PNAS
www.pnas.org

Supplementary Information for

Electrostatic Lock in the Transport Cycle of the Multi-Drug Resistance Transporter EmrE

Josh V. Vermaas, Susan B. Rempe, and Emad Tajkhorshid

Susan B. Rempe
Email:slrempe@sandia.gov
Emad Tajkhorshid
E-mail: emad@life.illinois.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Supporting Information Text

Additional data table S1 (faRM.pdb.txt)
Final faRM model in pdb format.

Additional data table S2 (top_tpp.txt)
TPP^+ topology file.

Additional data table S3 (par_tpp.txt)
TPP^+ parameter file.

Movie S1. Supplementary Animation 1. Direct structural comparison of the initial structure based on a naïve refinement of the original crystal structure (lighter colors) with the fully atomic Refined Model (faRM) obtained after interactive MDFF simulations (darker colors), rotated about the membrane normal. The protein structure is represented as a cartoon, where α-helical secondary structure elements as determined by `Stride` are clearly demarcated from loop regions. The extent of the hydrophobic acyl chains of the modeled DMPC bilayer is shown as a transparent surface around the EmrE homodimer. The monomers are identified by their color (A is blue, B is red), using the chain identifier from the X-ray structure of the EmrE TPP-bound state (PDB: 3B5D) and the cryo-EM based model (PDB: 2I68). A green sphere is drawn on the N-terminus of each monomer to help identify the loops. Proline residues 3, 32, 55, and 86 are also drawn to highlight their important role in terminating helices in the faRM. This animation is directly related to Fig. S1.

Movie S2. Supplementary Animation 2. The cryo-EM derived electron density (black wireframe) overlaid on the faRM, with the addition of a stick model for the side chain heavy atoms and rotated around the membrane normal. Carbon atoms are gray, nitrogen atoms are blue, oxygen atoms are red, and sulfur atoms are yellow. The extent of the membrane hydrophobic core surrounding the protein is shown as a transparent surface. Note that the reported cryo-EM densities show additional density outside the dimer. This animation is directly related to Fig. S2.

Movie S3. Supplementary Animation 3. Animation of a single 500 ns TPP^+-bound trajectory, where lipids (orange) split apart the EmrE dimer (red and blue cartoon) and interact directly with TPP^+ (gray). Lipid hydrogen atoms, water and ions have been omitted from this representation, and the motion of the particles has been smoothed in the visualization.

Movie S4. Supplementary Animation 4. Animation of a single 500 ns singly protonated (B^+) trajectory, highlighting the motion of Helix 3 (yellow) within monomer A (blue cartoon) that opens a continuous water channel. Monomer B is shown in red, and water within the membrane interface is shown as a blue surface.

Movie S5. Supplementary Animations 5A-E. Visualizations of the mean water occupancy across all five replicates within the lumen of EmrE under different loading conditions. Regions where water is frequently present are represented by isosurfaces of different colors (gray for the apo state (A), blue for the TPP^+ bound state (B), orange for A^+ and red for B^+ states (C and D), and green for the A^+B^+ state (E)), along with a cartoon representation of EmrE for context. These animations are directly related to Fig. 4.

Movie S6. Supplementary Animation 6. Demonstration of the different motions of the C-terminus, through a comparison of the motion of R106 from each monomer relative to the lumen, represented here by $E14^B$. The position of R106 across an apo trajectory is shown through a rainbow of residues superimposed on the same dimeric structure (transparent cartoon). This animation is directly related to Fig. S14.

Movie S7. Supplementary Animation 7. The lipid occupancy across all simulation conditions for semitransparent slices along the membrane normal, with higher lipid occupancies colored in red, and lower occupancies in blue, with the protein (drawn as a cartoon) provided for context. The high occupancy in certain regions highlights the specificity of lipid intercalation to the "open" side of the transporter. The snapshots shown all have the "open" side of the transporter (the side accessible to solution) oriented toward the viewer. This animation is directly related to Fig. 8.

Movie S8. Supplementary Animation 8. An example of a driven transition, with the static lighter dimer representing the initial state, and the darker EmrE dimer showing a smoothed animation of the transition between the inward- and outward- facing states.

Construction of the fully atomic Refined Model (faRM)

Construction of the faRM (fully atomic Refined Model) of the EmrE homodimer was challenging due to the low resolution of the crystallographic X-ray structure ($\sim$ 4 Å)(1). The X-ray structure only reported backbone alpha-carbon (C_α) positions for a subset of residues, requiring that the remaining backbone atoms as well as all side chains be modeled. In an initial naïve approach, the positions of the C_α atoms from the TPP^+-bound X-ray crystal structure (PDB: 3B5D)(1) were used to restrain a model of EmrE generated using a combination of `psfgen`(2) and `Modeller`(3, 4) software. In this manner, the available C_α positions served as a template to guide the positions of the side chains as they settled into their preferred rotameric states. The resulting structure from this naïve approach presented a number of unsatisfying features. Due to the strength of backbone hydrogen bonds in a low dielectric environment within the hydrophobic core of the membrane, transmembrane helices are expected to be largely α-helical (5, 6). This initial structure lacked membrane-spanning α-helices throughout the hydrophobic transmembrane regions of the protein (Fig. S1). Furthermore, the X-ray crystal structure placed a number of proline residues in the loop regions between helical domains (Fig. S1), rather than as helix terminators where they are far more commonly found(7). Thus, considerable effort was invested to use available experimental information and trends from other membrane proteins to refine the model to recapitulate a native-like state.

Beyond the X-ray crystal structure, a cryo-EM-based model of the EmrE dimer is available(9), as well as the original electron density used to construct that model(11). We combined this cryo-EM map with the MDFF(12) approach to refine the initial structural model derived from the X-ray structure. MDFF uses a set of external forces determined from an experimentally-derived electron density map to steer atoms to areas of high electron density(12). The refinement proceeded through two broad steps: simulation of the naked protein in a high-dielectric ($\epsilon = 80$) implicit solvent(13, 14), and embedding this model into an atomistic membrane representation for further relaxation while retaining electron-density restraints. Splitting the refinement into two parts permitted us the GPU implementation of implicit solvation in `NAMD`(15) to correct secondary structure and make the structural changes needed to locally refold EmrE interactively. After these large structural adjustments were achieved, further equilibration was carried out in an explicit membrane environment. During both steps, local structural restraints maintained the proper amino acid chirality and prevented the formation of *cis*-peptide bonds(16).

The changes in protein structure were driven through fitting the heavy atoms of the EmrE dimer into the experimental electron density (Fig. S2A) assisted by applying interactive temporary forces via `VMD`(10, 17). The additional forces were designed to drive the system toward favorable secondary structure. A common scenario for adding temporary forces interactively involved flipping the orientation of carbonyl oxygen atoms to promote the formation of α-helical structures, which would otherwise be disfavored due to the large energetic barrier to flipping the peptide backbone. Due to the geometrical constraints placed upon each residue by the protein environment, the flipping process has a high transition barrier, and therefore is unlikely to occur during conventional simulation. Other temporary interactive forces involved rotating bonds to bring aromatic side chains into the hydrophobic core of the membrane, where those side chains have been shown to partition naturally into lipid bilayers(18). By the exploratory nature of the refinement process, many parameters were changed during the simulation. Most notably, the coupling constant between the electron density and the heavy atoms (GSCALE in `NAMD` parlance) was not always held fixed, increasing briefly from 0.3 to 10 before being reduced back to 0.3 (Fig. S2B). In this case, high coupling between the atoms and the electron density drew protein termini to stray density elements on the periphery of the simulation box (Fig. S2A). Thus, while the cross-correlation coefficient in this case is lower (Fig. S2B), a reduced coupling constant yielded a compact EmrE dimer that was suitable to embed into a lipid environment. Nevertheless, the electron density correlation that results from this procedure (0.45) is a significant improvement on the starting value of 0.165 for the published cryo-EM model(9), which is itself an improvement over the X-ray based structure (0.138).

The lipid embedding process was carried out in `CHARMM-GUI`(19, 20), where eighty 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) lipids were added to each of the top and bottom leaflets to form the full bilayer. This shorter lipid was chosen to mimic more closely the experimental conditions under which EmrE has been studied, where shorter lipid tails predominate(21–25), rather than the longer lipids found in native bacterial *E. coli* membranes(26). Even in this comparatively thin membrane, the EmrE dimer is fully embedded within the hydrophobic core of the bilayer due to its small size (Fig. S1). The membrane-protein system was solvated with 9,942 TIP3 water molecules(27) and sufficient sodium chloride (Na^+Cl^-) for a concentration of 150 mM and a neutral simulation system. The total system size was approximately 50,700 atoms, with dimensions of 82 Å × 82 Å × 85 Å.

This membrane-embedded system was advanced forward through 18.8 ns of simulation time with the MDFF biases applied non-interactively. `NAMD 2.10`(2) was used to propagate atomic coordinates with 2 fs time steps using the CHARMM36 force field for proteins(28) and lipids(29) with TIP3 water(27). The equilibration simulation was maintained in a constant number, pressure, and temperature (NPT) ensemble using a Langevin thermostat(30, 31) with a damping coefficient of 1/ps to maintain the temperature at 310 K and a Nosé-Hoover Langevin piston barostat(32, 33) set to hold the membrane aspect ratio constant. The cryo-EM-derived electron density for EmrE(11), applied as a grid potential(12), maintained the overall topology of the helices as lipids equilibrated around them. Initially, a serine residue (S64) on Helix 3 of monomer "A" ($S64^A$) was oriented in such a way that it interacted directly with the membrane, exposing its hydroxyl to the membrane. Given that S64 of the "B" monomer ($S64^B$) was buried within the protein, such an orientation would require that Helix 3 rotates substantially about the helical axis during the conformational transition, in addition to exposing the serine hydroxyl to membrane acyl tails. This unnatural conformation was alleviated through a rotation of 60 degrees about the helical axis using the collective variables module of `NAMD`(2, 34) to allow S64 interaction with S64 of the opposite monomer (Fig. S3). The final faRM is presented in Supporting Information in both binary (js) and human readable (pdb) formats.

Development of the faRM represents a key initial step in our computational approach toward understanding the mechanism

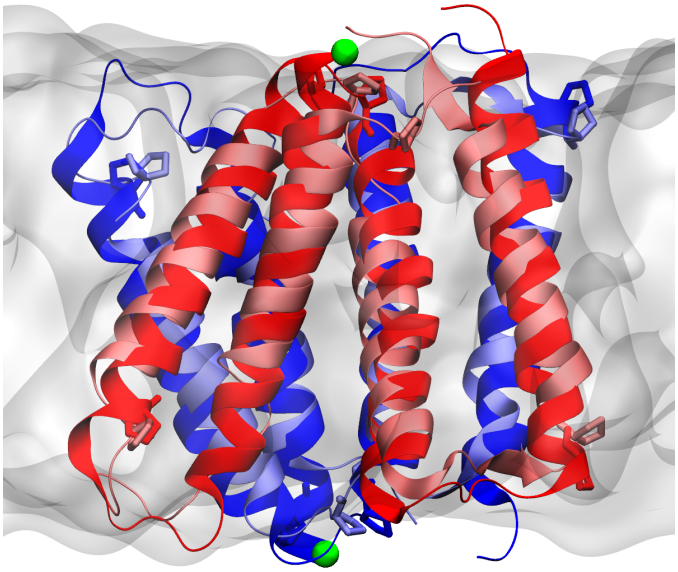

Fig. S1. Direct structural comparison of the initial structure based on a naïve refinement of the original crystal structure (lighter colors) with final faRM obtained after interactive MDFF simulations (darker colors). The protein structure is represented as a cartoon, where α-helical secondary structures determined by `Stride`(8) are clearly demarcated from loop regions. The extent of the hydrophobic acyl chains of the modeled lipid (DMPC) bilayer after simulation is shown as a transparent surface around the EmrE homodimer. The monomers are identified by their color (A is blue, B is red), using the chain identifier from the X-ray structure of the EmrE TPP^+-bound state (PDB: 3B5D)(1), which is also consistent with the cryo-EM based model (PDB: 2I68)(9). A green sphere has been drawn on the N-terminus of each monomer to help identify the loops and the orientation of each monomer. Proline residues 3, 32, 55, and 86 are also drawn in a stick representation to highlight their important role in terminating helices in the faRM. Supplementary Animation 1 shows this representation rotated around the membrane normal, permitting for close inspection of the faRM relative to its membrane environment. The backbone root-mean-square deviation (RMSD) between these two model structures is 3.8 Å.

