solid state. *Sci. Rep.* **9**, 11082 (2019).

70. Bak, M., Rasmussen, J. T. & Nielsen, N. C. SIMPSON: A general simulation program for solid-state NMR spectroscopy. *J. Magn. Reson.* **147**, 296–330 (2000).

71. Bak, M. & Nielsen, N. C. REPULSION, a novel approach to efficient powder averaging in solid-state NMR. *J. Magn. Reson.* **125**, 132–139 (1997).

72. Mandala, V. S., Liao, S. Y., Kwon, B. & Hong, M. Structural basis for asymmetric conductance of the influenza M2 proton channel investigated by solid-state NMR spectroscopy. *J. Mol. Biol.* **429**, 2192–2210 (2017).

73. Dregni, A. J., Duan, P. & Hong, M. Hydration and dynamics of full-length Tau amyloid fibrils investigated by solid-state nuclear magnetic resonance. *Biochemistry* **59**, 2237–2248 (2020).

74. Mandala, V. S., Loftis, A. R., Shcherbakov, A. A., Pentelute, B. L. & Hong, M. Atomic structures of closed and open influenza B M2 proton channel reveal the conduction mechanism. *Nat. Struct. Mol. Biol.* **27**, 160–167 (2020).

75. van Zundert, G. C. P. et al. The HADDOCK2.2 Web Server: User-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* **428**, 720–725 (2016).

76. de Vries, S. J., van Dijk, M. & Bonvin, A. M. J. J. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **5**, 883–897 (2010).

77. Sharoni, M., Steiner-Mordoch, S. & Schuldiner, S. Exploring the binding domain of EmrE, the smallest multidrug transporter. *J. Biol. Chem.* **280**, 32849–32855 (2005).

78. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).

79. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

80. 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).

81. Jo, S., Kim, T. & Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One* **2**, e880 (2007).

82. Wu, E. L. et al. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).

83. Mark, P. & Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **105**, 9954–9960 (2001).

84. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

85. Maciejewski, M. W. et al. NMRbox: A resource for biomolecular NMR computation. *Biophys. J.* **112**, 1529–1534 (2017).

86. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

87. 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).

88. 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).

89. Huang, J. et al. CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).

90. Khan, H. M., MacKerell, A. D. & Reuter, N. Cation–π interactions between methylated ammonium groups and tryptophan in the CHARMM36 additive force field. *J. Chem. Theory Comput.* **15**, 7–12 (2019).

Author contributions

M.H. and K.A.H.-W. designed the experiments. A.A.S. conducted the solid-state NMR experiments, data analysis, and structure calculation; P.S. purified the protein and

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-28556-6.

Correspondence and requests for materials should be addressed to Mei Hong.

Peer review information *Nature Communications* thanks Markus Weingarth and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Supplementary Information

High-pH Structure of EmrE Reveals the Mechanism of Proton-Coupled Substrate Transport

Alexander A. Shcherbakov [1], Peyton J. Spreacker [2], Aurelio J. Dregni [1], Katherine A. Henzler-Wildman [2], and Mei Hong [1]*

[1] Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139

[2] Department of Biochemistry, University of Wisconsin at Madison, Madison, WI 53706

This PDF file includes:
>   Tables S1-S6
>   Supplementary Figures 1 to 11

Supplementary Table 1. Explanation of the eight ^{1}H-detected 3D correlation MAS NMR experiments for assigning the chemical shifts of high-pH EmrE in DMPC bilayers. OaB = Out-and-Back. All H-N and H-C transfers used dipolar cross polarization.

Experiment	F1-F2-F3	Connectivity	Polarization transfer
hCANH	$CA_i – N_i – H^N_i$	Intra-residue	Dipolar transfer for CA-N
hcaCBcaNH	$CB_i – N_i – H^N_i$	Intra-residue	OaB INEPT for CB-CA, dipolar CA-N
hCOcaNH	$CO_i – N_i – H^N_i$	Intra-residue	Dipolar CO-CA, CA-N
hCAcoNH	$CA_{i-1} – N_i – H^N_i$	Inter-residue	Dipolar CA-CO and CO-N
hCONH	$CO_{i-1} – N_i – H^N_i$	Inter-residue	Dipolar CO-N
hcaCBcacoNH	$CB_{i-1} – N_i – H^N_i$	Inter-residue	OaB INEPT for CB-CA, dipolar CA-CO, CO-N
hNcacoNH	$N_{i-1} – N_i – H^N_i$	Inter-residue	Dipolar N-CA, CA-CO, CO-N
HncacoNH	$H^N_{i-1} – N_i – H^N_i$	Inter-residue	Dipolar N-CA, CA-CO, CO-N

Supplementary Table 2. Detailed H^N-F distance constraints between EmrE and F$_4$-TPP$^+$ used for HADDOCK docking of the ligand into the protein and for structure-based assignment of the H^N-^{19}F pairs.

Residue	Atom	r_{HF} (Å)	δ_{low} (Å)	δ_{high} (Å)	Assigned ^{19}F Atoms (Run 2)			
G8A	HN	9.7	2.0	40.0	F3	F13	F18	F8
G9A	HN	12.0	3.8	40.0	F3	F13	F18	F8
A10A	HN	12.0	3.9	40.0	F3	F13	F18	F8
I11A	HN	8.4	3.1	40.0	F3	F13	F18	F8
A13A	HN	8.7	1.3	40.0	F3	F13	F18	F8
E14A	HN	7.8	2.8	1.8	F8			
V15A	HN	6.3	2.5	0.5	F8			
I16A	HN	9.3	1.6	40.0	F3	F13	F18	F8
G17A	HN	7.2	2.7	1.1	F8			
T18A	HN	7.7	0.8	2.2	F8			
T19A	HN	8.7	2.3	40.0	F3	F13	F18	F8
L20A	HN	9.2	2.6	40.0	F3	F13	F18	F8
M21A	HN	9.6	2.9	40.0	F3	F13	F18	F8
K22A	HN	12.0	3.8	40.0	F3	F13	F18	F8
F23A	HN	12.0	3.8	40.0	F3	F13	F18	F8
S24A	HN	10.7	3.8	40.0	F3	F13	F18	F8
E25A	HN	12.0	3.8	40.0	F3	F13	F18	F8
G26A	HN	12.0	3.8	40.0	F3	F13	F18	F8
F27A	HN	12.0	3.8	40.0	F3	F13	F18	F8
R29A	HN	11.5	3.4	40.0	F3	F13	F18	F8
L30A	HN	11.9	3.9	40.0	F3	F13	F18	F8
W31A	HN	11.3	3.3	40.0	F3	F13	F18	F8
S33A	HN	10.3	2.5	40.0	F3	F13	F18	F8
V34A	HN	10.6	2.6	40.0	F3	F13	F18	F8
G35A	HN	12.0	4.9	40.0	F3	F13	F18	F8
C39A	HN	8.9	3.4	40.0	F3	F13	F18	F8
Y40A	HN	7.3	2.6	1.1	F8			
C41A	HN	12.0	5.9	40.0	F3	F13	F18	F8
S43A	HN	4.3	2.3	1.3	F8			
F44A	HN	4.8	1.8	0.8	F8			
Q49A	HN	10.8	3.9	40.0	F3	F13	F18	F8
G57A	HN	8.1	3.0	40.0	F3	F13	F18	F8
I58A	HN	11.7	3.6	40.0	F3	F13	F18	H^{N}F8
A59A	HN	9.5	3.8	40.0	F3	F13	F18	F8
Y60A	HN	6.3	2.7	0.5	F3			
I62A	HN	8.2	3.1	40.0	F3	F13	F18	F8
W63A	HN	8.2	3.1	40.0	F3	F13	F18	F8
V64A	HN	6.2	2.7	0.5	F3			
G65A	HN	6.7	0.6	0.7	F3			
V66A	HN	10.3	4.4	40.0	F3	F13	F18	F8
G67A	HN	5.9	0.7	0.7	F3			
I68A	HN	7.2	2.6	0.9	F3			
V69A	HN	9.7	4.0	40.0	F3	F13	F18	F8
I71A	HN	8.8	1.3	40.0	F3	F13	F18	F8
S72A	HN	12.0	5.9	40.0	F3	F13	F18	F8
L74A	HN	11.4	3.4	40.0	F3	F13	F18	F8
S75A	HN	7.9	2.9	2.7	F18			
W76A	HN	12.0	4.8	40.0	F3	F13	F18	F8
G77A	HN	9.4	1.7	40.0	F3	F13	F18	F8
F79A	HN	8.7	2.2	40.0	F3	F13	F18	F8
G80A	HN	11.1	3.1	40.0	F3	F13	F18	F8
Q81A	HN	11.8	3.8	40.0	F3	F13	F18	F8

R82A	HN	12.0	3.9	40.0	F3	F13	F18	F8
L83A	HN	11.5	3.5	40.0	F3	F13	F18	F8
D84A	HN	9.5	1.8	40.0	F3	F13	F18	F8
A87A	HN	8.6	1.3	40.0	F3	F13	F18	F8
G90A	HN	11.6	3.6	40.0	F3	F13	F18	F8
C95A	HN	10.9	2.9	40.0	F3	F13	F18	F8
G97A	HN	12.0	3.8	40.0	F3	F13	F18	F8
L103A	HN	8.8	1.4	40.0	F3	F13	F18	F8
L104A	HN	12.0	3.9	40.0	F3	F13	F18	F8
S105A	HN	12.0	3.8	40.0	-			
R106A	HN	10.7	2.7	40.0	-			
Y6B	HN	9.2	1.6	40.0	F3	F13	F18	F8
G8B	HN	10.2	2.4	40.0	F3	F13	F18	F8
G9B	HN	7.3	2.8	1.2	F13			
A10B	HN	8.3	1.1	40.0	F3	F13	F18	F8
I11B	HN	8.5	3.2	40.0	F3	F13	F18	F8
G17B	HN	7.6	2.8	1.8	F13			
T18B	HN	7.4	0.4	0.6	F13			
T19B	HN	11.3	3.2	40.0	F3	F13	F18	F8
K22B	HN	12.0	3.8	40.0	F3	F13	F18	F8
E25B	HN	11.2	3.2	40.0	F3	F13	F18	F8
G26B	HN	10.4	2.5	40.0	F3	F13	F18	F8
R29B	HN	10.1	2.3	40.0	F3	F13	F18	F8
L30B	HN	11.0	3.0	40.0	F3	F13	F18	F8
W31B	HN	9.8	2.1	40.0	F3	F13	F18	F8
S33B	HN	10.1	2.3	40.0	F3	F13	F18	F8
V34B	HN	12.0	3.9	40.0	F3	F13	F18	F8
G35B	HN	12.0	3.8	40.0	F3	F13	F18	F8
S43B	HN	6.9	2.7	0.8	F13			
L47B	HN	8.2	3.0	40.0	F3	F13	F18	F8
Q49B	HN	9.6	3.9	40.0	F3	F13	F18	F8
L51B	HN	9.0	1.5	40.0	F3	F13	F18	F8
A52B	HN	11.2	3.2	40.0	F3	F13	F18	F8
Y53B	HN	9.6	1.9	40.0	F3	F13	F18	F8
G57B	HN	11.9	3.8	40.0	F3	F13	F18	F8
I58B	HN	11.8	3.7	40.0	F3	F13	F18	F8
A59B	HN	9.4	1.8	40.0	F3	F13	F18	F8
Y60B	HN	6.6	2.5	0.6	F18			
A61B	HN	9.9	3.1	40.0	F3	F13	F18	F8
I62B	HN	9.3	3.6	40.0	F3	F13	F18	F8
W63B	HN	9.2	3.7	40.0	F3	F13	F18	F8
V66B	HN	8.5	2.2	40.0	F3	F13	F18	F8
G67B	HN	9.1	1.6	40.0	F3	F13	F18	F8
I68B	HN	9.0	2.5	40.0	F3	F13	F18	F8
V69B	HN	9.8	2.0	40.0	F3	F13	F18	F8
S75B	HN	11.5	3.4	40.0	F3	F13	F18	F8
W76B	HN	9.1	1.5	40.0	F3	F13	F18	F8
G77B	HN	12.0	3.9	40.0	F3	F13	F18	F8
F78B	HN	12.0	3.8	40.0	F3	F13	F18	F8
F79B	HN	12.0	3.8	40.0	F3	F13	F18	F8
G80B	HN	10.6	2.7	40.0	F3	F13	F18	F8
Q81B	HN	10.0	2.2	40.0	F3	F13	F18	F8
R82B	HN	12.0	3.8	40.0	F3	F13	F18	F8
L83B	HN	11.6	3.6	40.0	F3	F13	F18	F8
L85B	HN	12.0	3.9	40.0	F3	F13	F18	F8
A87B	HN	12.0	3.8	40.0	F3	F13	F18	F8

G90B	HN	12.0	3.9	40.0	F3	F13	F18	F8
G97B	HN	12.0	3.8	40.0	F3	F13	F18	F8
L103B	HN	12.0	3.8	40.0	F3	F13	F18	F8
L104B	HN	11.2	3.1	40.0	F3	F13	F18	F8
S105B	HN	10.1	2.3	40.0	-			
R106B	HN	12.0	3.9	40.0	-			
W63B	HE1	3.8	3.8	2.0	F13			

Supplementary Table 3. Distance violation statistics for the 25 lowest-violation structures from HADDOCK structure calculation and detailed violation analysis for two complexes selected for further refinement in MD simulations. The violation detail tables indicate distances from specified protein atoms to each of the four ligand fluorines, to score against the original four-fold ambiguous dataset.