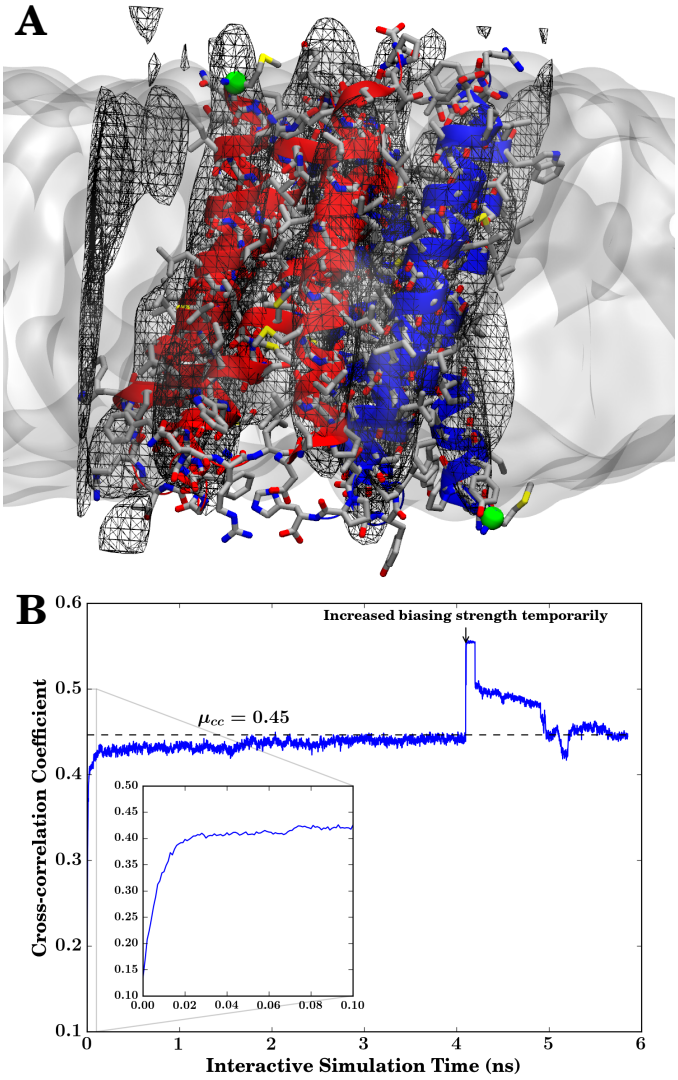

Fig. S2. (A) The cryo-EM derived electron density (black wireframe) overlaid on the faRM (colored as in Fig. S1), with the addition of a stick model for the side chain heavy atoms. Carbon atoms are gray, nitrogen atoms are blue, oxygen atoms are red, and sulfur atoms are yellow. The extent of the membrane hydrophobic core surrounding the protein within the simulation system is shown as a transparent surface. Note that the reported cryo-EM densities show additional density outside the dimer. Supplementary Animation 2 shows this representation rotated around the membrane normal. (B) Cross-correlation of simulated density maps from the simulation structure during interactive MD(10) in implicit solvent against the experimental map over time. At t=0, the cross-correlation coefficient is very low (0.165) for the published cryo-EM model(9), compared with 0.138 for the naïve crystal structure model(1). The mean cross-correlation at the end of the interactive MD in implicit solvent substantially improved ($\mu_{cc} = 0.45$), with the level shown by the dashed black line. The cross-correlation coefficient remains consistent during membrane equilibration in the subsequent step.

of proton-coupled transport in EmrE, as the previously available structural information is incomplete and contradictory. Alongside a TPP^+-bound structure, there are two apo crystal structures where two helices align perpendicularly to the others, and would lie along the membrane-water interface(1). In this case, crystallographic contacts may have forced EmrE to adopt a non-native structure that perturbs its secondary structure, resulting in a compressed protein structure with partially unwound helices. The refinement process used here promotes helical secondary structure and helix extension across the full span of the bilayer hydrophobic region through additional restraints (Fig. S1). With additional experimental inputs, such as ongoing NMR studies, future models can more accurately determine the extent of these helices. By leveraging available experimental inputs to guide the model, we arrive at a stable EmrE dimer that corrects secondary structure deficiencies in the crystal structure. This structure represents a significant improvement upon previously reported EmrE models, where secondary structure disintegrates upon simulation(35).

Simulation Protocols

The faRM described above forms the basis for studying the impact of protonation and substrate binding on the intramolecular interactions within EmrE, and how those interactions may regulate conformational changes consistent with the alternating access mechanism. The first step to prepare the structural model for simulation is to construct apo, substrate-bound, and singly and doubly protonated EmrE faRM. Three different EmrE protonation states were prepared using `psfgen`: 1) only protonating $E14^A$ (A^+), 2) only protonating $E14^B$ (B^+), and 3) protonating both E14 residues (A^+B^+). The tetraphenylphosphonium (TPP^+) cation was modeled in by aligning the TPP^+-bound X-ray structure (PDB: 3B5D)(1) to the faRM, and adopting the TPP^+ coordinates from the aligned structure. Water molecules within 1.5 Å of the newly placed TPP^+ were then removed.

In addition to the CHARMM36 force field for proteins(28), lipids(29), and TIP3 water(27), parameters for the substrate TPP^+ were determined using CGenFF(36). TPP^+ contains a phosphorus atom at the center of four phenyl rings. This phosphorus atom has no analogous atom types in CGenFF(37, 38), and required parameterization. Using the ffTK force field development toolkit(39), the charges and missing parameters were optimized to approximate the quantum mechanical interactions inherent to the TPP^+ cation in a classical force field, focusing on the central phosphorus atom. The parameters obtained are provided as Supporting Information.

Equilibrium Simulation. The five systems (apo, TPP^+-bound, and 3 different E14 protonation states) were equilibrated for 25 ns using `NAMD 2.10`(2). During equilibration, the protein was restrained to the EM electron density using MDFF with coupling constant at 0.3, which maintained the protein structure near that of the faRM to allow lipids and waters to equilibrate around the EmrE dimer. During equilibration under NPT conditions, the temperature was maintained at 310 K, and the pressure at 1 atm by Langevin dynamics and Langevin piston Nosé-Hoover methods, respectively(40, 41). The pressure operated on the membrane-normal and membrane-parallel dimensions independently. Short-range electrostatics were coupled to long-range particle mesh Ewald (PME)(42, 43) electrostatics at 10 Å, with the PME grid set to a 1.2 Å spacing. Dynamics used a 2 fs time step, and the SETTLE algorithm(44) constrained the bond lengths to hydrogen atoms.

The equilibrated structures for each of the five simulation systems were converted into a `GROMACS`-compatible format using TopoGromacs(45) for further simulation. Using `GROMACS 5.0.4`(46–48), five 500 ns simulations were performed for each of the five loading states of EmrE, resulting in an aggregate run time of 12.5 μs. The extensive sampling took advantage of new optimizations in GROMACS 5 that reduce computational walltimes for equilibrium simulations(46). `GROMACS` simulations were carried out in a constant volume and temperature (NVT) ensemble, using a Nosé-Hoover thermostat(49–51) set to 310 K. Electrostatic interactions were computed as described above, with 1.2 Å grid spacing for PME after a 12 Å cutoff for short-range electrostatic and non-bonded interactions.

Driven EmrE Structural Transitions. EmrE is unusual in that it is an antisymmetric homodimer(21), which causes the outward- and inward-facing states of EmrE to be related by symmetry. Driven simulations were used to determine the relative ease of transition between the inward- and outward-facing states depending on protonation and substrate binding. These driven simulations used the final frames of the equilibrium simulations as their initial state, and used symmetry operations to construct the target state from these initial states. To create the target state, the conformations of the individual monomers were swapped, with monomer A given the coordinates of monomer B and vice versa. The target state was then fit to the initial state to minimize the overall conformational change required, which preserved the original membrane topology in the final state.

Biases to drive the conversion from the initial state to the target state were applied over 20 ns of simulation using the collective variables (COLVARS) module(34) of `NAMD` such that the root-mean-square deviation (RMSD) to the target reached 2 Å. The TPP^+ molecule, if present, was translocated to the other side of the membrane by an additional bias applied to the center-of-mass of the molecule. Combined, this protocol represents an approximate way of driving the transition of EmrE from the outward- to inward-facing state. In this setup with no electrical or chemical gradient driving the transition, the free energy difference between the two states is implicitly zero due to the symmetry relations between the two states. Thus, rather than using the non-equilibrium work for the transition to supply an upper bound on the free energy difference(52), we use the non-equilibrium work instead to rank the height of the free energy barrier for each transition, connecting trends and interactions observed in equilibrium simulation and NMR experiments(25).

Intermonomer E14 pK_a Calculations. Prior experiments identified a pK_a difference between the two symmetry-related glutamate residues present within the transmembrane portion of EmrE. These residues display pK_a values shifted by 1.2 to 1.7 units with

respect to one another, depending on the temperature(53). To determine which monomer has which pK_a, a replica exchange thermodynamic integration (RETI)(54) calculation was conducted in `NAMD 2.10`(2) to determine whether E14 from monomer A or B ($E14^A$ or $E14^B$) will be protonated first during the transport process. When the transition parameter $\lambda = 0$ in the perturbation calculation, the proton occupies $E14^B$ (equivalent to the B^+ state). That proton transits in an alchemical manner to $E14^A$ when $\lambda = 1$ (the A^+ state), with $\lambda = 0.5$ representing a state where half of a proton is simultaneously on both $E14^A$ and $E14^B$. `NAMD` uses the dual-topology paradigm for alchemical simulations. Typically, at intermediate λ, there would be a total of four distinct E14 side chains simulated: two representing the protonated and deprotonated forms of $E14^A$, and two representing the protonated and deprotonated forms of $E14^B$. However, aside from the proton that differs between the states and the different charges on glutamate heavy atoms that result, the heavy atoms are identical. Thus, to limit sampling of largely decoupled side chains at extreme λ, we applied a zero-length 20 kcal/mol/Å^2 harmonic bond between equivalent heavy atoms to "pin" the heavy atoms of the protonated and deprotonated side chains on each monomer together. This additional set of restraints eliminates unphysical conformational sampling that can result in slow convergence of the calculations(55–57).

In order to optimize the transition rate while being mindful of the computational cost, the λ values for RETI simulation were chosen to be $\lambda \in \{0.0, 0.02, 0.08, 0.14, 0.2, 0.26, 0.32, 0.38, 0.44, 0.50, 0.56, 0.62, 0.68, 0.74, 0.80, 0.86, 0.92, 0.98, 1.0\}$, yielding exchange acceptance probabilities of at least 7% between adjacent replicas with frequent exchanges between replicas (Fig. S4). The initial coordinates were drawn from the A^+ and B^+ simulations (10 states from A^+, 9 from B^+) to minimize perturbation to the water network during the alchemical transition between these two states. Each of the 19 replicas was simulated for 10 ns to obtain the final result.

Analysis Methodology. Purpose-built `VMD`(17) scripts were written to analyze protein dynamics, water permeation, and connectivity during the equilibrium trajectories. These scripts utilized the python interface of `VMD`(17) to facilitate interoperability with the `NumPy`(58), `SciPy`, `NetworkX`(59), and `Matplotlib`(60) `python` packages for further analysis. The analyses performed include simple RMSD, root mean square fluctuation (RMSF), and distance measurements, as well as more elaborate evaluation of the geometries observed during simulation. To determine if water wires formed and if they might be conducive to proton transfer, directional networks of hydrogen bonds were constructed using `NetworkX`(59), and evaluated for connectivity between distant water molecules on either side of the transporter.

Specifically, we consider a conformation to be leaky to water if a trail of water molecules whose oxygen atoms are within 3.2 Å of their neighbors exists such that they connect any pair of water molecules on opposite sides of the lumen. In numerical terms, the analysis examines water molecule pairs where $|z| > 11.8$ Å for each molecule and $\Delta z > 23.6$ Å between the two molecules, with the membrane midplane located at z=0 Å. Additionally, a single snapshot of EmrE is assumed to permit proton leakage if any two water molecules fulfilling the same criteria can be connected via a series of directional hydrogen bonds that might allow rapid proton traversal of the membrane span. The statistics are computed by evaluating the existence or absence of water pores or proton conduction pathways every 5 ps, the frequency at which snapshots from the trajectory were saved. Since proton-conduction pathways are uniquely short lived, their mean duration is likely overestimated.

A similar approach was employed to compute the water-mediated hydrogen bonds between residues within EmrE. For every snapshot from the trajectory, directional hydrogen bonding networks were generated based on the geometry of each protein residue or water molecule, using only the protein side chains, excluding hydrogen bonds to the backbone. The pathlength between two residues was determined from this directional network. The pathlength represents the number of edges in the path between two residues, or equivalently the number of intermediate waters in the interaction plus one. Note that since the hydrogen bonding pathway is directional from donor to acceptor, we only count the interaction if a donor-acceptor relationship might exist between the residues. This excludes water mediated interactions between two donors or two acceptors.

Contacts between TPP^+ and EmrE were evaluated using a distance-weighted contact criteria:

$$C_i = \sum_{j \in \mathrm{TPP}^+} \frac{1}{1 + \exp\left(5\,\text{Å}^{-1}\left(d_{ij} - 4\,\text{Å}\right)\right)} \qquad [S1]$$

In this manner, every contact between protein heavy atoms (i) and heavy atoms in TPP^+ (j) can be aggregated together to visualize and quantify the extent of the contacts between residues and TPP^+. This formulation has been used previously for quantifying lipid-protein(61) and protein-lignin interactions (62), as it more strongly weights shorter interactions such as hydrogen bonds or π-stacking relative to longer-range, usually nonspecific contacts.

Data Availability. Highly strided equilibrium trajectories are provided for public use at http://www.ks.uiuc.edu/~jvermaas/EmrE/trajectories.zip. Non-strided trajectories are availabile upon request.

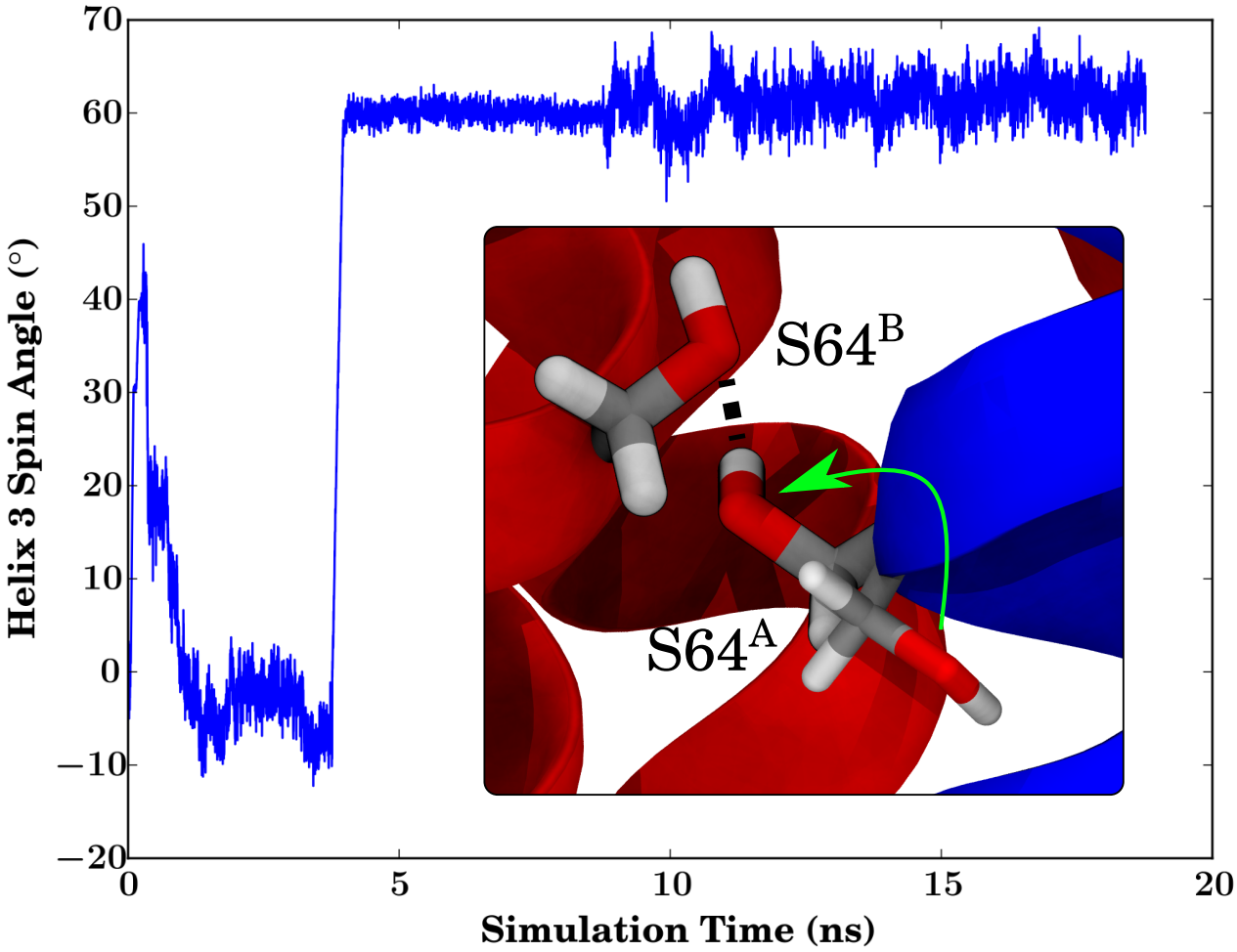

Fig. S3. Time series showing the rotation of helix 3 until S64A no longer faces the membrane. This transition was attempted twice using the Colvars module of NAMD. The initial 30 degree rotation was unstable. The 60 degree rotation was restrained for a time, and was stable over the subsequent 10 ns of simulation. The comparison between the initial and final states is shown in the inset, where the position of the hydroxyl oxygen follows the green arrow. Initially, $S64^{A}$ (thinner stick representation) pointed into the membrane. After rotation, $S64^{A}$ can interact with the $S64^{B}$ (dotted black line).

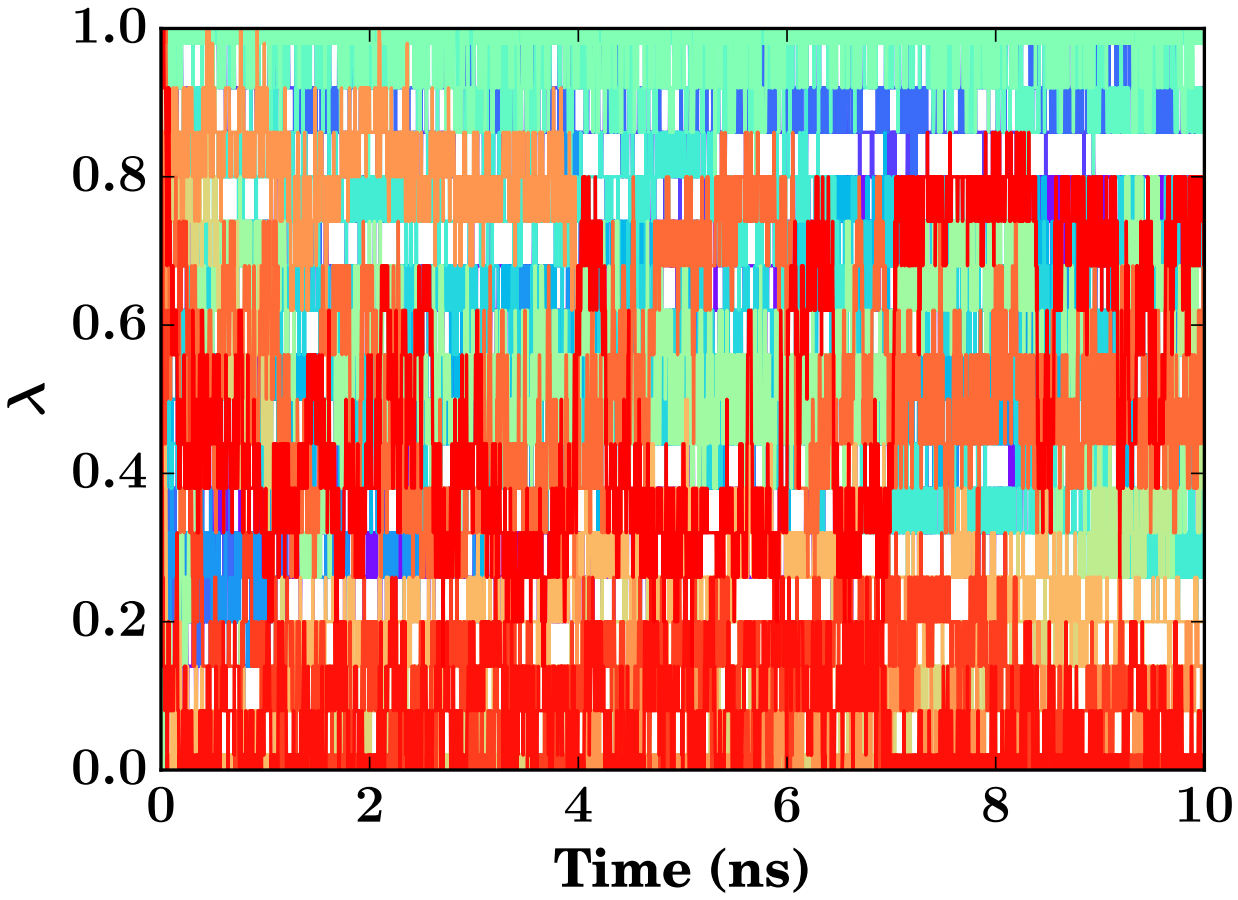

Fig. S4. Time series of the replica exchange thermodynamic integration (RETI) calculations showing how each replica exchanges through values of the alchemical transition parameter (λ) space over time. Each replica is drawn as its own color, based on the initial value for λ of the replica (red for larger λ, blue for smaller λ). Exchanges are frequent on the time scale of the simulation, and no isolated λ values are observed.

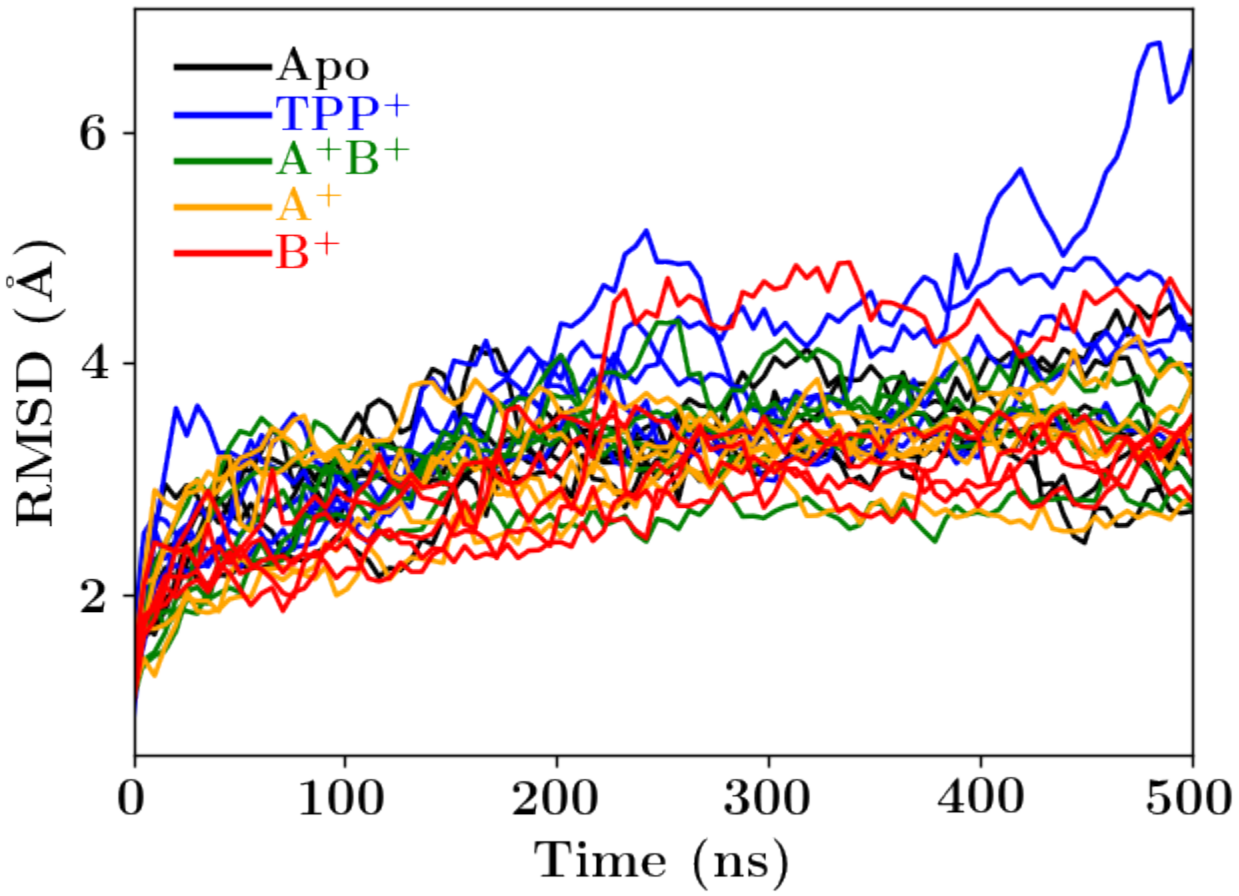

Fig. S5. RMSD of each trajectory compared against the starting faRM structure, using only the protein backbone. Similar to Fig. 2A, RMSD values comparable to the original resolution of the crystal structure are observed. The color of the lines indicates the loading state of EmrE: black for the apo state, blue for the TPP^+-bound state, green for the doubly-protonated state (A^+B^+), orange for the state where the proton is bound to monomer A (A^+), and red when the proton is bound to monomer B (B^+).