Output Model	Number of Violations	Average Violation (Å)	Std. Dev. Violation (Å)	Sum Total of Violations (Å)
complex_167w	6	0.4	0.2	2.3
complex_189w	6	0.4	0.2	2.3
complex_165w	7	0.3	0.2	2.3
complex_191w	7	0.3	0.2	2.3
complex_161w	6	0.4	0.2	2.4
complex_175w	6	0.4	0.2	2.4
complex_183w	8	0.3	0.2	2.4
complex_180w	6	0.4	0.1	2.5
complex_174w	7	0.4	0.2	2.5
complex_172w	6	0.4	0.2	2.6
complex_179w	7	0.4	0.2	2.6
complex_160w	7	0.4	0.2	2.7
complex_158w	10	0.3	0.1	2.8
complex_176w	7	0.4	0.4	2.9
complex_159w	8	0.4	0.3	3.2
complex_169w	7	0.5	0.6	3.3
complex_166w	6	0.6	0.8	3.5
complex_154w	9	0.4	0.2	3.5
complex_171w	7	0.5	0.6	3.7
complex_170w	7	0.5	0.5	3.8
complex_95w	7	0.6	0.3	3.9
complex_163w	5	0.8	1.3	4.1
complex_148w	6	0.7	1.1	4.1
complex_49w	7	0.6	0.4	4.1
complex_177w	7	0.6	0.8	4.1

Violation Details for complex_167w (lowest sum total of violations)

Residue	Atom	Restraint Range (Å)	Model Distances (Å)	Violation Magnitude (Å)
V15A	H^N	3.8 - 6.8	9.0, 14.7, 7.2, 11.8	0.4
G67A	H^N	5.2 - 6.6	10.4, 15.4, 11.9, 6.9	0.3
S75A	H^N	5.0 - 10.6	11.0, 20.9, 17.2, 16.5	0.4
G9B	H^N	4.5 - 8.5	16.9, 8.7, 18.1, 16.1	0.2
T18B	H^N	7.0 - 8.0	16.6, 8.7, 10.9, 11.1	0.7
Y60B	H^N	4.1 - 7.2	7.5, 10.5, 15.4, 11.1	0.3

Violation Details for complex_163w (lowest number of violations)

Residue	Atom	Restraint Range (Å)	Model Distances (Å)	Violation Magnitude (Å)
V15A	H^N	3.8 – 6.8	11.7, 6.9, 15.6, 11.2	0.1
Y60A	H^N	3.6 - 6.8	6.9, 13.4, 13.1, 16.4	0.1
G67A	H^N	5.2 - 6.6	6.8, 11.9, 16.1, 12.6	0.2
S75A	H^N	5.0 - 10.6	16.3, 17.3, 22.5, 14.0	3.4
G9B	H^N	4.5 - 8.5	16.0, 18.3, 8.8, 14.8	0.3

Supplementary Table 4. Detailed H^N-F distance constraints between EmrE and F$_4$-TPP$^+$ used in the GROMACS refinement stage of structure calculation.

Constraint Atoms			Constr. Range		Constraint Atoms			Constr. Range	
Res.	Atom	Fluorine	r_0 (Å)	r_1 (Å)	Res.	Atom	Fluorine	r_0 (Å)	r_1 (Å)
G8A	HN	F3	7.7	40	Y6B	HN	F3	7.6	40
G8A	HN	F13	7.7	40	Y6B	HN	F13	7.6	40
G8A	HN	F18	7.7	40	Y6B	HN	F18	7.6	40
G8A	HN	F8	7.7	40	Y6B	HN	F8	7.6	40
G9A	HN	F3	8.2	40	G8B	HN	F3	7.8	40
G9A	HN	F13	8.2	40	G8B	HN	F13	7.8	40
G9A	HN	F18	8.2	40	G8B	HN	F18	7.8	40
G9A	HN	F8	8.2	40	G8B	HN	F8	7.8	40
A10A	HN	F3	8.1	40	G9B	HN	F13	4.5	8.5
A10A	HN	F13	8.1	40	A10B	HN	F3	7.2	40
A10A	HN	F18	8.1	40	A10B	HN	F13	7.2	40
A10A	HN	F8	8.1	40	A10B	HN	F18	7.2	40
I11A	HN	F3	5.3	40	A10B	HN	F8	7.2	40
I11A	HN	F13	5.3	40	I11B	HN	F3	5.3	40
I11A	HN	F18	5.3	40	I11B	HN	F13	5.3	40
I11A	HN	F8	5.3	40	I11B	HN	F18	5.3	40
A13A	HN	F3	7.4	40	I11B	HN	F8	5.3	40
A13A	HN	F13	7.4	40	G17B	HN	F13	4.8	9.4
A13A	HN	F18	7.4	40	T18B	HN	F13	7	8
A13A	HN	F8	7.4	40	T19B	HN	F3	8.1	40
E14A	HN	F8	5	9.6	T19B	HN	F13	8.1	40
V15A	HN	F8	3.8	6.8	T19B	HN	F18	8.1	40
I16A	HN	F3	7.7	40	T19B	HN	F8	8.1	40
I16A	HN	F13	7.7	40	K22B	HN	F3	8.2	40
I16A	HN	F18	7.7	40	K22B	HN	F13	8.2	40
I16A	HN	F8	7.7	40	K22B	HN	F18	8.2	40
G17A	HN	F8	4.5	8.3	K22B	HN	F8	8.2	40
T18A	HN	F8	6.9	9.9	E25B	HN	F3	8	40
T19A	HN	F3	6.4	40	E25B	HN	F13	8	40
T19A	HN	F13	6.4	40	E25B	HN	F18	8	40
T19A	HN	F18	6.4	40	E25B	HN	F8	8	40
T19A	HN	F8	6.4	40	G26B	HN	F3	7.9	40
L20A	HN	F3	6.6	40	G26B	HN	F13	7.9	40
L20A	HN	F13	6.6	40	G26B	HN	F18	7.9	40
L20A	HN	F18	6.6	40	G26B	HN	F8	7.9	40
L20A	HN	F8	6.6	40	R29B	HN	F3	7.8	40
M21A	HN	F3	6.7	40	R29B	HN	F13	7.8	40
M21A	HN	F13	6.7	40	R29B	HN	F18	7.8	40
M21A	HN	F18	6.7	40	R29B	HN	F8	7.8	40
M21A	HN	F8	6.7	40	L30B	HN	F3	8	40
K22A	HN	F3	8.2	40	L30B	HN	F13	8	40
K22A	HN	F13	8.2	40	L30B	HN	F18	8	40
K22A	HN	F18	8.2	40	L30B	HN	F8	8	40
K22A	HN	F8	8.2	40	W31B	HN	F3	7.7	40
F23A	HN	F3	8.2	40	W31B	HN	F13	7.7	40
F23A	HN	F13	8.2	40	W31B	HN	F18	7.7	40
F23A	HN	F18	8.2	40	W31B	HN	F8	7.7	40
F23A	HN	F8	8.2	40	S33B	HN	F3	7.8	40
S24A	HN	F3	6.9	40	S33B	HN	F13	7.8	40
S24A	HN	F13	6.9	40	S33B	HN	F18	7.8	40
S24A	HN	F18	6.9	40	S33B	HN	F8	7.8	40

S24A	HN	F8	6.9	40	V34B	HN	F3	8.1	40
E25A	HN	F3	8.2	40	V34B	HN	F13	8.1	40
E25A	HN	F13	8.2	40	V34B	HN	F18	8.1	40
E25A	HN	F18	8.2	40	V34B	HN	F8	8.1	40
E25A	HN	F8	8.2	40	G35B	HN	F3	8.2	40
G26A	HN	F3	8.2	40	G35B	HN	F13	8.2	40
G26A	HN	F13	8.2	40	G35B	HN	F18	8.2	40
G26A	HN	F18	8.2	40	G35B	HN	F8	8.2	40
G26A	HN	F8	8.2	40	S43B	HN	F13	4.2	7.7
F27A	HN	F3	8.2	40	L47B	HN	F3	5.2	40
F27A	HN	F13	8.2	40	L47B	HN	F13	5.2	40
F27A	HN	F18	8.2	40	L47B	HN	F18	5.2	40
F27A	HN	F8	8.2	40	L47B	HN	F8	5.2	40
R29A	HN	F3	8.1	40	Q49B	HN	F3	5.7	40
R29A	HN	F13	8.1	40	Q49B	HN	F13	5.7	40
R29A	HN	F18	8.1	40	Q49B	HN	F18	5.7	40
R29A	HN	F8	8.1	40	Q49B	HN	F8	5.7	40
L30A	HN	F3	8	40	L51B	HN	F3	7.5	40
L30A	HN	F13	8	40	L51B	HN	F13	7.5	40
L30A	HN	F18	8	40	L51B	HN	F18	7.5	40
L30A	HN	F8	8	40	L51B	HN	F8	7.5	40
W31A	HN	F3	8	40	A52B	HN	F3	8	40
W31A	HN	F13	8	40	A52B	HN	F13	8	40
W31A	HN	F18	8	40	A52B	HN	F18	8	40
W31A	HN	F8	8	40	A52B	HN	F8	8	40
S33A	HN	F3	7.8	40	Y53B	HN	F3	7.7	40
S33A	HN	F13	7.8	40	Y53B	HN	F13	7.7	40
S33A	HN	F18	7.8	40	Y53B	HN	F18	7.7	40
S33A	HN	F8	7.8	40	Y53B	HN	F8	7.7	40
V34A	HN	F3	8	40	G57B	HN	F3	8.1	40
V34A	HN	F13	8	40	G57B	HN	F13	8.1	40
V34A	HN	F18	8	40	G57B	HN	F18	8.1	40
V34A	HN	F8	8	40	G57B	HN	F8	8.1	40
G35A	HN	F3	7.1	40	I58B	HN	F3	8.1	40
G35A	HN	F13	7.1	40	I58B	HN	F13	8.1	40
G35A	HN	F18	7.1	40	I58B	HN	F18	8.1	40
G35A	HN	F8	7.1	40	I58B	HN	F8	8.1	40
C39A	HN	F3	5.5	40	A59B	HN	F3	7.6	40
C39A	HN	F13	5.5	40	A59B	HN	F13	7.6	40
C39A	HN	F18	5.5	40	A59B	HN	F18	7.6	40
C39A	HN	F8	5.5	40	A59B	HN	F8	7.6	40
Y40A	HN	F8	4.7	8.4	Y60B	HN	F18	4.1	7.2
C41A	HN	F3	6.1	40	A61B	HN	F3	6.8	40
C41A	HN	F13	6.1	40	A61B	HN	F13	6.8	40
C41A	HN	F18	6.1	40	A61B	HN	F18	6.8	40
C41A	HN	F8	6.1	40	A61B	HN	F8	6.8	40
S43A	HN	F8	2	5.6	I62B	HN	F3	5.7	40
F44A	HN	F8	3	5.6	I62B	HN	F13	5.7	40
Q49A	HN	F3	6.9	40	I62B	HN	F18	5.7	40
Q49A	HN	F13	6.9	40	I62B	HN	F8	5.7	40
Q49A	HN	F18	6.9	40	W63B	HN	F3	5.5	40
Q49A	HN	F8	6.9	40	W63B	HN	F13	5.5	40
G57A	HN	F3	5.1	40	W63B	HN	F18	5.5	40
G57A	HN	F13	5.1	40	W63B	HN	F8	5.5	40
G57A	HN	F18	5.1	40	V66B	HN	F3	6.3	40
G57A	HN	F8	5.1	40	V66B	HN	F13	6.3	40

I58A	HN	F3	8.1	40	V66B	HN	F18	6.3	40
I58A	HN	F13	8.1	40	V66B	HN	F8	6.3	40
I58A	HN	F18	8.1	40	G67B	HN	F3	7.5	40
I58A	HN	F8	8.1	40	G67B	HN	F13	7.5	40
A59A	HN	F3	5.7	40	G67B	HN	F18	7.5	40
A59A	HN	F13	5.7	40	G67B	HN	F8	7.5	40
A59A	HN	F18	5.7	40	I68B	HN	F3	6.5	40
A59A	HN	F8	5.7	40	I68B	HN	F13	6.5	40
Y60A	HN	F3	3.6	6.8	I68B	HN	F18	6.5	40
I62A	HN	F3	5.1	40	I68B	HN	F8	6.5	40
I62A	HN	F13	5.1	40	V69B	HN	F3	7.8	40
I62A	HN	F18	5.1	40	V69B	HN	F13	7.8	40
I62A	HN	F8	5.1	40	V69B	HN	F18	7.8	40
W63A	HN	F3	5.1	40	V69B	HN	F8	7.8	40
W63A	HN	F13	5.1	40	S75B	HN	F3	8.1	40
W63A	HN	F18	5.1	40	S75B	HN	F13	8.1	40
W63A	HN	F8	5.1	40	S75B	HN	F18	8.1	40
V64A	HN	F3	3.5	6.7	S75B	HN	F8	8.1	40
G65A	HN	F3	6.1	7.4	W76B	HN	F3	7.6	40
V66A	HN	F3	5.9	40	W76B	HN	F13	7.6	40
V66A	HN	F13	5.9	40	W76B	HN	F18	7.6	40
V66A	HN	F18	5.9	40	W76B	HN	F8	7.6	40
V66A	HN	F8	5.9	40	G77B	HN	F3	8.1	40
G67A	HN	F3	5.2	6.6	G77B	HN	F13	8.1	40
I68A	HN	F3	4.6	8.1	G77B	HN	F18	8.1	40
V69A	HN	F3	5.7	40	G77B	HN	F8	8.1	40
V69A	HN	F13	5.7	40	F78B	HN	F3	8.2	40
V69A	HN	F18	5.7	40	F78B	HN	F13	8.2	40
V69A	HN	F8	5.7	40	F78B	HN	F18	8.2	40
I71A	HN	F3	7.5	40	F78B	HN	F8	8.2	40
I71A	HN	F13	7.5	40	F79B	HN	F3	8.2	40
I71A	HN	F18	7.5	40	F79B	HN	F13	8.2	40
I71A	HN	F8	7.5	40	F79B	HN	F18	8.2	40
S72A	HN	F3	6.1	40	F79B	HN	F8	8.2	40
S72A	HN	F13	6.1	40	G80B	HN	F3	7.9	40
S72A	HN	F18	6.1	40	G80B	HN	F13	7.9	40
S72A	HN	F8	6.1	40	G80B	HN	F18	7.9	40
L74A	HN	F3	8	40	G80B	HN	F8	7.9	40
L74A	HN	F13	8	40	Q81B	HN	F3	7.8	40
L74A	HN	F18	8	40	Q81B	HN	F13	7.8	40
L74A	HN	F8	8	40	Q81B	HN	F18	7.8	40
S75A	HN	F18	5	10.6	Q81B	HN	F8	7.8	40
W76A	HN	F3	7.2	40	R82B	HN	F3	8.2	40
W76A	HN	F13	7.2	40	R82B	HN	F13	8.2	40
W76A	HN	F18	7.2	40	R82B	HN	F18	8.2	40
W76A	HN	F8	7.2	40	R82B	HN	F8	8.2	40
G77A	HN	F3	7.7	40	L83B	HN	F3	8	40
G77A	HN	F13	7.7	40	L83B	HN	F13	8	40
G77A	HN	F18	7.7	40	L83B	HN	F18	8	40
G77A	HN	F8	7.7	40	L83B	HN	F8	8	40
F79A	HN	F3	6.5	40	L85B	HN	F3	8.1	40
F79A	HN	F13	6.5	40	L85B	HN	F13	8.1	40
F79A	HN	F18	6.5	40	L85B	HN	F18	8.1	40
F79A	HN	F8	6.5	40	L85B	HN	F8	8.1	40
G80A	HN	F3	8	40	A87B	HN	F3	8.2	40
G80A	HN	F13	8	40	A87B	HN	F13	8.2	40