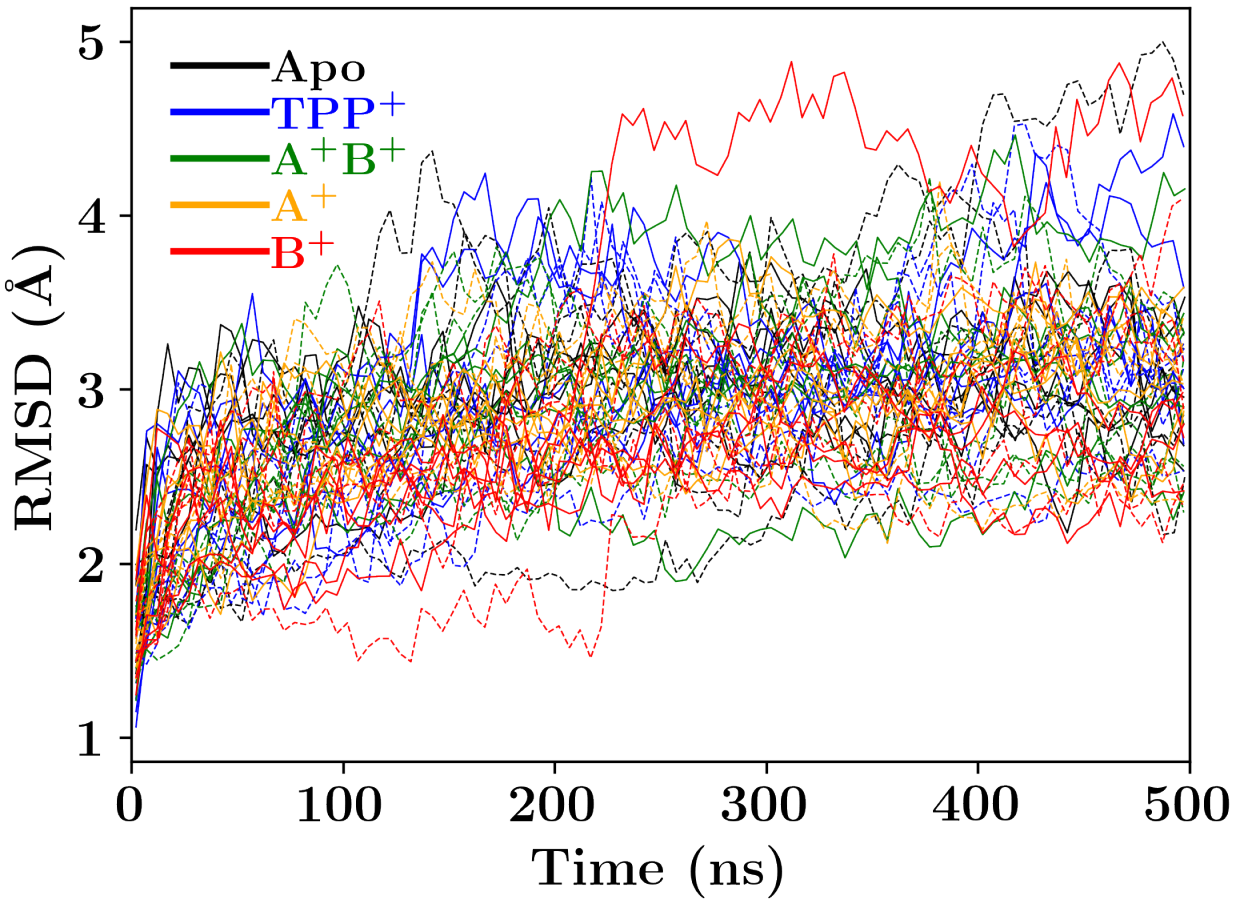

Fig. S6. RMSD of each trajectory compared against the starting faRM structure, using only the protein backbone, calculated independently for each monomer. Similar to Fig. 2A, RMSD values comparable to the original resolution of the crystal structure are observed. Solid lines indicate the RMSDs for monomer A, whereas dashed lines indicate the RMSDs considering only monomer B. The color of the lines indicates the loading state of EmrE: black for the apo state, blue for the TPP^+-bound state, green for the doubly-protonated state (A^+B^+), orange for the state where the proton is bound to monomer A (A^+), and red when the proton is bound to monomer B (B^+).

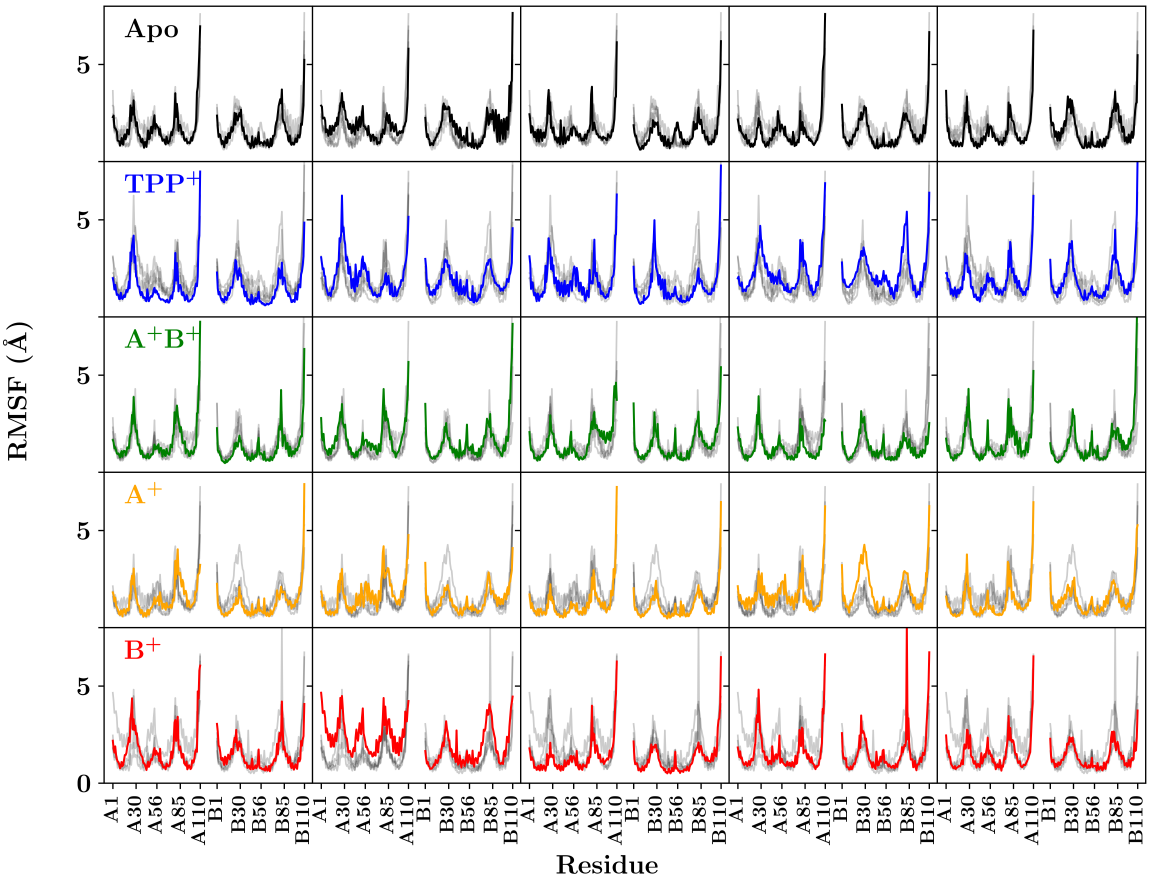

Fig. S7. RMSF per residue for individual equilibrium trajectory runs. Each different EmrE loading state is one row of the matrix, with the 5 independent trajectories occupying the columns of the matrix. The colored trace indicates the RMSF for that specific trajectory, while the gray background traces provide context for how the other 4 copies of the same loading state behaved.

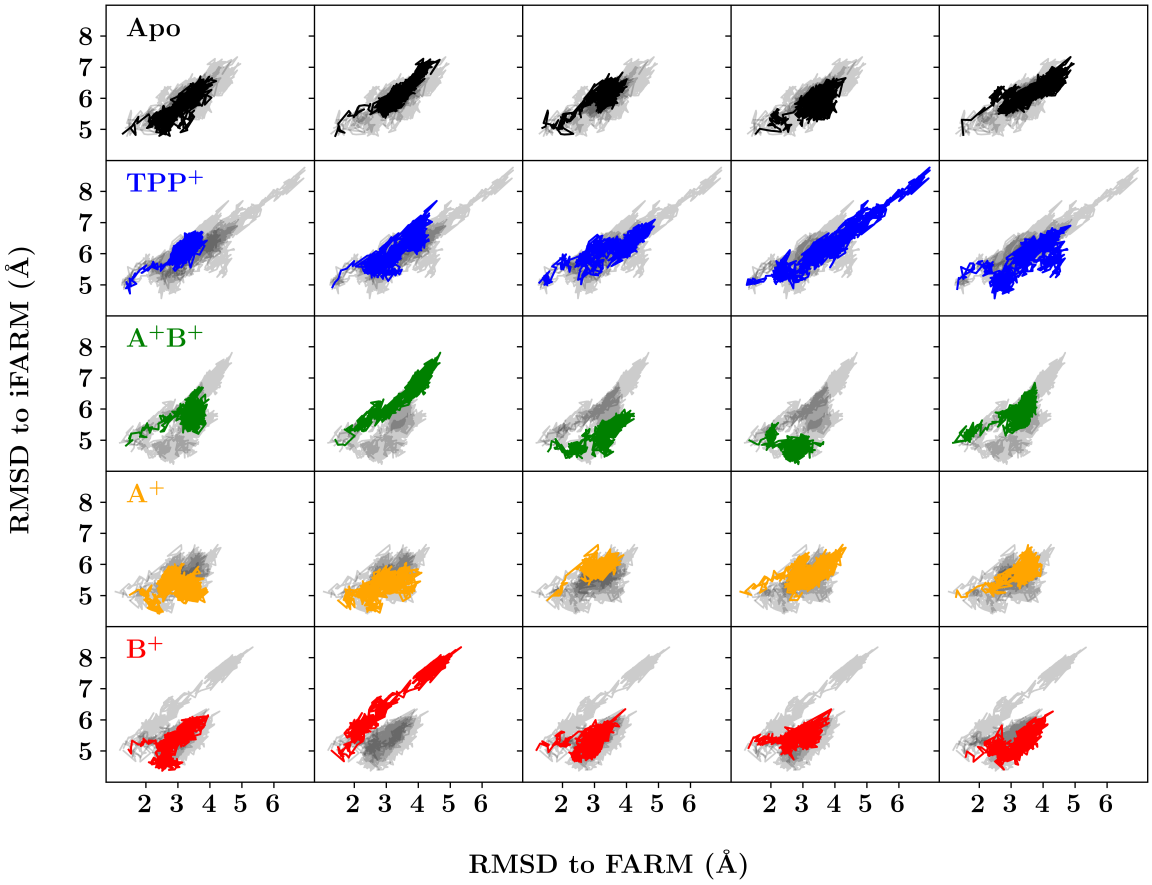

Fig. S8. RMSD matrix highlighting the states sampled by the equilibrium trajectories with respect to both the original faRM (x-axis) and the inverted faRM (y-axis). Motion towards the inverted model (low RMSD states) would be suggestive of the transition EmrE undergoes as part of its functional cycle.