G80A	HN	F18	8	40		A87B	HN	F18	8.2	40
G80A	HN	F8	8	40		A87B	HN	F8	8.2	40
Q81A	HN	F3	8	40		G90B	HN	F3	8.1	40
Q81A	HN	F13	8	40		G90B	HN	F13	8.1	40
Q81A	HN	F18	8	40		G90B	HN	F18	8.1	40
Q81A	HN	F8	8	40		G90B	HN	F8	8.1	40
R82A	HN	F3	8.1	40		G97B	HN	F3	8.2	40
R82A	HN	F13	8.1	40		G97B	HN	F13	8.2	40
R82A	HN	F18	8.1	40		G97B	HN	F18	8.2	40
R82A	HN	F8	8.1	40		G97B	HN	F8	8.2	40
L83A	HN	F3	8	40		L103B	HN	F3	8.2	40
L83A	HN	F13	8	40		L103B	HN	F13	8.2	40
L83A	HN	F18	8	40		L103B	HN	F18	8.2	40
L83A	HN	F8	8	40		L103B	HN	F8	8.2	40
D84A	HN	F3	7.7	40		L104B	HN	F3	8.1	40
D84A	HN	F13	7.7	40		L104B	HN	F13	8.1	40
D84A	HN	F18	7.7	40		L104B	HN	F18	8.1	40
D84A	HN	F8	7.7	40		L104B	HN	F8	8.1	40
A87A	HN	F3	7.3	40		W63B	HE1	F13	0	5.8
A87A	HN	F13	7.3	40						
A87A	HN	F18	7.3	40						
A87A	HN	F8	7.3	40						
G90A	HN	F3	8	40						
G90A	HN	F13	8	40						
G90A	HN	F18	8	40						
G90A	HN	F8	8	40						
C95A	HN	F3	8	40						
C95A	HN	F13	8	40						
C95A	HN	F18	8	40						
C95A	HN	F8	8	40						
G97A	HN	F3	8.2	40						
G97A	HN	F13	8.2	40						
G97A	HN	F18	8.2	40						
G97A	HN	F8	8.2	40						
L103A	HN	F3	7.4	40						
L103A	HN	F13	7.4	40						
L103A	HN	F18	7.4	40						
L103A	HN	F8	7.4	40						
L104A	HN	F3	8.1	40						
L104A	HN	F13	8.1	40						
L104A	HN	F18	8.1	40						
L104A	HN	F8	8.1	40						

Supplementary Table 5. Distance constraint violation statistics for the 10 lowest-violation structures from the GROMACS refinement stage of structure calculation and detailed violation analysis for two complexes selected from the two MD runs. Violation detail tables indicate distances from specified protein atoms to each of the four ligand fluorines, to score against the original four-fold ambiguous data set.

Ensemble Conformer #	MD Timepoint (ns)	MD Run #	Number of Violations	Average Violation (Å)	Std. Dev. Violation (Å)	Sum Total of Violations (Å)
1	280	2	4	0.2	0.1	0.8
2	200	2	5	0.3	0.2	1.5
3	210	2	5	0.3	0.3	1.7
4	335	2	4	0.4	0.2	1.7
5	330	1	4	0.4	0.2	1.8
6	335	1	4	0.5	0.3	1.8
7	355	1	5	0.4	0.1	1.8
8	260	2	3	0.6	0.2	1.9
9	385	1	4	0.5	0.2	1.9
10	315	1	5	0.4	0.2	1.9

Violation Details for Conformer 1 (Run 2, 280 ns timepoint)

Residue	Atom	Restraint Range (Å)	Model Distances (Å)	Violation Magnitude (Å)
V15A	H^N	3.8 - 6.8	10.9, 7.0, 15.0, 9.6	0.2
Y60A	H^N	3.6 – 6.8	6.9, 14.1, 13.9, 16.6	0.1
G9B	H^N	4.5 – 8.5	16.8, 18.4, 8.9, 14.8	0.4
T18B	H^N	7.0 – 8.0	10.8, 13.7, 8.1, 16.6	0.1

Violation Details for Conformer 5 (Run 1, 330 ns timepoint)

Residue	Atom	Restraint Range (Å)	Model Distances (Å)	Violation Magnitude (Å)
V15A	H^N	3.8 - 6.8	12.6, 16.8, 7.3, 13.5	0.5
Y60A	H^N	3.6 – 6.8	17.0, 14.4, 13.4, 7.0	0.2
G9B	H^N	4.5 – 8.5	15.1, 9.2, 18.3, 17.7	0.7
T18B	H^N	7.0 – 8.0	17.8, 8.4, 13.7, 13.6	0.4

Supplementary Table 6. Detailed parameters for the solid-state NMR experiments for resonance assignment and distance measurements of F_4-TPP$^+$ bound CDN-labeled EmrE in DMPC bilayers at pH 8.0. The protein : lipid molar ratio is 1 : 25.

| Experiment	NMR Parameters	Expt. Time
^{1}H^{15}N solution TROSY-HSQC	B_0 = 17.6 T; T_{eff} = 318 K; ns = 16 (32 for pH 8.4); d1 = 2 s	2 hr * 3+ 4 hr = 10 hr
2D hNH	B_0 = 14.1 T; 1.3 mm HCN; T_{eff} = 280 K (T_{set} = 248 K); v_{MAS} = 55 kHz; ns = 48; τ_{rd} = 1.2 s; τ_{CP-HN} = 1.2 ms; $v_{1H-CP,HN}$ = 93.5 kHz; $v_{15N-CP,HN}$ = 38.5 kHz; $t_{1,max}$ = 45 ms; $t_{1,inc}$ = 300.0 μs; $v_{1H-DD,t1evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-NH} = 1.0 ms; $v_{1H-CP,NH}$ = 93.5 kHz; $v_{15N-CP,NH}$ = 38.5 kHz; τ_{dwell} = 10.0 μs; τ_{acq} = 20.48; $v_{15N-DD,acq}$ = 10 kHz; $v_{13C-DD,acq}$ = 10 kHz	6 hr
3D hCANH	B_0 = 14.1 T; 1.3 mm HCN; T_{eff} = 280 K (T_{set} = 248 K); v_{MAS} = 55 kHz; ns = 24 (8 * 3 blocks); τ_{rd} = 1.8 s; τ_{CP-HC} = 1.1 ms; $v_{1H-CP,HC}$ = 90.0 kHz; $v_{13C-CP,HC}$ = 35 kHz; $t_{1,max}$ = 4.32 ms; $t_{1,inc}$ = 160.0 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{DCP-CN} = 7.4 ms; $v_{13C-DCP}$ = 35.0 kHz; $v_{15N-DCP}$ = 20.0 kHz; $v_{1H-DD,DCP}$ = 10 kHz; $t_{2,max}$ = 10.5 ms; $t_{2,inc}$ = 300 μs; $v_{1H-DD, t2evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-NH} = 1.0 ms; $v_{1H-CP,NH}$ = 93.5 kHz; $v_{15N-CP,NH}$ = 38.5 kHz; τ_{dwell} = 10.0 μs; τ_{acq} = 20.48 ms; $v_{15N-DD,acq}$ = 10 kHz; $v_{13C-DD,acq}$ = 10 kHz	17 hr* 3 = 51 hr
3D hCOcaNH	B_0 = 14.1 T; 1.3 mm HCN; T_{eff} = 280 K (T_{set} = 248 K); v_{MAS} = 55 kHz; ns = 40 (16 + 24); τ_{rd} = 1.4 s; τ_{CP-HC} = 1.4 ms; $v_{1H-CP,HC}$ = 93.5 kHz; $v_{13C-CP,HC}$ = 38.5 kHz; $t_{1,max}$ = 5.0 ms; $t_{1,inc}$ = 250.0 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{DREAM} = 6.5 ms; τ_{DCP-CN} = 5.0 ms; $v_{13C-DCP}$ = 20.0 kHz; $v_{15N-DCP}$ = 35.0 kHz; $v_{1H-DD,DCP}$ = 10 kHz; $t_{2,max}$ = 10.5 ms; $t_{2,inc}$ = 300.0 μs; $v_{1H-DD, t2evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-NH} = 0.70 ms; $v_{1H-CP,NH}$ = 93.5 kHz; $v_{15N-CP,NH}$ = 38.5 kHz; τ_{dwell} = 10.0 μs; τ_{acq} = 20.48 ms; $v_{15N-DD,acq}$ = 10 kHz; $v_{13C-DD,acq}$ = 10 kHz	20 + 30 = 50 hr
3D hCONH	B_0 = 14.1 T; 1.3 mm HCN; T_{eff} = 280 K (T_{set} = 248 K); v_{MAS} = 55 kHz; ns = 16; τ_{rd} = 1.2 s; τ_{CP-HC} = 1.5 ms; $v_{1H-CP,HC}$ = 93.5 kHz; $v_{13C-CP,HC}$ = 38.5 kHz; $t_{1,max}$ = 5.0 ms; $t_{1,inc}$ = 250.0 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{DCP-CN} = 7.1 ms; $v_{13C-DCP}$ = 35.0 kHz; $v_{15N-DCP}$ = 20.0 kHz; $v_{1H-DD,DCP}$ = 10 kHz; $t_{2,max}$ = 10.5 ms; $t_{2,inc}$ = 300.0 μs; $v_{1H-DD, t2evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-NH} = 0.8 ms; $v_{1H-CP,NH}$ = 93.5 kHz; $v_{15N-CP,NH}$ = 38.5 kHz; τ_{dwell} = 10.0 μs; τ_{acq} = 20.48 ms; $v_{15N-DD,acq}$ = 10 kHz; $v_{13C-DD,acq}$ = 10 kHz	17.4 hr
3D hCAcoNH	B_0 = 14.1 T; 1.3 mm HCN; T_{eff} = 280 K (T_{set} = 245 K); v_{MAS} = 55 kHz; ns = 32 (16 * 2 blocks); τ_{rd} = 1.4 s; τ_{CP-HC} = 1.0 ms; $v_{1H-CP,HC}$ = 93.5 kHz; $v_{13C-CP,HC}$ = 38.5 kHz; $t_{1,max}$ = 4.32 ms; $t_{1,inc}$ = 160.0 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{DREAM} = 5.0 ms; τ_{DCP-CN} = 6.0 ms; $v_{13C-DCP}$ = 35.0 kHz; $v_{15N-DCP}$ = 20.0 kHz; $v_{1H-DD,DCP}$ = 10 kHz; $t_{2,max}$ = 10.5 ms; $t_{2,inc}$ = 300.0 μs; $v_{1H-DD, t2evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-NH} = 1.1 ms; $v_{1H-CP,NH}$ = 93.5 kHz; $v_{15N-CP,NH}$ = 38.5 kHz; τ_{dwell} = 10.0 μs; τ_{acq} = 20.48 ms; $v_{15N-DD,acq}$ = 10 kHz; $v_{13C-DD,acq}$ = 10 kHz	2 * 27 hr = 54 hr
3D hcaCBcaNH	B_0 = 14.1 T; 1.3 mm HCN; T_{eff} = 280 K (T_{set} = 245 K); v_{MAS} = 55 kHz; ns = 24 (8 * 3 blocks); τ_{rd} = 1.2 s; τ_{CP-HC} = 1.1 ms; $v_{1H-CP,HC}$ = 93.5 kHz; $v_{13C-CP,HC}$ = 38.5 kHz; $t_{1,max}$ = 4.48 ms; $t_{1,inc}$ = 102.0 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{INEPT} = 6.5 ms * 2; τ_{ZF} = 3.0 ms; τ_{DCP-CN} = 7.8 ms; $v_{13C-DCP}$ = 30.0 kHz; $v_{15N-DCP}$ = 25.0 kHz; $t_{2,max}$ = 10.5 ms; $t_{2,inc}$ = 300.0 μs; $v_{1H-DD, t2evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-NH} = 1.2 ms; $v_{1H-CP,NH}$ = 93.5 kHz; $v_{15N-CP,NH}$ = 38.5 kHz; τ_{dwell} = 10.0 μs; τ_{acq} = 20.48 ms; $v_{15N-DD,acq}$ = 10 kHz; $v_{13C-DD,acq}$ = 10 kHz	3 * 19 hr = 57 hr
3D hcaCBcacoNH	B_0 = 14.1 T; 1.3 mm HCN; T_{eff} = 280 K (T_{set} = 245 K); v_{MAS} = 55 kHz; ns = 32 (16 * 2 blocks); τ_{rd} = 1.5 s; τ_{CP-HC} = 1.1 ms; $v_{1H-CP,HC}$ = 93.5 kHz; $v_{13C-CP,HC}$ = 38.5 kHz; $t_{1,max}$ = 4.48 ms; $t_{1,inc}$ = 102.0 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{INEPT} = 6.5 ms * 2; τ_{ZF} = 3.0 ms; τ_{DREAM} = 3.8 ms; τ_{DCP-CN} = 6.0 ms; $v_{13C-DCP}$ = 30.0 kHz; $v_{15N-DCP}$ = 25.0 kHz; $t_{2,max}$ = 10.5 ms; $t_{2,inc}$ = 300.0 μs; $v_{1H-DD, t2evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-NH} = 1.2 ms; $v_{1H-CP,NH}$ = 93.5 kHz; $v_{15N-CP,NH}$ = 38.5 kHz; τ_{dwell} = 10.0 μs; τ_{acq} = 20.48 ms; $v_{15N-DD,acq}$ = 10 kHz; $v_{13C-DD,acq}$ = 10 kHz	2 * 46.5 hr = 93 hr
2D ^{19}F-^{19}F exchange, (Intermediate T)	B_0 = 14.1 T; 1.9 mm HFX; T_{eff} = 285 K (T_{set} = 250 K); v_{MAS} = 38 kHz; τ_{rd} = 2.0 s; $t_{1,max}$ = 1.579 ms; $t_{1,inc}$ = 78.95 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{dwell} = 2.5 μs; τ_{acq} = 7.68 ms; $v_{1H-DD,acq}$ = 10 kHz; ns = [256, 256, 256, 256, 384, 512, 256], τ_{mix} = [0.1, 1.0, 5.0, 10.0, 20.0, 40.0, 75.0] ms	5.7 hr per ns = 256 block, total 48.5 hr
2D H-N H^N-F REDOR, S_0 and S pair	B_0 = 14.1 T; 1.9 mm HFX; T_{eff} = 282 K (T_{set} = 243 K); v_{MAS} = 38 kHz; τ_{rd} = 2.0 s; τ_{CP-HN} = 0.7 ms; $v_{1H-CP,HN}$ = 88.0 kHz; $v_{15N-CP,HN}$ = 50.0 kHz; $t_{1,max}$ = 28 ms; $t_{1,inc}$ = 200.0 μs; $v_{1H-DD, t1evol}$ = 10 kHz; τ_{solsup} = 200 ms; $v_{1H-solsup}$ = 15 kHz; τ_{CP-}	31.5 * 2 = 63 hr, 28.7 * 2 = 57.5 hr, 32.9 * 2 = 65.7 hr