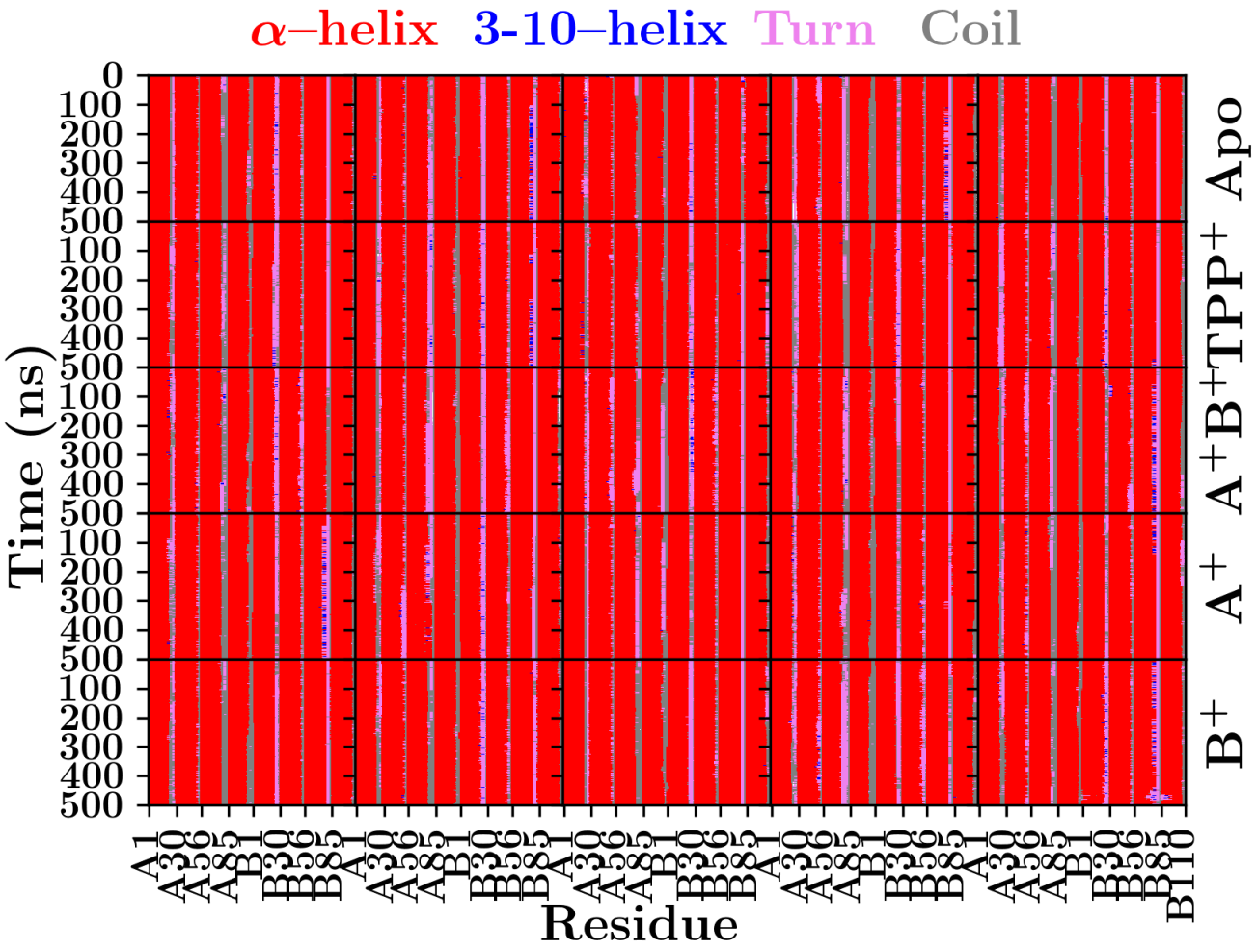

Fig. S9. Secondary structure evolution over time for the 25 individual trajectories, with the five for each loading states shown in one row, using the same color legend as Fig. 3.

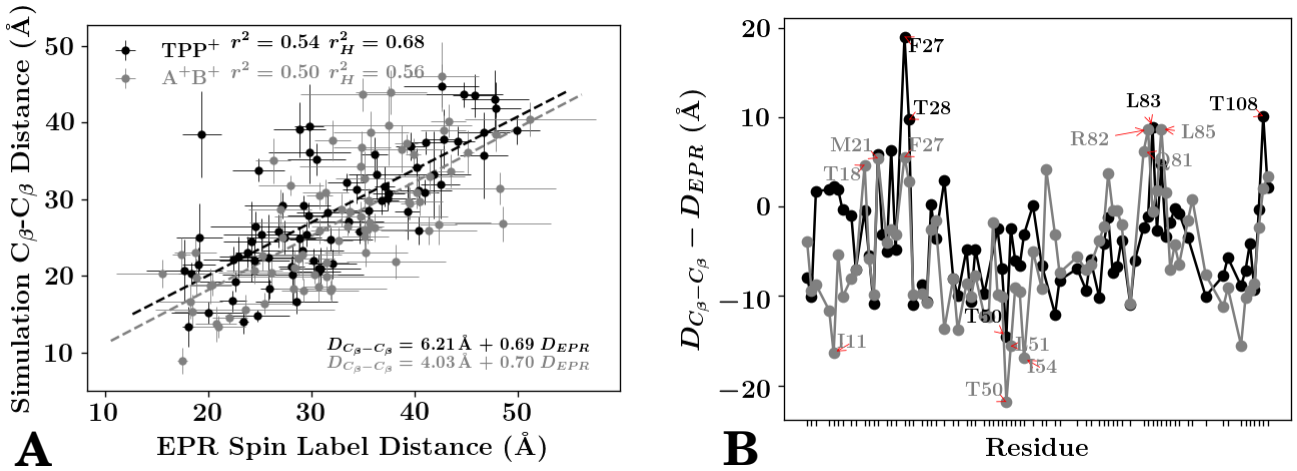

Fig. S10. Comparison between measured β carbon distances from simulation and the equivalent spin label distances from EPR studies(63). (A) Correlation plot between equivalent conditions, using the TPP$^+$-bound and pH 5 states from Dastvan et al.(63) as a comparison for the TPP$^+$-bound (black) and A$^+$B$^+$ (grey) simulated states. The equation for the dashed trendlines is given in the lower right corner, with the correlation coefficient given next to the in-figure legend for both the entire protein (r^2) as well as that of only the helical domains (r_H^2). (B) Identification of residues with the largest deviation between our simulated distances and the EPR measurements. Individual residues that differ by more than 10 Å from the mean difference are annotated by red arrows.

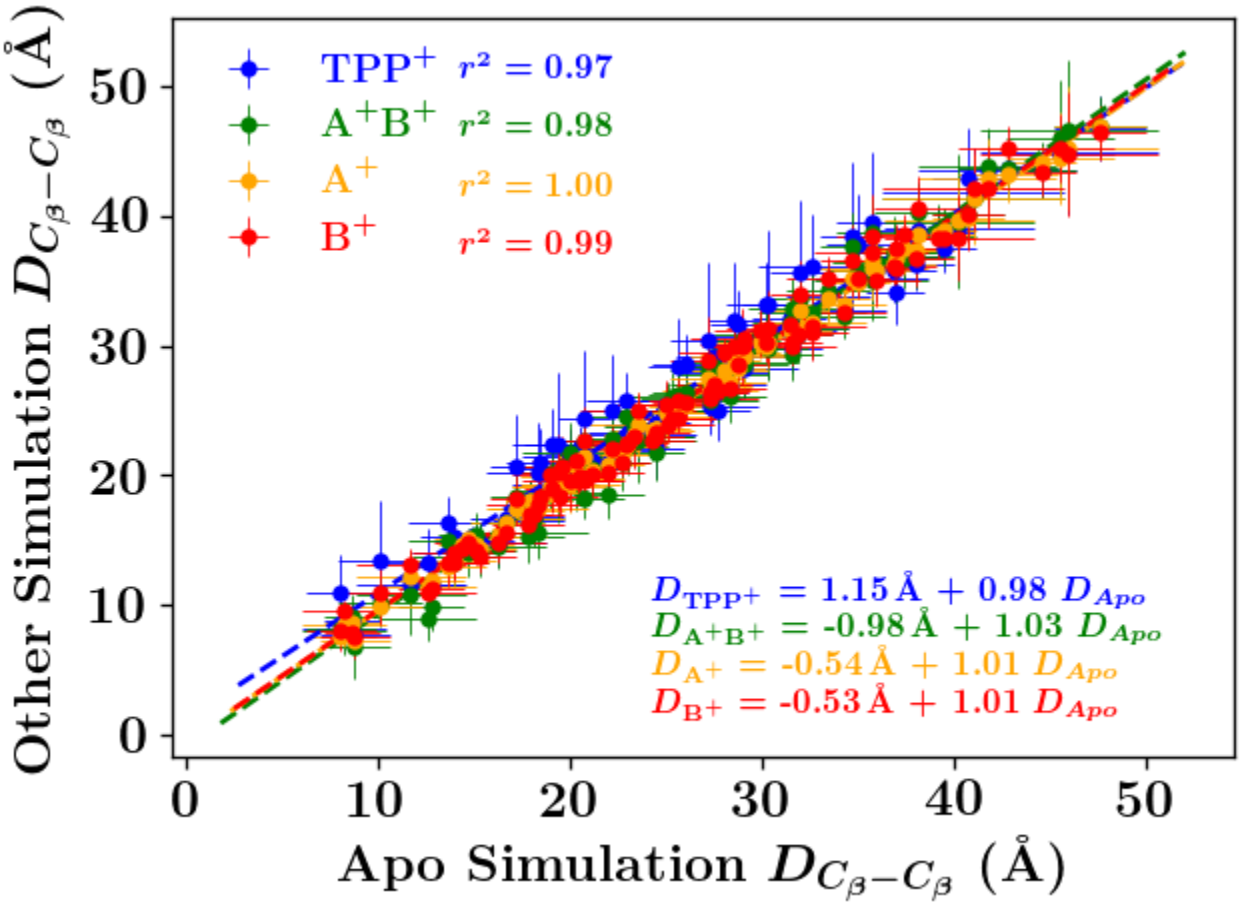

Fig. S11. Comparison of the distances observed between C_β-C_β from residue pairs on opposite monomers observed in the MD simulations. This figure is equivalent to Fig. S10A, but compares the distances from different loading conditions against the distances seen in the apo state.

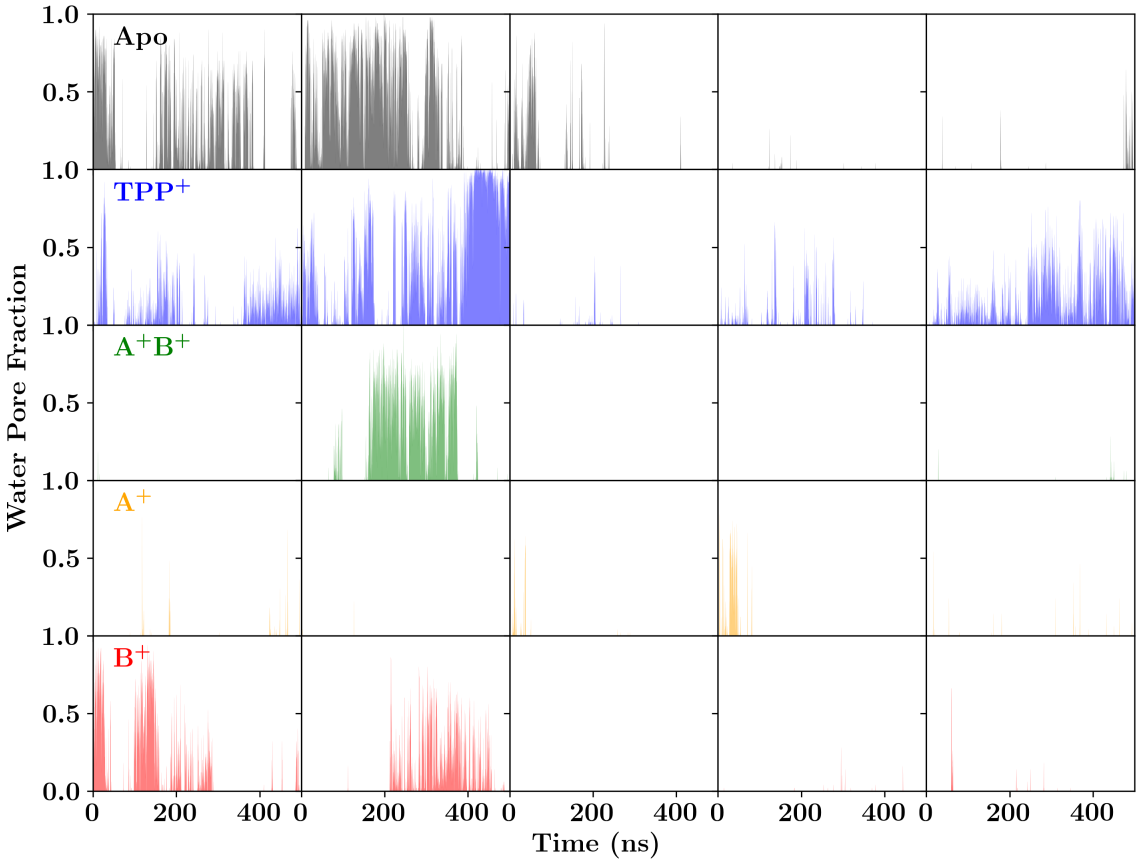

Fig. S12. Time resolved water pore formation for each simulation trajectory. Since the time resolution of the data (5 ps steps) far exceeds the limits of what is easily resolvable graphically, the individual datapoints are aggregated together in lots of 50, such that each feature is equivalent to 250 ps of simulation time. Each remaining point thus represents the fraction of time within that simulation snapshot that the water pore, as defined in Table 1 of the main text, was formed during the 250 ps of simulation.

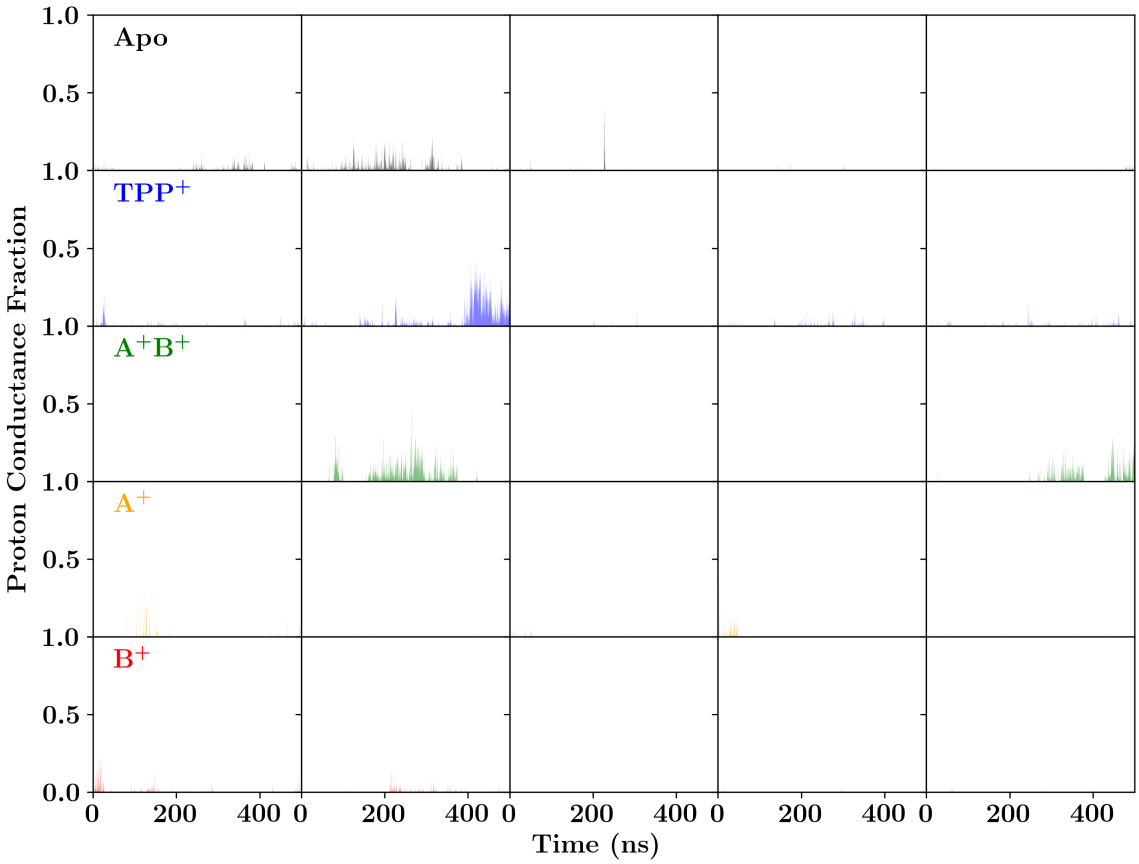

Fig. S13. Time resolved existence of water pores that can conduct protons for each simulation trajectory. Since the time resolution of the data (5 ps steps) far exceeds the limits of what is easily resolvable graphically, the individual datapoints are aggregated together in lots of 50, such that each feature is equivalent to 250 ps of simulation time. Each remaining point thus represents the fraction of time within that simulation snapshot that the water pore, as defined in Table 1 of the main text, was formed and could conduct protons during the 250 ps of simulation.

Discussion of the Environment Around H110

Contacts between the C-terminus of monomer A and $E14^B$ unexpectedly appear in some states (Table 2). These interactions do not always exist. Instead, they are the result of heterogeneous populations of conformations in the C-terminal region of EmrE. Due to the closure of the lumen to one side, only the C-terminus of monomer A can make these interactions (Fig. S14, Supplementary Animation 6). As a result, the C-terminus of both monomers effectively experience independent environments (Fig. S14), and these environments may change rapidly as the terminus is exposed to bulk solution or the protein lumen. The environmental diversity of the C-terminus may affect the terminal histidine, whose exposure to different environments can vary greatly. The heterogeneity of histidine exposure may result in alternative protonation states rather than the assumed N_ϵ protonation in this model. Further experimental studies or quantum mechanical simulation would be needed to explore this aspect.

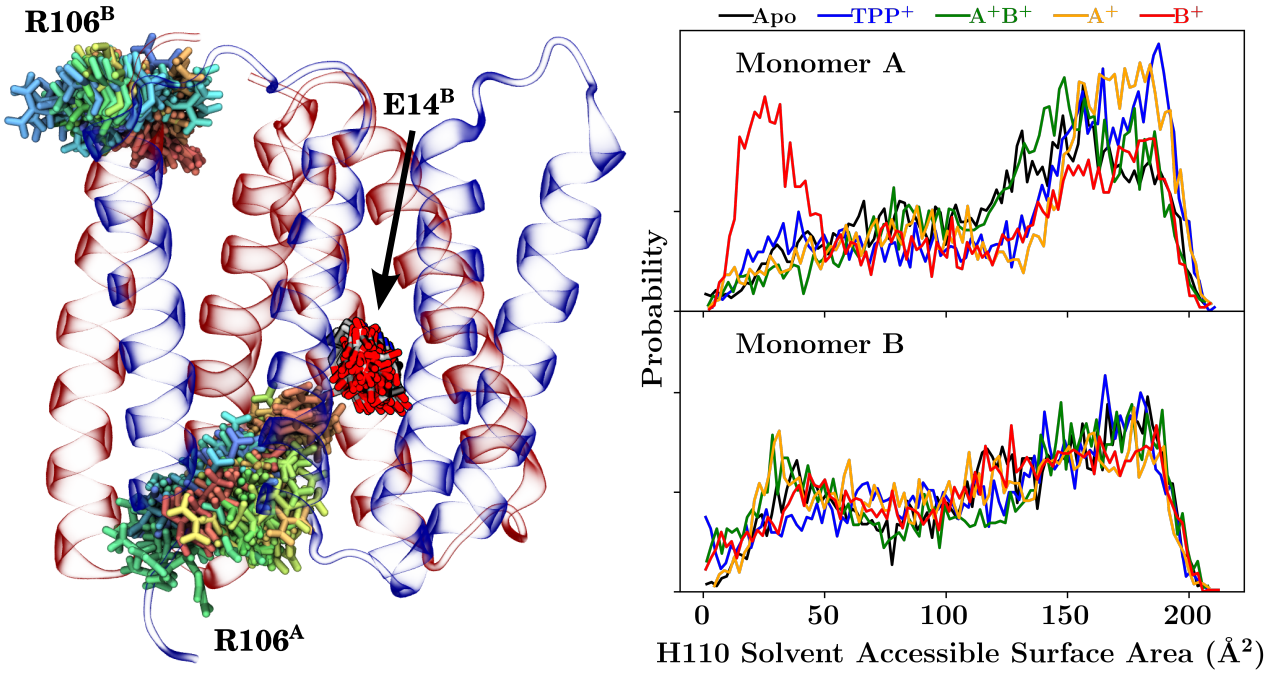

Fig. S14. Demonstration of the different motions of the C-terminus. (Left) A comparison of the motion of R106 from each monomer relative to the lumen, represented here by $E14^{B}$. The position of R106 across an apo trajectory is shown through a rainbow of residues superimposed on the same dimeric structure (transparent cartoon). A rotation of this view around the membrane normal is shown in Supplmentary Animation 6. (Right) Probability distribution of the solvent-accessible surface area of the terminal H110 residue.

Table S1. Hydrogen bond propensity for observed donor-acceptor pairs, normalized such that 100% equates to the exisitence of a single hydrogen bond between two residues for the entire simulation. The hydrogen bonds reported here are only for cases where an amino acid side chain is the donor, thereby excluding helix-stabilizing backbone hydrogen bonds.

		Sidechain Hydrogen Bond Propensity				
Acceptor	Donor	Apo	TPP-bound	$H_A^+H_B^+$	H_A^+	H_B^+
$Y6^A$	$N2^A$	7.5	7.4	9.2	5.6	6.6
$A10^A$	$S43^A$	81.8	65.6	84.4	73.9	81.5
$E14^A$	$T18^A$	63.2	65.8	75.2	60.4	63.5
$E14^A$	$Y40^A$	14.5	10.9	0.7	1.5	7.5
$E14^A$	$S43^A$	5.4	10.7	–	0.4	1.5
$E14^A$	$W63^A$	32.8	13.8	3.1	5.1	57.1
$V15^A$	$T19^A$	90.4	88.9	85.9	89.8	87.3
$G17^A$	$T36^A$	1.5	9.0	7.0	3.1	3.5
$T18^A$	$Y40^A$	5.2	13.1	33.7	19.5	16.9
$L20^A$	$S24^A$	14.4	4.0	43.6	28.2	38.8
$S24^A$	$R29^A$	0.5	1.6	11.2	7.0	13.1
$E25^A$	$K22^A$	3.7	–	29.9	20.3	19.8
$E25^A$	$R29^A$	7.6	0.2	9.0	5.1	9.8
$G26^A$	$R29^A$	2.9	4.1	24.9	20.6	14.0
$T28^A$	$S24^A$	2.0	2.7	0.1	5.5	1.3
$R29^A$	$S33^A$	74.5	56.4	59.8	74.2	67.9
$P32^A$	$T36^A$	87.4	78.4	82.4	88.5	86.7
$G35^A$	$C39^A$	63.9	62.5	60.8	61.8	65.1
$I37^A$	$C41^A$	56.9	59.6	55.5	54.9	53.0
$C39^A$	$S43^A$	1.3	1.0	5.6	7.7	1.8
$Y40^A$	$T18^A$	3.7	2.0	3.4	5.5	3.4
$S43^A$	$W63^A$	–	2.1	31.2	28.6	–
$W45^A$	$Q49^A$	33.9	30.5	37.2	36.5	31.2
$L46^A$	$T50^A$	86.3	88.0	71.0	78.7	86.1
$L47^A$	$T50^A$	3.8	0.5	16.4	11.6	1.1
$G57^A$	$N102^A$	–	0.3	10.4	–	–
$Y60^A$	$S64^A$	68.8	51.5	74.3	60.9	60.9
$Y60^A$	$N102^A$	9.3	5.2	0.2	12.0	15.2
$W63^A$	$E14^A$	–	–	35.5	35.9	–
$S64^A$	$N102^A$	1.9	4.2	8.7	14.4	7.4
$I68^A$	$S72^A$	83.8	86.2	87.8	90.0	90.5
$I71^A$	$S75^A$	63.5	64.9	63.1	67.1	67.6
$S75^A$	$Q81^A$	13.6	2.6	1.8	3.6	2.4
$W76^A$	$Q81^A$	6.9	9.1	1.7	0.3	0.8
$W76^A$	$R82^A$	–	5.4	2.8	5.1	–
$F78^A$	$Q81^A$	10.9	10.2	1.5	2.5	1.6
$F79^A$	$Q81^A$	3.2	8.9	4.2	5.4	4.1
$G80^A$	$S75^A$	5.7	3.6	0.4	0.1	0.1
$Q81^A$	$R82^A$	5.6	5.0	2.0	0.4	0.7
$D84^A$	$R82^A$	11.4	21.6	21.5	28.7	18.4
$M91^A$	$C95^A$	33.5	50.6	42.6	40.8	28.9
$V98^A$	$N102^A$	5.0	2.7	13.3	3.0	2.8
$I101^A$	$S105^A$	58.4	58.0	53.7	63.4	61.1
$N102^A$	$R106^A$	28.5	26.6	10.6	34.4	17.5
$N102^A$	$T108^A$	–	–	9.8	–	–
$L103^A$	$S107^A$	38.9	34.2	27.0	27.4	38.3
$L104^A$	$S107^A$	2.1	4.0	4.8	7.8	2.1
$L104^A$	$T108^A$	15.1	19.0	11.1	6.9	15.2
$R106^A$	$T56^A$	1.4	2.2	13.5	6.7	3.7
$T108^A$	$H110^A$	8.5	9.0	9.2	9.6	11.1
$H110^A$	$T56^A$	–	–	11.3	–	–
$H110^A$	$R106^A$	–	–	29.2	–	–
$E14^A$	$Y60^B$	8.6	0.1	10.3	3.5	32.2
$T18^A$	$Y60^B$	1.3	–	7.8	9.2	2.9
$Y60^A$	$Y40^B$	–	–	4.5	3.3	7.0
$S64^A$	$S64^B$	0.1	–	8.8	0.3	0.6
$S75^A$	$T56^B$	6.5	11.8	3.2	4.3	1.9
$Q81^A$	$T56^B$	–	–	0.1	9.1	0.3
$D84^A$	$T108^B$	4.5	10.2	13.3	12.1	13.8
$N102^A$	$S72^B$	7.8	–	0.1	–	3.1
$H110^A$	$K22^B$	–	–	15.4	10.7	–
$H110^A$	$R82^B$	32.8	32.5	43.1	65.4	32.3