$_{HN}$ = 0.7 ms; $v_{1H-CP,NH}$ = 88.0 kHz; $v_{15N-CP,HN}$ = 50.0 kHz; $v_{19F,pulse}$ = 83.3 kHz; $v_{15N-DD, REDOR}$ = 10 kHz; τ_{dwell} = 15.0 μs; τ_{acq} = 30.7 ms; $v_{15N-DD,acq}$ = 10 kHz, ns (S and S_0 each) = [56+64+64 = 184, 40+64+64 = 168, 128+64 = 192], $N\tau_{r,REDOR}$ = [64, 96, 144], $\tau_{mix, REDOR}$ = [1.68, 2.52, 3.78] ms	Total = 186 hr

Symbols: B_0 = magnetic field; NMR probe (rotor diameter, channels); T_{eff} = effective sample temperature; T_{set} = thermocouple-reported bearing gas temperature; v_{MAS} = MAS frequency; ns = number of scans (transients) per free induction decay (FID); τ_{rd} = recycle delay between scans; $t_{1,max}$ = maximum t_1 (indirect dimension 1) evolution time; $t_{1,inc}$ = increment for t_1 (indirect dimension 1) evolution time; $t_{2,max}$ = maximum t_2 (indirect dimension 2) evolution time; $t_{2,inc}$ = increment for t_2 (indirect dimension 2) evolution time; τ_{dwell} = dwell time during direct FID acquisition; τ_{acq} = maximum acquisition time during direct FID detection; τ_{CP-XY} = cross polarization (CP) contact time during CP *from* channel X *to* channel Y; $v_{nuc-CP,XY}$ = radiofrequency field strength for CP spin lock on *nuc* (*nuc* = ^1H, ^{13}C, ^{15}N, or ^{19}F) during CP from X to Y; $v_{1H-DD, t1evol}$ = ^1H dipolar decoupling field strength during t_1 evolution; $v_{1H-DD, t2evol}$ = ^1H dipolar decoupling field strength during t_2 evolution; $v_{nuc-DD, acq}$ = dipolar decoupling field strength during FID acquisition on channel *nuc*; τ_{DCP-XY} = CP contact time during heteronuclear (specific / double) CP from channel X to channel Y; $v_{nuc-DCP}$ = radiofrequency spin lock field strength on *nuc* during heteronuclear CP; $v_{1H-DD,DCP}$ = ^1H dipolar decoupling field strength during heteronuclear CP; τ_{DREAM} = spin lock contact time during homonuclear coherence transfer with the DREAM condition; τ_{INEPT} = delay for antiphase coherence evolution/reconversion for Ca-Cb INEPT; τ_{ZF} = post-INEPT Z-filter time for removal of unwanted coherences; τ_{solsup} = solvent suppression mixing time during MISSISIPPI period; $v_{1H-solsup}$ = radiofrequency field strength during solvent suppression; $N_{N\tau,REDOR}$ = number of rotor periods of REDOR recoupling; $v_{19F,pulse}$ = ^{19}F pulse field strength during REDOR recoupling; $v_{15N-DD, REDOR}$ = ^{15}N dipolar decoupling field strength during REDOR recoupling.
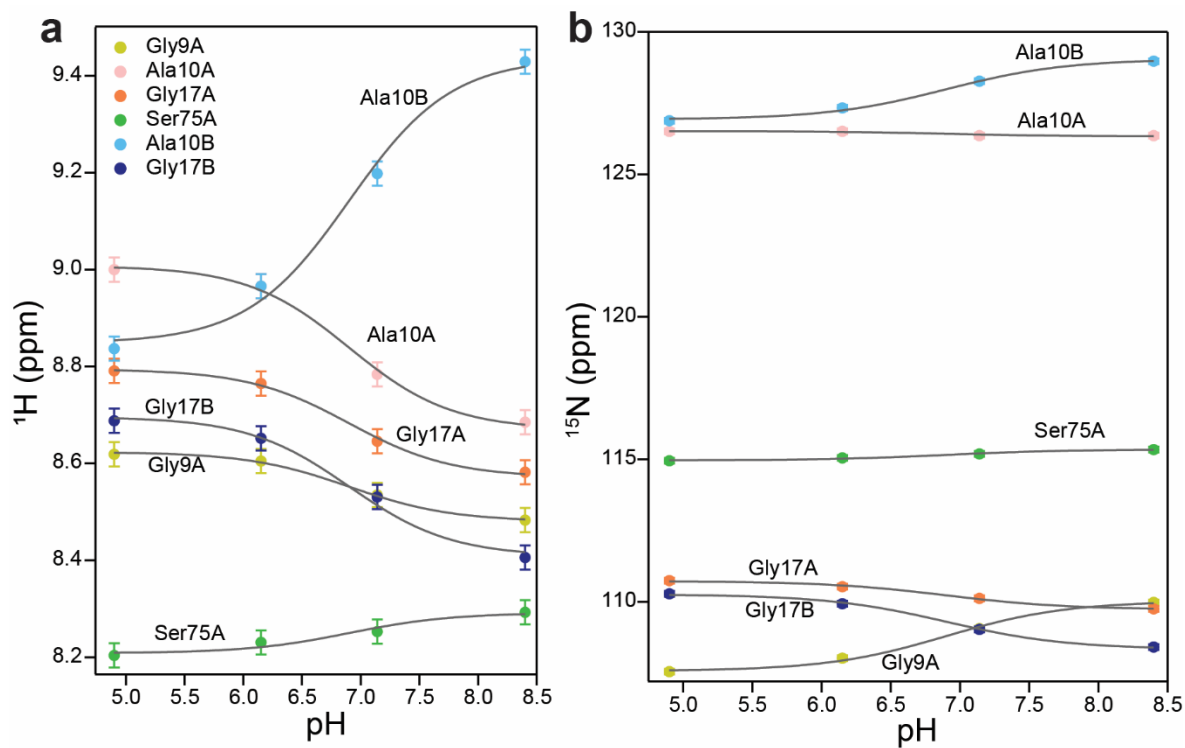
Supplementary Figure 1. pK$_a$ determination of F$_4$-TPP$^+$-bound S64V-EmrE by solution NMR in DMPC/DHPC bicelles. **(a)** ^{1}H chemical shifts for six well resolved residues as a function of pH. **(b)** ^{15}N chemical shifts for the same residues as a function of pH. Global fits (grey lines) were calculated, yielding a pK$_a$ value of 6.9 ± 0.1. The pH-dependent ^{1}H and ^{15}N chemical shifts are provided as a Source Data file.
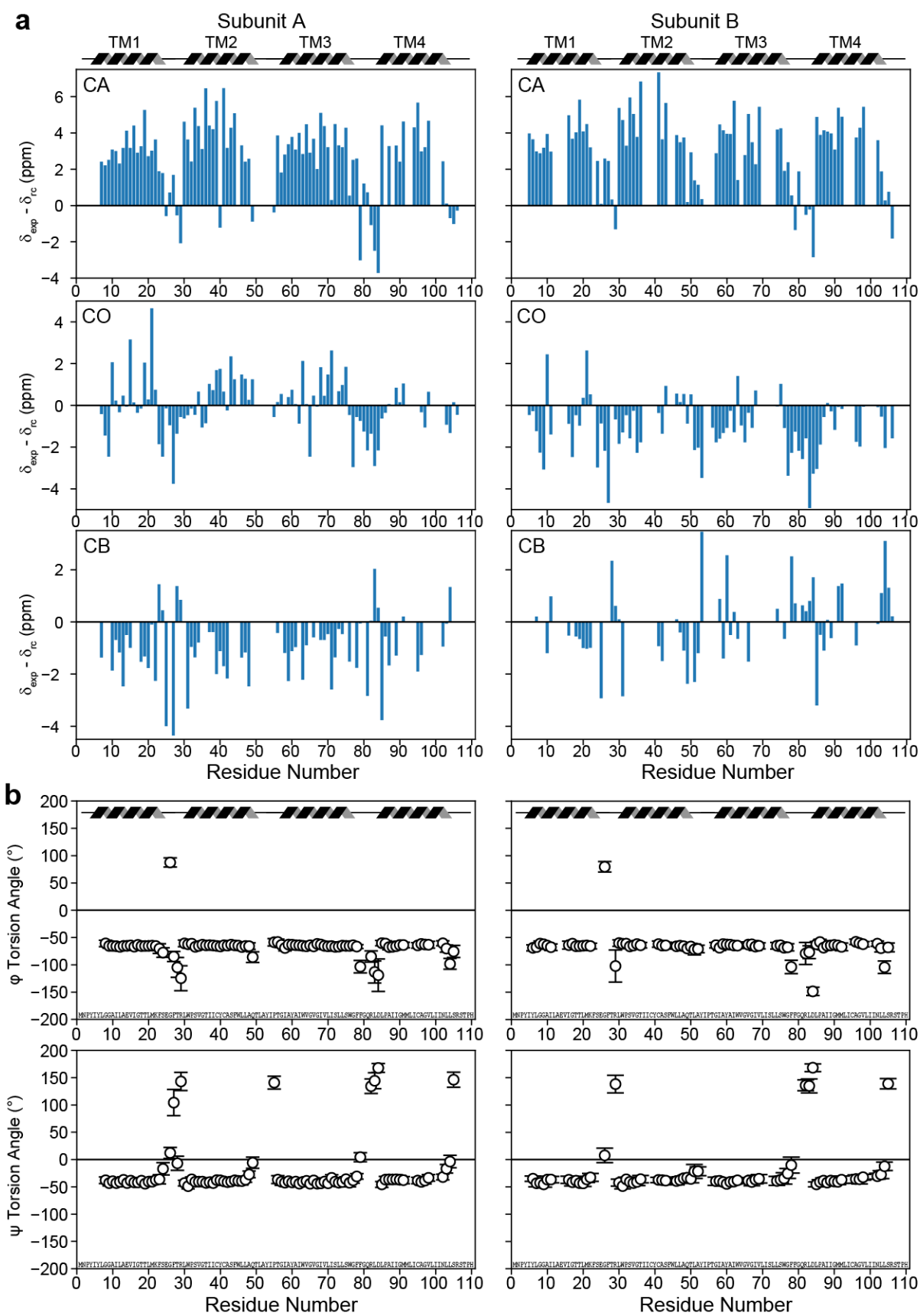
Supplementary Figure 2. Secondary structure of the pH 8.0 EmrE-TPP complex in DMPC bilayers derived from chemical shifts. (**a**) Secondary chemical shifts of Cα, Cβ, CO of EmrE at pH 8.0. (**b**) TALOS