E14^{B}	Y60^{A}	–	11.7	15.9	0.1	0.1
E14^{B}	R106^{A}	6.1	27.2	–	0.2	0.3
A48^{B}	K22^{A}	–	–	0.3	8.0	0.1
Q49^{B}	K22^{A}	–	3.4	–	3.1	5.4
L51^{B}	K22^{A}	7.4	2.3	4.3	6.3	10.9
T56^{B}	K22^{A}	3.4	1.1	6.2	3.8	2.8
Y60^{B}	Y40^{A}	–	–	5.2	0.4	0.6
Y60^{B}	S64^{A}	0.6	7.4	0.8	1.7	0.4
S64^{B}	N102^{A}	10.3	2.8	2.5	0.1	0.5
S72^{B}	N102^{A}	4.8	1.1	0.2	3.8	11.9
S72^{B}	S105^{A}	0.3	0.4	1.6	10.6	–
D84^{B}	R106^{A}	–	–	19.9	5.3	–
D84^{B}	T108^{A}	7.1	–	1.8	0.5	0.8
H110^{B}	R82^{A}	93.3	38.7	65.7	54.5	68.3
Y4^{B}	N102^{B}	6.9	10.7	9.3	8.3	8.6
Y6^{B}	N2^{B}	6.1	5.8	4.7	3.5	6.2
A10^{B}	S43^{B}	59.6	65.5	87.0	74.0	65.0
E14^{B}	T18^{B}	57.8	58.7	71.4	60.1	54.2
E14^{B}	Y40^{B}	6.8	9.2	0.2	9.4	1.5
E14^{B}	S43^{B}	22.1	16.2	–	9.0	1.4
E14^{B}	W63^{B}	15.9	8.7	2.0	44.2	0.4
V15^{B}	T19^{B}	87.2	83.5	87.8	89.3	81.9
G17^{B}	T36^{B}	5.1	13.4	1.8	3.4	6.4
T18^{B}	Y40^{B}	6.1	13.1	32.5	6.7	4.2
L20^{B}	S24^{B}	17.2	17.8	35.9	21.4	8.1
M21^{B}	S24^{B}	15.1	22.0	13.0	10.6	20.6
S24^{B}	T28^{B}	7.6	1.6	1.4	4.2	3.1
E25^{B}	T28^{B}	26.7	44.6	25.9	17.1	32.2
E25^{B}	R29^{B}	91.0	74.7	38.0	34.0	93.3
G26^{B}	S24^{B}	4.1	–	2.2	19.2	15.5
R29^{B}	S33^{B}	82.1	77.3	68.8	75.6	83.0
P32^{B}	T36^{B}	82.0	72.3	84.9	84.2	84.9
S33^{B}	S24^{B}	2.2	10.2	1.5	1.3	7.6
G35^{B}	C39^{B}	62.5	58.9	61.6	62.3	61.2
I37^{B}	C41^{B}	61.4	54.4	50.1	62.4	51.9
C39^{B}	S43^{B}	1.1	2.0	1.8	3.0	8.7
Y40^{B}	T18^{B}	2.3	2.0	3.8	4.2	5.6
S43^{B}	W63^{B}	5.6	0.9	61.7	3.9	41.8
W45^{B}	Q49^{B}	24.0	22.7	33.3	31.5	31.4
L46^{B}	T50^{B}	93.3	92.5	66.3	91.4	83.3
L47^{B}	T50^{B}	0.1	0.3	23.7	1.7	9.2
Y60^{B}	S64^{B}	58.4	68.5	57.9	57.6	49.2
W63^{B}	E14^{B}	–	–	69.7	–	50.6
I68^{B}	S72^{B}	78.3	87.8	85.3	87.1	84.3
I71^{B}	S75^{B}	74.7	71.6	70.2	77.4	73.2
W76^{B}	Q81^{B}	2.6	5.1	5.2	0.3	1.0
F79^{B}	R82^{B}	12.4	8.4	6.9	5.6	10.2
G80^{B}	R82^{B}	3.2	2.4	1.2	0.8	11.8
Q81^{B}	W76^{B}	11.6	5.9	8.0	12.2	20.3
Q81^{B}	R82^{B}	7.2	6.7	8.5	3.0	3.2
D84^{B}	R82^{B}	58.9	68.5	100.1	119.9	25.5
M91^{B}	C95^{B}	57.2	56.6	60.4	53.0	58.5
V98^{B}	N102^{B}	8.1	8.1	7.6	4.7	6.7
I101^{B}	S105^{B}	46.8	29.1	44.4	45.4	55.5
N102^{B}	Y4^{B}	9.9	3.8	6.3	6.4	4.5
N102^{B}	R106^{B}	17.1	13.7	17.6	7.2	18.8
L103^{B}	S107^{B}	40.9	36.9	40.9	43.0	39.3
L104^{B}	S107^{B}	9.6	2.1	2.0	2.5	2.7
L104^{B}	T108^{B}	26.6	32.7	23.0	24.7	23.8
S105^{B}	Y4^{B}	1.3	5.4	2.5	1.7	2.1
T108^{B}	H110^{B}	7.3	8.6	6.2	7.8	7.1
P109^{B}	H110^{B}	3.6	2.7	4.6	5.2	3.3
H110^{B}	R106^{B}	–	16.7	–	–	–

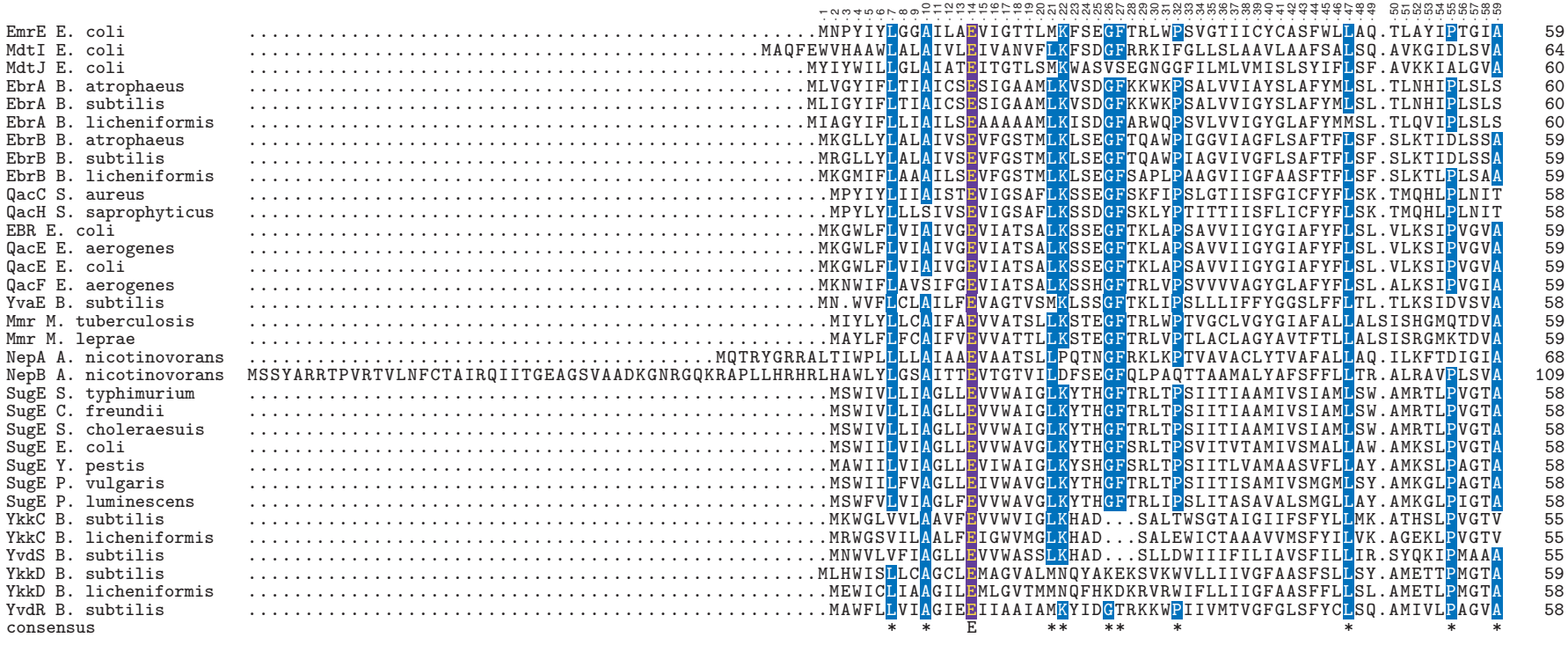

Fig. S15. Sequence alignment for diverse members of the small multi-drug resistance (SMR) transporter family, highlighting regions of sequence conservation above 75% (blue) or always conserved (purple), with residue numbering relative to EmrE. Seqences taken from reviewed sequences within the Uniprot sequence database.

EmrE E. coli YAIWSGVGIVLISLLSWGFFGQRLDLPAIIGMMLICAGVLIINLLSRSTPH..... 110
MdtI E. coli YALWGGFGIAATLAAGWILFGQRLNRKGWIGLVLLLAGMIMVKLA..... 109
MdtJ E. coli YALWEGIGILFITLFSVLLFDESLSLMKIAGLTTLVAGIVLIKSGTRKARKPELEVNHGAV 121
EbrA B. atrophaeus YATWSGVGTVLTAVIGVKWFKELNAKGLIGILLLISGVVLLNWQ..... 105
EbrA B. subtilis YATWSGAGTVLTTVIGVKWFKEDLNAKGLIGILLLLSGVVLLNWP..... 105
EbrA B. licheniformis YATWSGAGTVLTAIIGVLWFQEKLNRRNIAGIICLVSGVVLINLS..... 105
EbrB B. atrophaeus YATWSGVGTALTAIVGFLLFGETISLKGVFGLTLVIAGVVVLNQSKAPAKEKK.QTVCE.. 117
EbrB B. subtilis YATWSGVGTALTAIVGFLLFGETISLKGVFGLTLVIAGVVVLNQSKAHAEDKK.QTACE.. 117
EbrB B. licheniformis YATWAGTGTALTAAIGHFIFQEPFNLKTLIGLTLIIGGVFLLNSKRTEAADQKAQLTIEI. 119
QacC S. aureus YATWAGLGLVLTTVVSIIIFKEQINLITIVSIVLIIVGVVSLNIFGTSH..... 107
QacH S. saprophyticus YASWAGLGLVLTTIVSVLIFKEQINLISIISIILIIFGVVLLNTFGSSH..... 107
EBR E. coli YAVWSGLGVVIITAIAWLLHGQKLDAWGFVGMGLIIAAFLLARSPSWKSLRRPTPW..... 115
QacE E. aerogenes YAVWSGLGVVIITAIAWLLHGQKLDAWGFVGMGLIVSGVVVLNLLSKASAH..... 110
QacE E. coli YAVWSGLGVVIITAIAWLLHGQKLDAWGFVGMGLIVSGVVVLNLLSKASAH..... 110
QacF E. aerogenes YAVWAGLGIVLVAAIAWIFHGQKLDFWAFIGMGLIVSGVAVLNLLSKVSAH..... 110
YvaE B. subtilis YAVWSGMGIVLITVVGFLFFQEHVSVMKVISIGLIIAGVVSLNLIEHVAVSEPVHKSGQYK 119
Mmr M. tuberculosis YALWSAIGTAAIVLVAVLFLGSPISVMKVVGVGLIVVGVVTLNLAGAH..... 107
Mmr M. leprae YALWSAIGTAAIVLIAVLFLDSPVSVAKVVGVALIIVGVITLNLADAH..... 107
NepA A. nicotinovorans YALWAGLGTASVAVIGVLFRNERFSWKHAIGLALVVTGVVTLNLQAGQ..... 116
NepB A. nicotinovorans YATWSGLGTVAVAFAGAIIHGEAVTLGRITAITAVIGGIVILNLATTRQHSARRKDV.... 166
SugE S. typhimurium YAVWTGIGAVGAAITGILLLGESASPARLLSLGLIVAGIIGLKLSTH..... 105
SugE C. freundii YAVWTGIGAVGAAITGILLLGESASPARLLSLGLIVAGIIGLKLSTH..... 105
SugE S. choleraesuis YAVWTGIGAVGAAITGILLLGESASPARLLSLGLIVAGIIGLKLSTH..... 105
SugE E. coli YAVWTGIGAVGAAITGIVLLGESANPMRLASLALIVLGIIGLKLSTH..... 105
SugE Y. pestis YAVWTGIGAVGTAILGIVLLGESASLARILSLGLILAGIIGLKLAS..... 104
SugE P. vulgaris YAIWTGIGAVGTAIFGIIVFGESANIYRLLSLAMIVFGIIGLKLAS..... 104
SugE P. luminescens YAVWTGIGAVGTAIFGIIVFGESASFARILSFALIIAGIIGLKLSS..... 104
YkkC B. subtilis YAVFTGLGTAGTVLSEIVLFHEPVGWPKLLLIGVLLIGVIGLKLVTQDETEEKGGEA.... 112
YkkC B. licheniformis YAVFTGLGTAGTVVCEIALFNEPANIAKLALIGVLLCGVIGLKLVTN...EEKGEAS.... 109
YvdS B. subtilis YTVFVGIGTVGTYLTGIVLG.ESFSAAQMFFLALLAGILGMKLFTKESKSQPGGEK.... 111
YkkD B. subtilis YAVWTGIGTAGGALIGILFYKEQKDAKRIFFIALILCSAVGLKILS..... 105
YkkD B. licheniformis YAVWTGIGTVGGALVGILFYGEPKDGKRIFFIALILGSAVGLKLIS..... 104
YvdR B. subtilis YAVWTGIGSIGVSAVGLIWFKERFQLSQVISLCLILAGVIGLRLTSSS..... 106
consensus Y* * * G * * * * *
60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110

References

1. Chen YJ, et al. (2007) X-ray structure of EmrE supports dual topology model. *Proceedings of the National Academy of Sciences of the United States of America* 104(48):18999–9004.
2. Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J. Comp. Chem.* 26:1781–1802.
3. Martí-Renom MA, et al. (2000) Comparative Protein Structure Modeling of Genes and Genomes. *Annual Review of Biophysics and Biomolecular Structure* 29(1):291–325.
4. Eswar N, et al. (2007) Comparative Protein Structure Modeling Using MODELLER in *Current Protocols in Protein Science*. (John Wiley & Sons, Inc., Hoboken, NJ, USA).
5. Barrera FN, Fendos J, Engelman DM (2015) Membrane physical properties influence transmembrane helix formation. *Proc. Natl. Acad. Sci. USA* 109(36):14422–14427.
6. Popot JL, Engelman DM (2000) Helical Membrane Protein Folding, Stability, and Evolution. *Annual Review of Biochemistry* 69(1):881–922.
7. Shelar A, Bansal M (2014) Sequence and conformational preferences at termini of α-helices in membrane proteins: Role of the helix environment. *Proteins: Structure, Function, and Bioinformatics* 82(12):3420–3436.
8. Frishman D, Argos P (1995) Knowledge-based secondary structure assignment. *Proteins* 23:566–579.
9. Fleishman SJ, et al. (2006) Quasi-symmetry in the Cryo-EM Structure of EmrE Provides the Key to Modeling its Transmembrane Domain. *Journal of Molecular Biology* 364(1):54–67.
10. Stone JE, Gullingsrud J, Grayson P, Schulten K (2001) A system for interactive molecular dynamics simulation in *2001 ACM Symposium on Interactive 3D Graphics*, eds. Hughes JF, Séquin CH. (ACM SIGGRAPH, New York), pp. 191–194.
11. Ubarretxena-Belandia I, Baldwin JM, Schuldiner S, Tate CG (2003) Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *The EMBO Journal* 22(23):6175–6181.
12. Trabuco LG, Villa E, Mitra K, Frank J, Schulten K (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16:673–683.
13. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112:6127–6129.
14. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins: Struct., Func., Bioinf.* 55:383–394.
15. Tanner DE, Phillips JC, Schulten K (2012) GPU/CPU algorithm for generalized Born / solvent-accessible surface area implicit solvent calculations. *J. Chem. Theory Comput.* 8:2521–2530.
16. Trabuco LG, Villa E, Schreiner E, Harrison CB, Schulten K (2009) Molecular Dynamics Flexible Fitting: A practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* 49:174–180.
17. Humphrey W, Dalke A, Schulten K (1996) VMD – Visual Molecular Dynamics. *J. Mol. Graphics* 14(1):33–38.
18. MacCallum JL, Bennett WFD, Tieleman DP (2007) Partitioning of amino acid side chains into lipid bilayers: results from computer simulations and comparison to experiments. *J. Gen. Physiol.* 129:371–377.
19. Jo S, Kim T, Im W (2007) Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One* 2:e880.
20. Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comp. Chem.* 29:1859–1865.
21. Morrison EA, et al. (2012) Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature* 481(7379):45–50.
22. Gayen A, Banigan JR, Traaseth NJ (2013) Ligand-Induced Conformational Changes of the Multidrug Resistance Transporter EmrE Probed by Oriented Solid-State NMR Spectroscopy. *Angewandte Chemie International Edition* 52(39):10321–10324.
23. Dutta S, Morrison EA, Henzler-Wildman KA (2014) Blocking Dynamics of the SMR Transporter EmrE Impairs Efflux Activity. *Biophysical Journal* 107(3):613–620.
24. Cho MK, Gayen A, Banigan JR, Leninger M, Traaseth NJ (2014) Intrinsic Conformational Plasticity of Native EmrE Provides a Pathway for Multidrug Resistance. *Journal of the American Chemical Society* 136(22):8072–8080.
25. Gayen A, Leninger M, Traaseth NJ (2016) Protonation of a glutamate residue modulates the dynamics of the drug transporter EmrE. *Nature Chemical Biology* 12(3):141–145.
26. Oursel D, et al. (2007) Lipid composition of membranes ofEscherichia coli by liquid chromatography/tandem mass spectrometry using negative electrospray ionization. *Rapid Communications in Mass Spectrometry* 21(11):1721–1728.
27. Jorgensen W, Chandrasekhar J, Maudura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
28. Best RB, et al. (2012) Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.* 8(9):3257–3273.
29. Klauda JB, et al. (2010) Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J. Phys. Chem. B* 114(23):7830–7843.
30. Kubo R, Toda M, Hashitsume N (1991) *Statistical Physics II: Nonequilibrium statistical mechanics.* (Springer, New York).
31. Brünger AT, Brooks III CL, Karplus M (1984) Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.* 105(5):495–498.
32. Martyna GJ, Tobias DJ, Klein ML (1994) Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* 101(5):4177.

33. Feller SE, Zhang Y, Pastor RW, Brooks BR (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics* 103(11):4613–4621.
34. Fiorin G, Klein ML, Hénin J (2013) Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* 111(22–23):3345–3362.
35. Padariya M, Kalathiya U, Baginski M (2015) Structural and dynamic changes adopted by EmrE, multidrug transporter protein–Studies by molecular dynamics simulation. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1848(10):2065–2074.
36. Vanommeslaeghe K, et al. (2010) CHARMM General Force Field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comp. Chem.* 31(4):671–690.
37. Vanommeslaeghe K, MacKerell, Jr. AD (2012) Automation of the CHARMM General Force Field (CGenFF) I: Bond perception and atom typing. *J. Chem. Inf. Model.* 52(12):3144–3154.
38. Vanommeslaeghe K, Raman EP, MacKerell, Jr. AD (2012) Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* 52(12):3155–3168.
39. Mayne CG, Saam J, Schulten K, Tajkhorshid E, Gumbart JC (2013) Rapid parameterization of small molecules using the Force Field Toolkit. *J. Comp. Chem.* 34:2757–2770.
40. Martyna GJ, Tobias DJ, Klein ML (1994) Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101(5):4177–4189.
41. Feller SE, Zhang Y, Pastor RW, Brooks BR (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* 103(11):4613–4621.
42. Darden T, York D, Pedersen LG (1993) Particle mesh Ewald: An $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98(12):10089–10092.
43. Essmann U, et al. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
44. Miyamoto S, Kollman PA (1992) SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water molecules. *J. Comp. Chem.* 13(8):952–962.
45. Vermaas JV, Hardy DJ, Stone JE, Tajkhorshid E, Kohlmeyer A (2016) TopoGromacs: Automated topology conversion from CHARMM to GROMACS within VMD. *J. Chem. Inf. Model.* 56(6):1112–1116.
46. Abraham MJ, et al. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1-2:19–25.
47. Pronk S, et al. (2013) Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29(7):845–854.
48. van der Spoel D, et al. (2005) Gromacs: Fast, flexible, and free. *J. Comp. Chem.* 26:1701–1718.
49. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* 81(1):511–519.
50. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* 31:1695–1697.
51. Nose S (1990) Constant-temperature molecular dynamics. *Journal of Physics: Condensed Matter* 2(S):SA115–SA119.
52. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* 78:2690–2693.
53. Morrison EA, Robinson AE, Liu Y, Henzler-Wildman KA (2015) Asymmetric protonation of EmrE. *The Journal of General Physiology* 146(6):445–461.
54. Woods CJ, Essex JW, King MA (2003) The Development of Replica-Exchange-Based Free-Energy Methods. *Journal of Physical Chemistry B* 107(49):13703–13710.
55. Axelsen PH, Li D (1998) Improved convergence in dual-topology free energy calculations through use of harmonic restraints. *Journal of Computational Chemistry* 19(11):1278–1283.
56. Deng Y, Roux B (2009) Computations of standard binding free energies with molecular dynamics simulations. *J. Phys. Chem. B* 113:2234–2246.
57. Radak BK, et al. (2017) Constant-pH Molecular Dynamics Simulations for Large Biomolecular Systems. *Journal of Chemical Theory and Computation* 13(12):5933–5944.
58. Van Der Walt S, Colbert SC, Varoquaux G (2011) The NumPy array: A structure for efficient numerical computation. *Computing in Science and Engineering* 13(2):22–30.
59. Hagberg A, Swart P, Chult D (2008) Exploring network structure, dynamics, and function using NetworkX in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, eds. Varoquaux G, Travis V, Millman J. (Pasedena, CA USA), No. SciPy, pp. 11–15.
60. Hunter JD (2007) Matplotlib: A 2d graphics environment. *Comput. in Sci. and Eng.* 9(3):90–95.
61. Vermaas JV, Tajkhorshid E (2014) Conformational heterogeneity of α-synuclein in membrane. *Biochim. Biophys. Acta Biomembr.* 1838(12):3107–3117.
62. Vermaas JV, et al. (2015) Mechanism of lignin inhibition of enzymatic biomass deconstruction. *Biotechnology for Biofuels* 8(1):217.
63. Dastvan R, Fischer AW, Mishra S, Meiler J, Mchaourab HS (2016) Protonation-dependent conformational dynamics of the multidrug transporter EmrE. *Proceedings of the National Academy of Sciences* 113(5):1220–1225.