predictions of Phi/Psi torsion angles. Errors for torsion angles are directly from TALOS output and represent 1σ error.
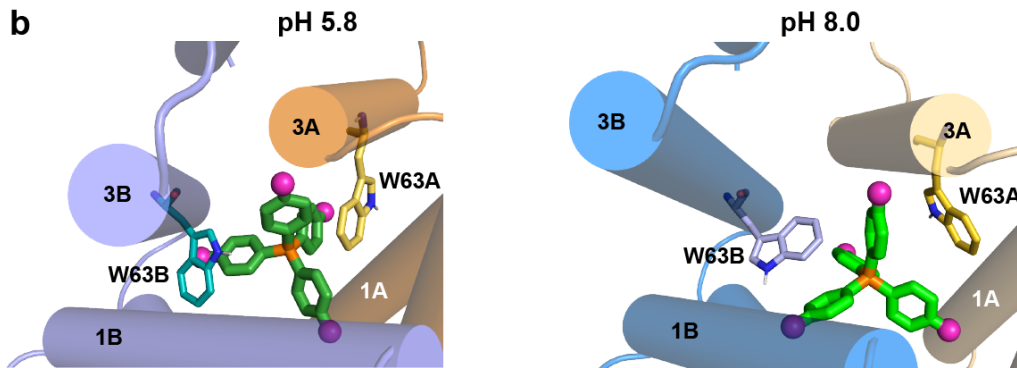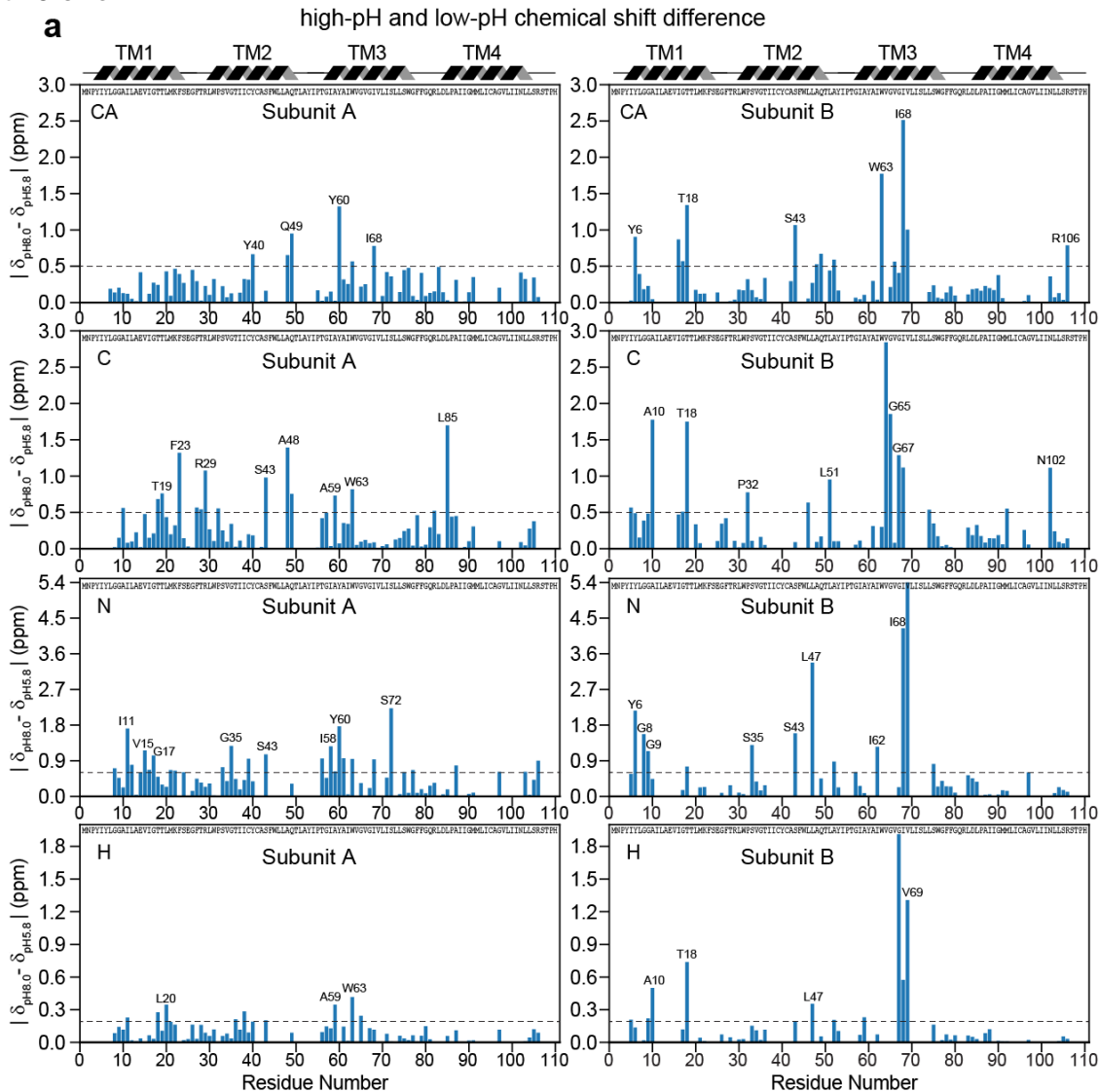
Supplementary Figure 3. Chemical shift differences of EmrE backbone atoms between pH 8.0 and pH 5.8. (**a**) CA, CO, N, and H chemical shift differences between pH 8.0 and pH 5.8. (**b**) The F_4-TPP$^+$ binding pocket in the low-pH and high-pH complex. The W63B indole ring is oriented differently with respect to the TM3B helix between the two structures, while the W63A indole has similar orientation relative to TM3A at both pH.
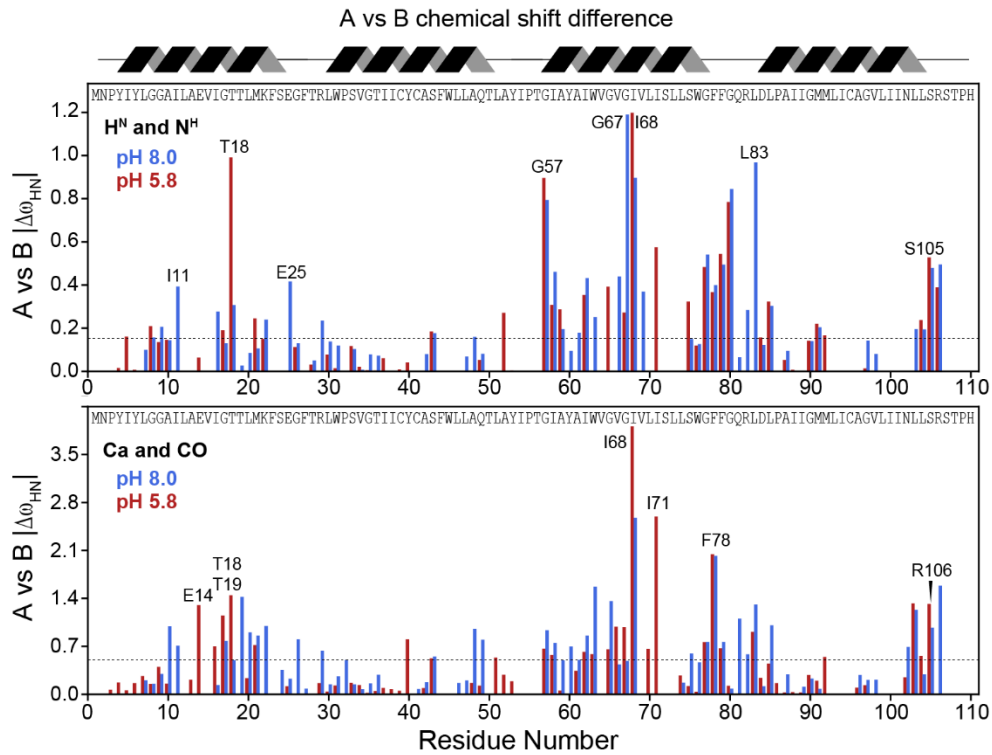
Supplementary Figure 4. Conformational asymmetry between subunits A and B of EmrE in DMPC bilayers. (**a**) Composite H^N and ^{15}N chemical shift differences between the two subunits. (**b**) Composite $C\alpha$ and CO chemical shift differences between the two subunits. Dashed lines indicate the linewidth-based estimate of the significance levels for chemical shift changes, which are 0.15 ppm for the composite amide and 0.5 ppm for ^{13}C.
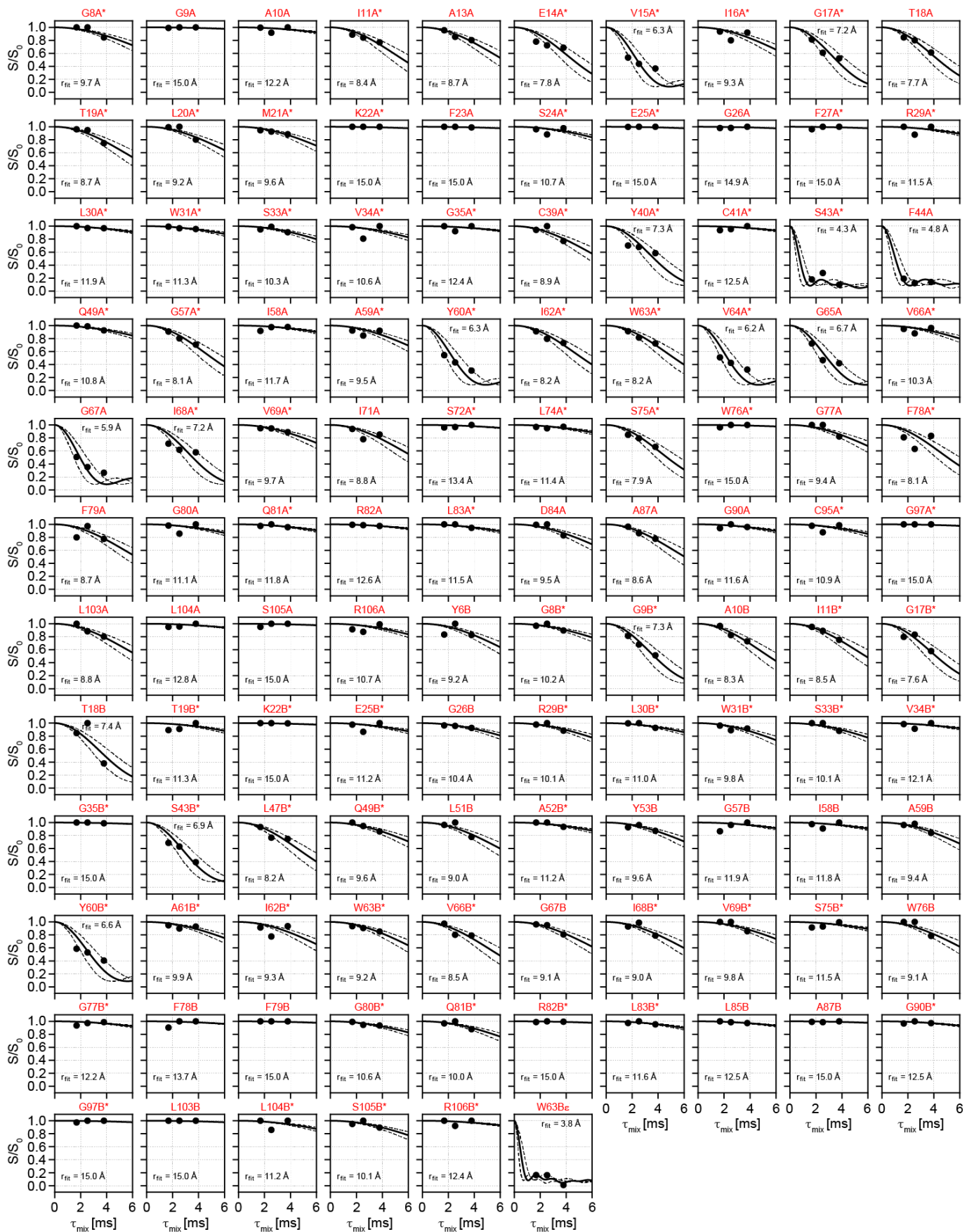
Supplementary Figure 5. All resolved ^{1}H-^{19}F REDOR dephasing curves for the EmrE-TPP complex at pH 8.0 in DMPC bilayers. For each resolved H^N peak, the simulated REDOR curve for the best-fit distance is overlaid with the experimental data. Asterisks indicate residues whose signals are partially overlapped in the 2D spectra.
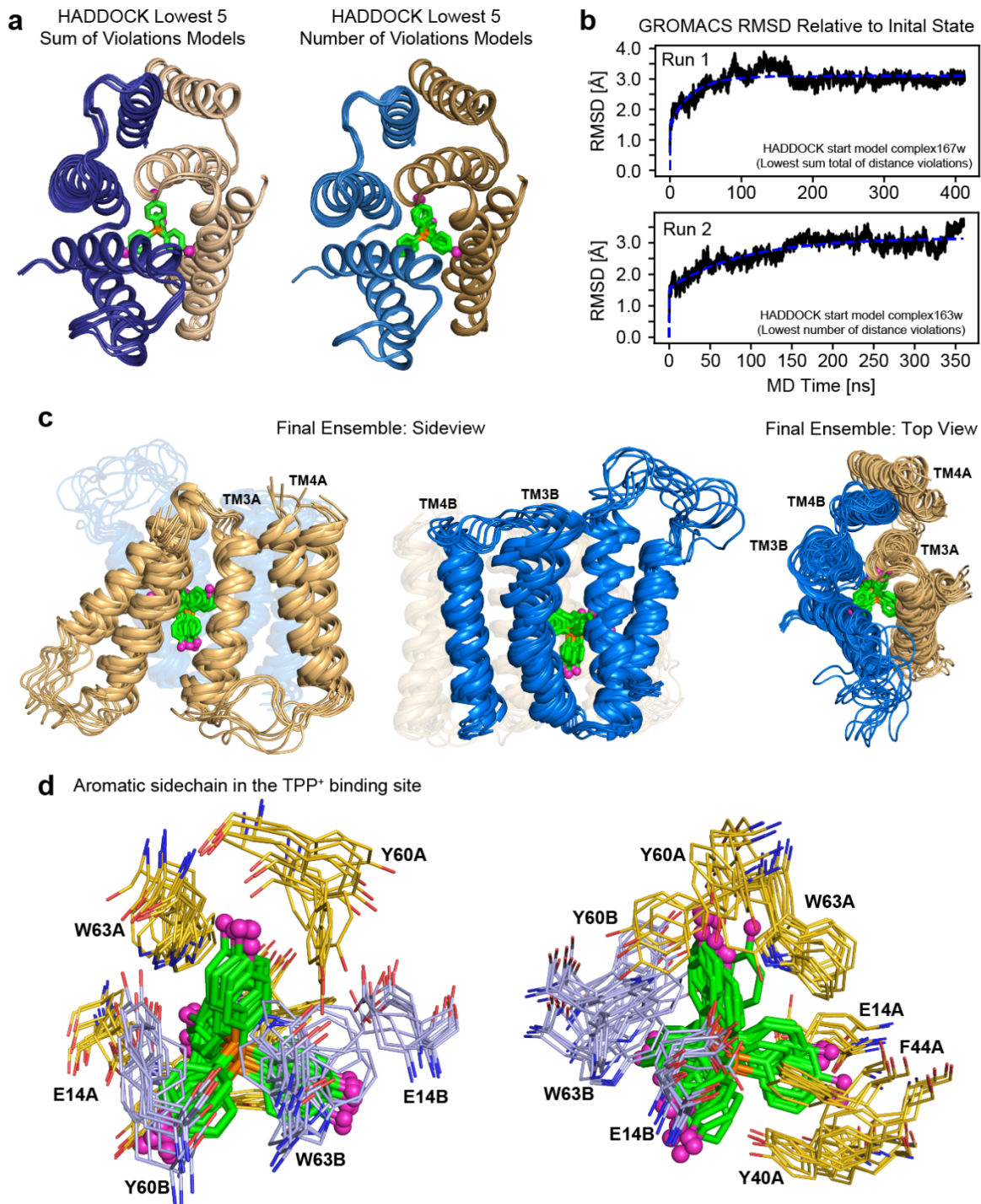
Supplementary Figure 6. Structure calculation of the pH 8.0 EmrE-TPP complex. (**a**) HADDOCK result of docking F_4-TPP$^+$ into the pH 5.8 structure with the distance constraints obtained at pH 8.0. The left view shows the 5 best models scored by the lowest sum of distance restraint violations (mean violation magnitude 2.2 ± 0.2 Å, mean number of violations = 8.4 ± 0.5). The right view shows the 5 best models scored by the lowest number of violations (mean violation magnitude 2.4 ± 0.3 Å, mean number of violations = 7.8 ± 0.4). The ligand center P atom has an RMSD of 0.2 ± 0.1 Å for the left ensemble and 0.5 ± 0.3 Å for the right ensemble. (**b**) Gromacs MD trajectory RMSD over 350-400 ns compared to the initial state. MD is equilibrated by 200 ns for both runs. (**c**) Side views and top view of the pH 8.0 EmrE-TPP ensemble. (**d**) Important binding-site sidechains (E14, Y40, F44, Y60 and W63) in the ensemble of 10 structures. Left view and right view differ by a ~180˚ rotation about the membrane normal.
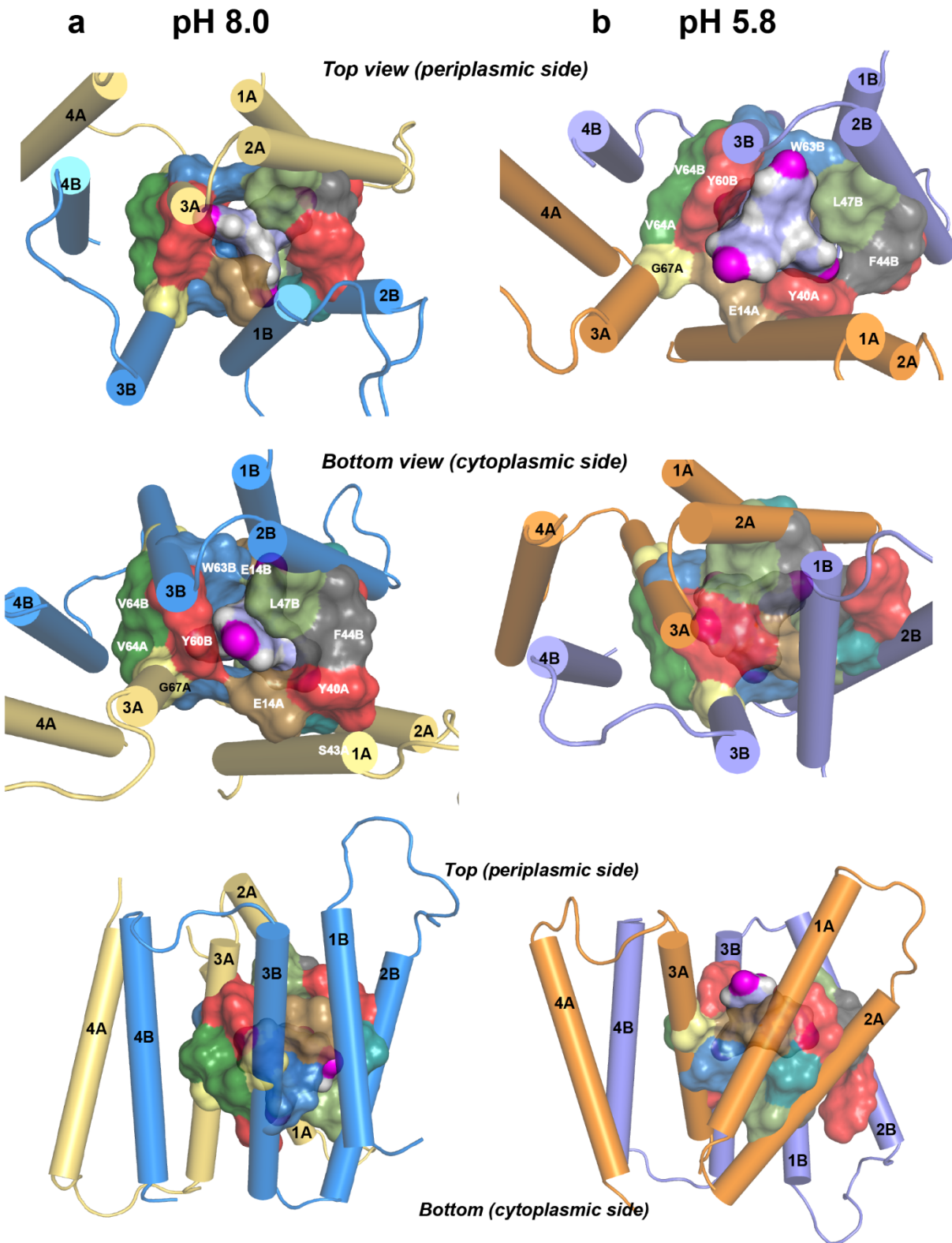
Supplementary Figure 7. Structural comparison of the EmrE-TPP complexes at high and low pH. (**a**) Surface views of the binding-site residues at high pH. (**b**) Surface views of the binding-site residues at low pH. Top and bottom sides are functionally defined (see main text). At high pH, F_4-TPP$^+$ is similarly and modestly exposed to the two sides of the helical bundle due to the relatively parallel orientations of the TM helices. In comparison, at low pH, F_4-TPP$^+$ is well exposed to the top of the helical bundle in a shallow binding pocket but is occluded at the bottom. We attribute the slower reorientation of the ligand at low pH to the bottom constriction in the low-pH complex.
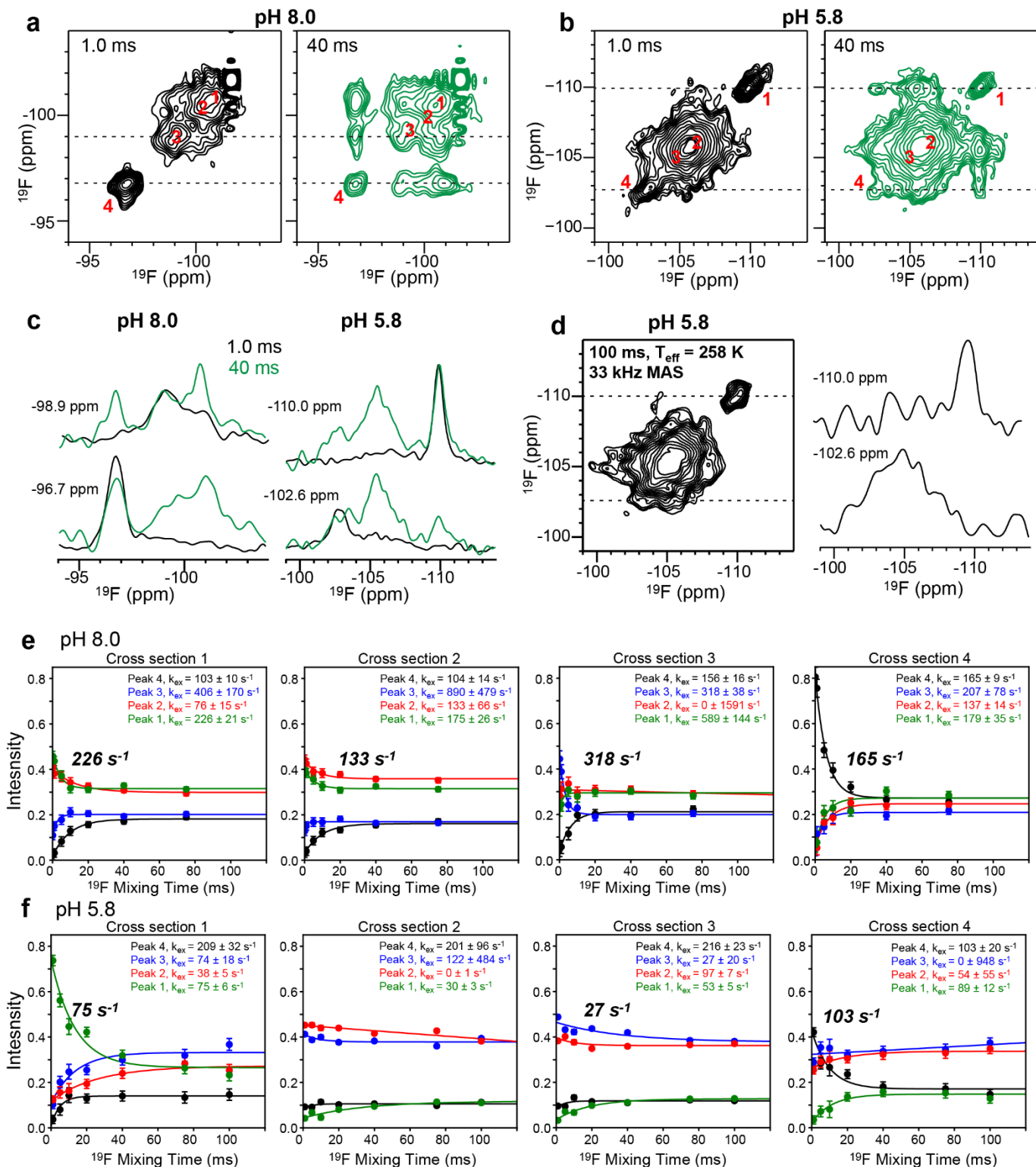
Supplementary Figure 8. Additional 2D ^{19}F-^{19}F exchange buildup curves of F_4-TPP$^+$ at high and low pH. (**a, b**) Representative 2D ^{19}F correlation spectra of F_4-TPP$^+$ with 1 ms and 40 ms mixing ms. (**a**) pH 8.0 data. (**b**) pH 5.8 data. These 2D spectra were measured at a sample temperature of ~285 K under 38 kHz MAS. (**c**) Two cross sections from the 2D spectra at both pH, showing the signal-to-noise ratios. (**d**) 100 ms 2D ^{19}F-^{19}F exchange spectrum of the pH 5.8 complex at an effective sample temperature of 258 K. No exchange cross peaks are observed, indicating that when immobilized, the inter-fluorine distances within each molecule are too long to be measured by spin diffusion on this timescale. Thus, the high-temperature cross peaks result from substrate reorientation. (**e**) Exchange buildup curves for the pH 8.0 sample. The diagonal peak's decay rates are indicated in each panel. (**f**) Exchange buildup curves for the pH 5.8

22

sample. The decay rates are smaller compared to the high-pH sample, indicating that the substrate reorients more slowly at acidic pH. Data are presented as mean values +/- 2σ. Error of intensity values was propagated from spectral signal to noise, while fitting parameter errors were estimated by Monte Carlo methods. The exchange peak intensities are provided as a Source Data file.
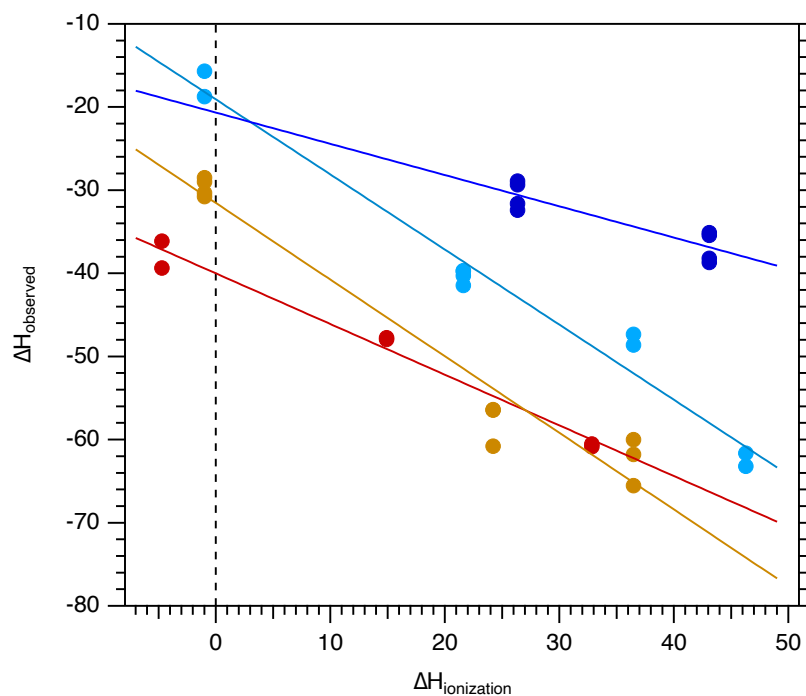
Supplementary Figure 9. Determination of pH-dependent binding enthalpy. Drug binding to EmrE releases protons, which are then bound by buffer, leading to a buffer-dependent contribution to the observed binding enthalpy. ITC experiments were performed for TPP$^+$ binding to WT EmrE in isotropic bicelles at 45°C at pH 5.5 (dark red), pH 6.5 (dark orange), pH 7.5 (cyan) and pH 8.5 (dark blue) with multiple buffers, as previously reported in (18). Extrapolation of the observed binding enthalpy to $\Delta H_{ionization}$ = 0 (dotted vertical line) allows determination of ΔH_{bind} due solely to the interaction of TPP$^+$ with the transporter, independent of the secondary protonation and deprotonation of the buffer. The thermodynamic data are provided as a Source Data file.
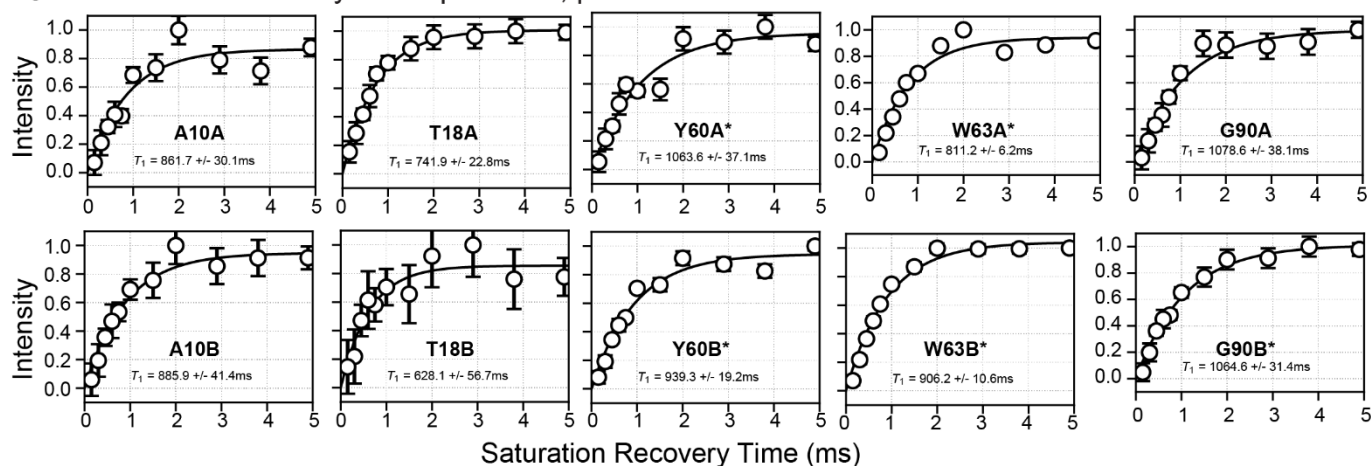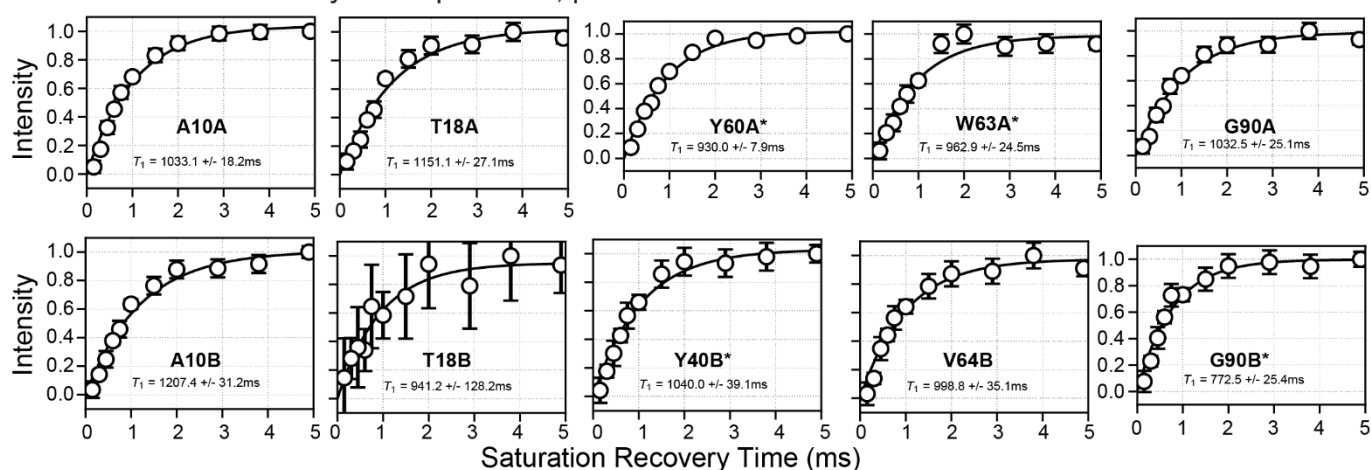
Supplementary Figure 10. Representative saturation recovery data of the EmrE-TPP complex. (**a**) High pH data. (**b**) Low pH data. Asterisks indicate residues that are partially overlapped in the 2D hNH spectra. Peak volumes were normalized by the largest value for each residue in the mixing time series and fit to a single-exponential plateau buildup in a two-parameter fit. Error bars were propagated from the signal-to-noise ratios (SNR) of each peak. Uncertainty of T_1 fitting parameter was determined by a Monte Carlo error estimation of 1000 trials and represents random uncertainty from the spectral SNRs and does not include systematic errors that result from peak shifts or spectral overlap. Data are presented as mean values +/- 2σ. Error of intensity values was propagated from spectral signal-to-noise ratios, while fitting parameter errors were estimated by Monte Carlo methods. Intensity values are provided as a Source Data file.
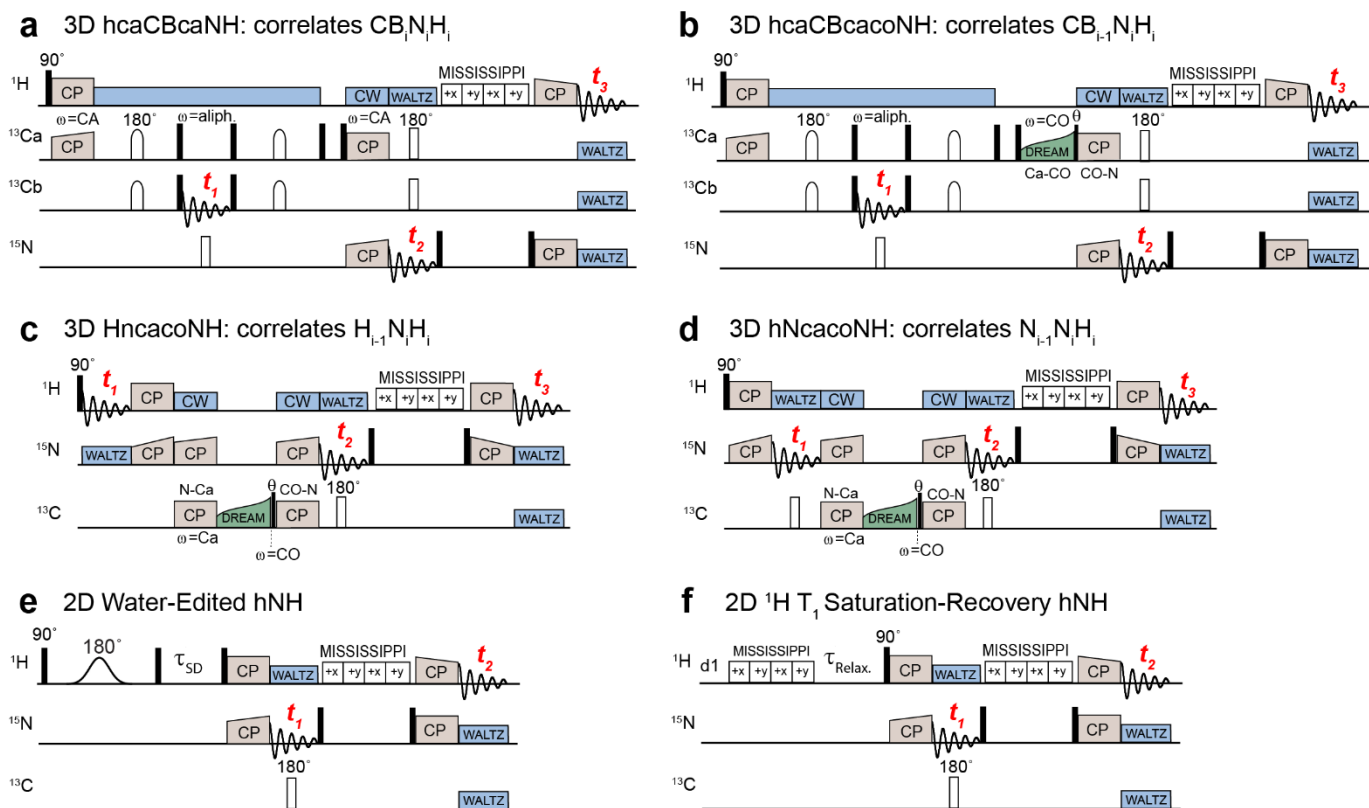
Supplementary Figure 11. Pulse diagrams for some of the 3D ^{1}H-detected correlation experiments for resonance assignment of EmrE at pH 8.0. (**a**) Intra-residue hcaCBcaNH experiment. (**b**) Inter-residue hcaCBcacoNH experiment. (**c**) Inter-residue HncacoNH experiment. (**d**) Inter-residue hNcacoNH experiment. (**e**) 2D hNH-resolved water-edited experiment. A selective 180° pulse is used to select the water magnetization, which is then allowed to diffuse to the protein amide protons during the variable delay τ_{SD}. (**f**) 2D hNH-resolved ^{1}H T_1 relaxation experiment. MISSISSIPPI solvent suppression is used prior to ^{1}H excitation to saturate the system, after which a variable delay τ_{relax} allows the system to undergo T_1 relaxation.

nature portfolio

Peer Review File

High-pH Structure of EmrE Reveals the Mechanism of Proton-Coupled Substrate Transport

Reviewer #1 (Remarks to the Author):

Structural knowledge of ErmE is imperative to elucidate the mode of action of SMR transporters and to develop inhibitors in the context of antimicrobial resistance. What is more, EmrE serves a model for understanding proton-coupled transport.

This study by Hong and co-workers presents high-quality insights into the high pH structure of EmrE, which is an important milestone to understand the transport mechanism of SMR transporters in a physiologically relevant lipid environment. This is an impressive tour-de-force and merits publication in Nature Communication.

Especially the 19F-based data open new vistas for solid-state NMR and add a distance dimension that was previously unheard of in our community. We will see 19F-based bio-ssNMR applications more often in the future, and the here presented data (along with the previous study by Hong/Henzler-Wildman) are among the very best that I have seen.

I have some comments that I would like to see addressed:

1. Figure 4: There seems some confusion regarding the perspective and/or colouring of subunits A and B. Why is the blue subunit in front in c) and in the back in d)? It seems that a 270° rotation was applied to a), yet a 90° rotation to b). Please align the representation.

2. I do not have practical experience in 19F-19F spin diffusion, and this applies to the vast majority of the biological ssNMR community. It would hence instil trust in the validity of the 19F-19F exchange data if the authors showed the data obtained at 265 K, where no exchange was observed. This was done in the previous low pH EmrE paper, but the longitudinal exchange/transfer time was longer in this manuscript. Please also provide cross-sections at the relevant peak positions. This could be done in the Supporting Information.

3. Please provide a detailed violation analysis for the structure determinations in the Supporting Information. Please also provide more information on how the distance restraints were implemented in HADDOCK, what were upper and lower boundaries, what were the force constants. It would also be useful to provide this kind of information for the MD simulations.

4. It is intriguing to see that the dynamics of the substrate increases at high pH, although the binding affinity for TPP+ is higher at high pH. Given that drug binding specificity usually rather relates to enthalpic contributions and given that the substrate specificity of EmrE is apparently low, I would expect that entropic contributions play a considerable role in substrate binding/binding affinity. Could the authors comment on this? Are ITC data known for EmrE?

Markus Weingarth

Reviewer #2 (Remarks to the Author):

Shcherbakov and co-authors determined a dimeric SMR transporter EmrE in a complex with F4- TPP+ at high pH (two E14 sites - deprotonated). They also determined recently determined another state of EmrE, with F4- TPP at low pH (one of the E14 should be protonated). By comparison with two structures, they discuss the structural changes and proton-substrate coupling mechanism.

General comments

EmrE has been well-characterized as a model system of the SMR transport family. This paper found that EmrE has a more elongated and spacious binding pocket in the high-pH, and the substrate is more dynamic than in the low-pH. They also found that the dissociation of protons induces a small opening on either side of the membrane. They propose that these protonation/deprotonation-dependent conformational differences explain how proton plays a role in conformation changes coupling with substrate release in EmrE.

High-resolution structural data of EmrE are already available recently by the same group (Scherbakov et al. Nat Commun. – Ref39), but this time "the missing link structure – a substrate-bound form in the high-pH condition" was determined. Therefore, the value of this paper should be more properly assessed by pairing together with a low-pH structure. Considering its importance as a potential future drug target, and a deeper insight into the transport mechanism of EmrE is of broad interest to the scientific community. The author's group is an expert on EmrE using solution or solid-state NMR technique; thus, their accumulated functional and structural analyses are sound.

Although there is still an argument of precise pKa value of E14 residues (e.g., Ref 21, 30 32), the results are expected to be largely in line with the results that have been obtained so far, such as asymmetric protonation, two independent pKa of E14…

In my view, the discussion and conclusions derived from the experiment are reasonable to present. I enjoyed reading and but I have some questions and suggestions that could improve the manuscript.

Page 5.

The authors discuss the hydration level in the binding pocket. This is shown by water-edited 2D hNH spectra. The authors are also experts in MD simulation, so it seems like the environment is ready. How about simulating how water molecules affect the conformation of the TM helices and the volume change of the binding pocket? The outcome might support further "The fact that these significant structural changes result exclusively from a pH change is remarkable. page 5"

Page 6. In the final paragraph, the author mentions

> This suggests that after a toxic substrate has bound to the protein from the cytoplasm and EmrE adopts a conformation similar to Fig. S7a, there is sufficient space for protons to enter and bind E14 from the periplasmic side. This could then trigger a conformational change to a state similar to our low-pH structure, with the transporter open to the periplasmic (low pH) side of the membrane. In this state, the toxic substrate is bound more peripherally to the central transport pore and is primed for release in accord with the enhanced off rate in the drug- and proton-bound state.

This is an exciting mechanistic proposal. Is it possible to add any supportive evidence from previous results or simulations? e.g., if you run the MD simulation using a high-pH structure as a starting model and then protonate one or both of E14 sites, does the protonation induces conformational changes and makes the substate possible to be released to the periplasmic side?

Minor comments

Results – first paragraph

the pKa of S64V-EmrE – do you mean the pI value? Alternatively, pKa value of specific residue, please specify.

Reviewer #3 (Remarks to the Author):

In this paper, the authors present the high pH structure of the small multidrug transporter EmrE.

EmrE has been a paradigm for secondory drug-protein antiporters but it raised many questions with respect

to structure, oligomerisation, transport cycle and stoichiometry. There is now converging evidence that EmrE

is an asymmtric, antiparallel dimer with two Glue (E14) coordinating substrate and proton binding.

In a previous paper, the authors have presented the low pH structure of EmrE in lipid bilayers based on solid-state NMR. Here, they extend this work towards the high pH structure. Both structure are essential to understand the transport cycle. The experimental strategy is rather elegant and is based on two approaches: (i) High spectral resolution in achieved by the S64V mutant, which is fully functional but shifts the protein dynamics into a more favourable region. (ii) Highly quantitative distance restraints for the 3D structure determination have been obtained by using a 19F-labelled ligand through which many 19F-13C distances between ligand and protein have been measured.

The obtained structure agrees with available biochemical data and shows interesting differences to the low pH state. In particular, the binding pocket appears elongated, spacious and hydrated. Overall, the protein structure is more symmetric. The binding pocket apperas to be open partially open to either side. It is amazing to see that a pH change causes such large differences.

Overall, I consider the experimental approach as elegant and highly appropriate, the data are of high quality and I fully agreed with the conclusions drawn from the results.

Reviewer #1 (Remarks to the Author):
Structural knowledge of ErmE is imperative to elucidate the mode of action of SMR transporters and to develop inhibitors in the context of antimicrobial resistance. What is more, EmrE serves a model for understanding proton-coupled transport.

This study by Hong and co-workers presents high-quality insights into the high pH structure of EmrE, which is an important milestone to understand the transport mechanism of SMR transporters in a physiologically relevant lipid environment. This is an impressive tour-de-force and merits publication in Nature Communication.

Especially the 19F-based data open new vistas for solid-state NMR and add a distance dimension that was previously unheard of in our community. We will see 19F-based bio-ssNMR applications more often in the future, and the here presented data (along with the previous study by Hong/Henzler-Wildman) are among the very best that I have seen.

I have some comments that I would like to see addressed:

1. Figure 4: There seems some confusion regarding the perspective and/or colouring of subunits A and B. Why is the blue subunit in front in c) and in the back in d)? It seems that a 270° rotation was applied to a), yet a 90° rotation to b). Please align the representation.

We thank the reviewer for this question. The previous bottom view for the high-pH complex and A/B designation were correct. The confusion about the A/B designation and orientations of the dimer arises from the fact that alternating access switches the definition of subunit A and subunit B between the two pH states! In other words, the same molecule switches their assignment. To minimize this confusion, we believe it is to maintain the functional definition that the "top" side of the TM helical bundle is the acidic periplasmic side and the "bottom" side is the neutral cytoplasmic side. Therefore, we have now chosen to show the top view of both the high pH and low pH structures. For the same reason, in the side views, the subunit that is closer to the reader switches from B (blue) at high pH to A (orange) at low pH, even though it is the same molecule. So the previous views were all correct. And our slightly updated figure, showing a different view at the top left, hopefully clarifies this complexity further.
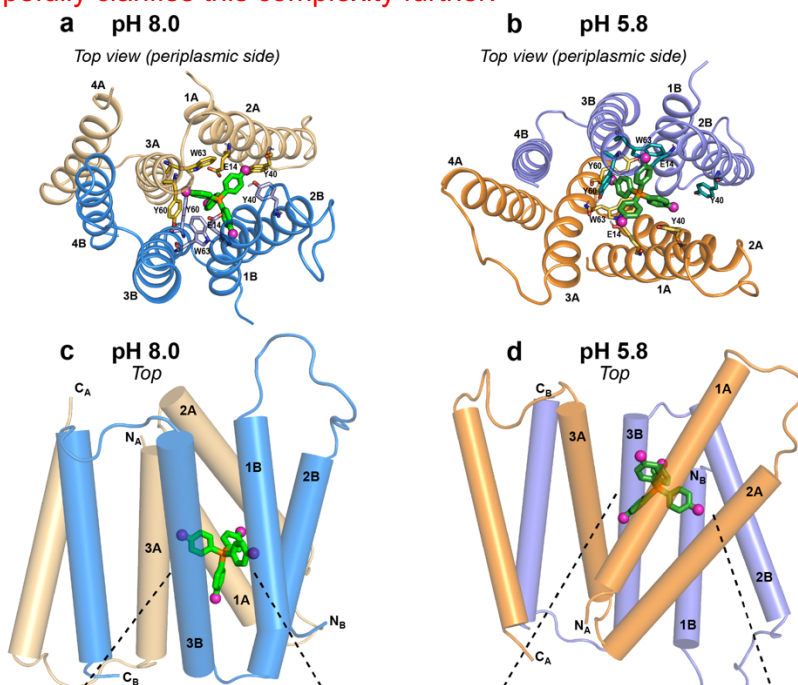
We have now revised the Figure 4 caption:

"(**a**) Top view seen from the periplasmic side of the high-pH structure, showing the crucial binding-site residues E14, Y40, Y60 and W63. (**b**) Top view of the low-pH structure for comparison. Note that conformer A (beige) in the high-pH complex has changed to conformer B (purple) in the low-pH complex. This conformational change switches the designation of the two subunits between high and low pH."

2. I do not have practical experience in 19F-19F spin diffusion, and this applies to the vast majority of the biological ssNMR community. It would hence instill trust in the validity of the 19F-19F exchange data if the authors showed the data obtained at 265 K, where no exchange was observed. This was done in the previous low pH EmrE paper, but the longitudinal exchange/transfer time was longer in this manuscript. Please also provide cross-sections at the relevant peak positions. This could be done in the Supporting Information.

In Figure S8 we have now added 1D ^{19}F cross sections and a control 2D spectrum measured at a low sample temperature of 258 K with 100 ms mixing. There is no significant exchange at this low temperature and long mixing time. This means that the ^{19}F-^{19}F distances within each ligand are too long to cause spin diffusion on the timescale of these 2D exchange experiments. Because the ligand geometry is fixed and is independent of pH, one control spectrum at low temperature is sufficient to prove that all cross peaks seen at high temperature arise form dynamic interconversion at both pH. We have now clarified this point in the caption:

"(**c**) Two cross sections from the 2D spectra at both pH, showing the signal-to-noise ratios. (**d**) 100 ms 2D ^{19}F-^{19}F exchange spectrum of the pH 5.8 complex at a sample temperature of 258 K. No exchange cross peaks are observed, indicating that when immobilized, the inter-fluorine distances within each molecule are too long to be measured by spin diffusion on this timescale. Thus, the high-temperature cross peaks result from substrate reorientation. "

3. Please provide a detailed violation analysis for the structure determinations in the Supporting Information. Please also provide more information on how the distance restraints were implemented in HADDOCK, what were upper and lower boundaries, what were the force constants. It would also be useful to provide this kind of information for the MD simulations.

We thank the reviewer for careful reading of the manuscript and for pointing out these details were missing from the paper. We have now added Tables S2 and S4 to give the list of distance constraints used during docking and MD, respectively. We have also added Tables S3 and S5 to give the detailed violation statistics and analysis, which were crucial in the structure calculation pipeline. We have also added details to the methods section about the implementation of distance constraints in the docking and MD simulations. With these details other people should be able to reproduce our structure calculation if needed.

4. It is intriguing to see that the dynamics of the substrate increases at high pH, although the binding affinity for TPP+ is higher at high pH. Given that drug binding specificity usually rather relates to enthalpic contributions and given that the substrate specificity of EmrE is apparently low, I would expect that entropic contributions play a considerable role in substrate binding/binding affinity. Could the authors comment on this? Are ITC data known for EmrE?

Markus Weingarth

We thank the reviewer for this interesting question. We do have a significant amount of ITC data regarding binding of multiple different substrates to EmrE under different conditions. ΔG, ΔH and

ΔS all vary with the individual substrate as expected. The data that we have acquired that most directly addresses the reviewer's question was acquired for TPP+ binding to WT EmrE as a function of pH. This data was previously published (Thomas *et al,* 2018 *J Biol Chem*, Ref 18). Analysis of the relative contribution of enthalpy and entropy to overall binding affinity is complicated by the fact that drug binding is coupled to proton release. Thus,

$$\Delta H_{observed} = \Delta H_{bind} + \Delta H_{ionization}$$

Where the observed enthalpy by ITC includes both a contribution from the binding event and the enthalpy of buffer ionization as protons are absorbed/released by the buffer. We previously performed TPP+ binding experiments in multiple buffers at 4 different pH values to determine the number of protons released upon drug binding. Since different buffers have different enthalpies of ionization, this can be used to determine the number of protons released upon drug binding to EmrE. However, we can reanalyze this data to instead extract the ΔH_{bind} values by extrapolating to $\Delta H_{ionization}$ =0. This is what we now report in **Table 3.**

We first note that at high pH where there is minimal proton release upon drug binding, ΔH and ΔS are nearly identical in all buffers confirming the accuracy of this method. Second, the affinity increases with increasing pH due to increasingly favorable ΔS that overcomes a less favorable ΔH. This increasingly favorable entropic contribution to binding at high pH is exactly consistent with the observed increase in ligand dynamics and hydration in the high pH structure reported here. We thank the reviewer for prompting us to re-examine the ITC data and add this important confirmation to the manuscript.

Reviewer #2 (Remarks to the Author):

Shcherbakov and co-authors determined a dimeric SMR transporter EmrE in a complex with F4-TPP+ at high pH (two E14 sites - deprotonated). They also determined recently determined another state of EmrE, with F4- TPP at low pH (one of the E14 should be protonated). By comparison with two structures, they discuss the structural changes and proton-substrate coupling mechanism.

General comments

EmrE has been well-characterized as a model system of the SMR transport family. This paper found that EmrE has a more elongated and spacious binding pocket in the high-pH, and the substrate is more dynamic than in the low-pH. They also found that the dissociation of protons induces a small opening on either side of the membrane. They propose that these protonation/deprotonation-dependent conformational differences explain how proton plays a role in conformation changes coupling with substrate release in EmrE.

High-resolution structural data of EmrE are already available recently by the same group (Scherbakov et al. Nat Commun. – Ref39), but this time "the missing link structure – a substrate-bound form in the high-pH condition" was determined. Therefore, the value of this paper should be more properly assessed by pairing together with a low-pH structure. Considering its importance as a potential future drug target, and a deeper insight into the transport mechanism of EmrE is of broad interest to the scientific community. The author's group is an expert on EmrE using solution or solid-state NMR technique; thus, their accumulated functional and structural analyses are sound.

Although there is still an argument of precise pKa value of E14 residues (e.g., Ref 21, 30 32), the results are expected to be largely in line with the results that have been obtained so far, such as asymmetric protonation, two independent pKa of E14…

In my view, the discussion and conclusions derived from the experiment are reasonable to present. I enjoyed reading and but I have some questions and suggestions that could improve the manuscript.

Page 5.
The authors discuss the hydration level in the binding pocket. This is shown by water-edited 2D hNH spectra. The authors are also experts in MD simulation, so it seems like the environment is ready. How about simulating how water molecules affect the conformation of the TM helices and the volume change of the binding pocket? The outcome might support further "The fact that these significant structural changes result exclusively from a pH change is remarkable. page 5"

We thank the reviewer for these important suggestions regarding the structures and mechanistic conclusions. We have added panels to show explicit water molecules in the MD simulations in Figure 6. This suggestion prompted us to find and show that the MD simulations are consistent with the hydration and 2D ^{19}F-^{19}F exchange NMR data. The high-pH complex has a much more hydrated binding site, with more than twice as many water molecules in the binding cavity relative to the low pH state. Further, the MD simulations support the mechanistic hypothesis, as they clearly show the high-pH state has water molecules from both sides of the membrane to access the ligand, whereas the low-pH state has water molecules only accessing from one side of the membrane. This result significantly strengthened our conclusion of the transport mechanism by EmrE.

We have added the following in the results section:

" These spectral observations are borne out by the MD equilibrated structural ensemble, as the structure calculation explicitly solvated the protein-ligand complex in lipid bilayers. **Fig. 6e** shows membrane-embedded water molecules whose oxygen atoms lie within 15 Å of any of the ligand atoms. Strikingly, the ligand-binding site is much more hydrated in the high-pH complex: a total of 69 water molecules are found in the ligand-binding pocket, and these water molecules approach the ligand from both sides of the lipid bilayer. In comparison, only 23 water molecules are found in the ligand-binding pocket in the low-pH complex; moreover they approach the ligand only from one side, the putative periplasmic side, of the lipid bilayer. "

Our discussion section now includes the following sentences:

" It is interesting to note that the high-pH structure of the EmrE-TPP complex is not fully closed on either side of the membrane. This is borne out by the water-accessibility data (**Fig. 6a-d**) and the MD simulations (**Fig. 6e**), which show larger water accessibility of the ligand-binding cavity at high pH than at low pH. Moreover, water molecules approach the ligand from both sides of the membrane at high pH, but only accesses the ligand from one side of the membrane at low pH. "

Page 6. In the final paragraph, the author mentions

> This suggests that after a toxic substrate has bound to the protein from the cytoplasm and EmrE adopts a conformation similar to Fig. S7a, there is sufficient space for protons to enter and bind E14 from the periplasmic side. This could then trigger a conformational change to a state similar to our low-pH structure, with the transporter open to the periplasmic (low pH) side of the membrane. In this state, the toxic substrate is bound more peripherally to the central transport

pore and is primed for release in accord with the enhanced off rate in the drug- and proton-bound state.

This is an exciting mechanistic proposal. Is it possible to add any supportive evidence from previous results or simulations? e.g., if you run the MD simulation using a high-pH structure as a starting model and then protonate one or both of E14 sites, does the protonation induces conformational changes and makes the substate possible to be released to the periplasmic side?

This is an interesting idea to pursue in the future but is outside the scope of the current study. The goal of the current study is to provide direct experimental data of the high-pH structure to complete the description of the alternating-access motion of EmrE as bound to a substrate. For this experimental structure determination, we do not need simulations as supporting evidence. But in the future, it will be very interesting to follow up with simulations to investigate the molecular pathways and potential intermediate states that connect the two structures during alternating access.

Minor comments
Results – first paragraph
the pKa of S64V-EmrE – do you mean the pI value? Alternatively, pKa value of specific residue, please specify.

This refers to the pK_a value of E14. We have now clarified this in the results section.

Reviewer #3 (Remarks to the Author):

In this paper, the authors present the high pH structure of the small multidrug transporter EmrE. EmrE has been a paradigm for secondory drug-protein antiporters but it raised many questions with respect to structure, oligomerisation, transport cycle and stoichiometry. There is now converging evidence that EmrE is an asymmtric, antiparallel dimer with two Glue (E14) coordinating substrate and proton binding.

In a previous paper, the authors have presented the low pH structure of EmrE in lipid bilayers based on solid-state NMR. Here, they extend this work towards the high pH structure. Both structure are essential to understand the transport cycle. The experimental strategy is rather elegant and is based on two approaches: (i) High spectral resolution in achieved by the S64V mutant, which is fully functional but shifts the protein dynamics into a more favourable region. (ii) Highly quantitative distance restraints for the 3D structure determination have been obtained by using a 19F-labelled ligand through which many 19F-13C distances between ligand and protein have been measured.

The obtained structure agrees with available biochemical data and shows interesting differences to the low pH state. In particular, the binding pocket appears elongated, spacious and hydrated. Overall, the protein structure is more symmetric. The binding pocket appears to be open partially open to either side. It is amazing to see that a pH change causes such large differences.

Overall, I consider the experimental approach as elegant and highly appropriate, the data are of high quality and I fully agreed with the conclusions drawn from the results.

We are glad the reviewer appreciates the quality and novelty of this study and agrees with our conclusions. No change is needed.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

All my comments have been expertly addressed. Especially the analysis of the binding energetics has strengthened this very interesting work even further. I support publication of the manuscript in its current form and congratulate the authors.

Reviewer #2 (Remarks to the Author):

Thank you for adding the MD simulation of water molecules in the substrate-binding site. It is a fascinating result. I am also glad to see that these results support the experimental evidence.

Although I have proposed another MD simulation, these simulations may require plenty of adjustments and detailed conditional studies in some cases. As the authors have stated, I accept this may be beyond the scope of this work, and do not ask further.

Overall, I consider the experimental data and the conclusions drawn from the results attractive and reasonable, significantly contributing to understanding this efflux system.

I have no further comments.