The C-terminus of the multi-drug efflux pump EmrE prevents proton leak by gating transport.

Merissa Brousseau[1‡], Da Teng[4‡], Nathan E. Thomas[1,2], Gregory A. Voth[4] and Katherine A. Henzler-Wildman [1,3*]

[1]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53703 USA
[2]*Current Address:* LabCorp, Madison, WI 53704
[3]Nuclear Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, Madison WI 53703, USA
[4] Department of Chemistry, Chicago Center for Theoretical Chemistry, Institute for Biophysical Dynamics, and The James Franck Institute, The University of Chicago, Chicago, Illinois, 60637
‡ Equal contribution authors

***Correspondence: henzlerwildm@wisc.edu**

Author Contributions: Conceptualization, N.E.T, M.B., and K.A.H.W; Methodology, N.E.T., M.B., D.T., G.V. and K.H.W,.; Formal Analysis, N.E.T, M.B., D.T., G.V. and K.A.H.W; Validation, M.B., N.E.T, and D.T.; Investigation, N.E.T, M.B., and D.T. ; Writing – Original Draft, M.B., D.T., G.V. and K.A.H.W.; Writing – Review and editing –M.B., N.E.T, D.T., G.V. and K.A.H.W.; Visualization, K.A.H.W., M.B., D.T. and N.E.T.; Supervision, N.E.T., G.V. and K.A.H.W., Funding Acquisition, K.A.H.W and G.V. All authors have read and agreed to the published version of the manuscript.

This PDF file includes:

Main Text
Figures 1 to 8

Abstract

The model multi-drug efflux pump from *Escherichia coli*, EmrE, can perform multiple types of transport leading to different biological outcomes, conferring resistance to some drug substrates and enhancing susceptibility to others. While transporters have traditionally been classified as antiporters, symporters, or uniporters, there is growing recognition that some transporters may exhibit mixed modalities. This raises new questions about the regulation and mechanisms of these transporters. Here we show that the C-terminal tail of EmrE acts as a secondary gate, preventing proton leak in the absence of drug. Substrate binding unlocks this gate, allowing transport to proceed. Truncation of the C-terminal tail (Δ107-EmrE) leads to altered pH regulation of alternating access, an important kinetic step in the transport cycle, as measured by NMR. Δ107-EmrE has increased proton leak in proteoliposome assays and bacteria expressing this mutant have reduced growth. MD simulations of Δ107-EmrE show formation of a water wire from the open face of the transporter to the primary binding site in the core, facilitating proton leak. In WT-EmrE, the C-terminal tail forms specific interactions that block formation of the water wire. Together these data strongly support the C-terminus of EmrE acting as a secondary gate that regulates access to the primary binding site in the core of the transporter.

Main Text

Introduction

Antibiotic resistance may be mediated by several mechanisms, including active export of drugs by promiscuous multidrug efflux pumps (1). Among these transporters, the small multi-drug resistance (SMR) family have been found throughout the bacterial kingdom and exhibit particularly promiscuous substrate profiles (2–4). The most well-studied SMR transporter is the *Escherichia coli* protein EmrE, which confers

resistance to a broad array of toxic polyaromatic cations and quaternary ammonium compounds through secondary active transport (5). These transporters couple the energetically favorable import of protons down the proton motive force (PMF) to drive active export of antibiotics and antiseptics (6, 7). As the archetype for the family of the smallest ion-coupled transporters, EmrE has become a model system for studying the molecular mechanism of proton-coupled transport.

EmrE transport is electrogenic for tetraphenyl-phosphonium (TPP^+) and electroneutral for methyl viologen (MV^{2+}) (8), consistent with a $2H^+$:1 drug antiport stoichiometry (5). Early mechanistic models focused on the minimal set of states and transitions necessary for such stoichiometric antiport. More recently, NMR studies of EmrE protonation and alternating access showed that many more states and transitions have populations and rates that are not insignificant at near-physiological pH and temperature (9). Inclusion of these states and transitions in the mechanistic model provides pathways that allow for alternative transport activity, including symport, drug-uniport, and proton-uniport (leak) (Fig. 1). In this model, different environmental conditions (pH) or small molecule substrates that shift the relative rates of different microscopic steps can alter the dominant transport behavior (9, 10). This has been confirmed experimentally: a small molecule substrate, harmane, triggers uncontrolled proton leak through EmrE to an extent that is detrimental to *E. coli* growth and NADH production (11). However, the question of how this transporter avoids catastrophic leak remains unanswered.

Direct measurements of proton release upon drug binding showed that drug-induced deprotonation occurs at the C-terminal histidine (H110) in addition to the essential glutamate-14 residues that defines the primary binding site for drug and proton (12). Additionally, the C-terminus on one protomer in the homodimer is highly sensitive to the identity of drug bound in the primary site (13). Early solid-state ^{31}P NMR experiments suggested a second, lower affinity TPP^+ binding site near the acidic loop residues E25 and D84 (14), which are likely to be in close spatial proximity to the C-terminal tail in this small transporter. Together, these data led us to propose a secondary gating model where the C-terminal tail prevents proton release until drug binding at a peripheral site on the transporter surface displaces the tail (12)

Unfortunately, none of the available EmrE structures provide high resolution data on the conformation of the C-terminal tail and adjacent loop regions. Early cryo-electron microscopy and crystal structures revealed the unique asymmetric arrangement of the transmembrane helices and antiparallel topology of the EmrE homodimer, but had very low resolution and limited density in the loops and tails (15, 16). Recent, higher-resolution crystal structures and NMR structures have provided more precision on substrate binding within the transport pore (17–19). However, the crystal structures used a monobody that required mutation of three residues in the TM1-TM2 loop (E25N, W31I, V34M), including E25, and there is limited or missing density for other loops on the open face of the transporter and the C-terminal tail after residue 104. In the NMR structures, distance restraints are primarily substrate-protein distances within the transmembrane helices lining the primary binding site and only chemical-shift-derived backbone torsion angles restrain the loops and tail (18, 19). The most recent NMR structures used a loop mutant, L51I, that disrupts the gating mechanism, locking the transporter open, and again had limited restraints in the loops and C-terminal tail (20). Thus, there is limited structural data for the C-terminal tail and loops although these regions are functionally important in gating access to the central binding site defined by residue E14. In such cases, molecular dynamics (MD) has proven to be an excellent tool, and the only atomic resolution model of the C-terminal tail is from MD simulations (21).

Here we use NMR, *in vitro* and *in vivo* biochemical assays, and MD simulations to characterize a C-terminal deletion mutant of EmrE truncated after residue 106, denoted Δ107-EmrE, to directly determine the regulatory role of the C-terminal tail. Comparisons of growth and resistance phenotypes, alternating access rates, and transport activities of Δ107- and WT-EmrE confirm the importance of the C-terminal tail in regulating tightly coupled antiport and minimizing proton leak. Simulations on these two systems also suggest differences in water structure and hydrogen bonding patterns when the C-terminus is truncated. Examination of interactions with the newly-discovered substrate harmane, which triggers uncoupled proton leak as the dominant transport mode (11), suggests interactions between the tail and a secondary site on the protein may allow for allosteric regulation of gating, reconciling the free exchange model with minimal

leak by the WT transporter in the absence of small molecule substrates. Further, MD simulations provided a possible secondary site and the structural basis for this regulation.

Results

EmrE is properly folded and functional when the C-terminal tail is truncated

We first assessed whether C-terminal tail truncation affected the ability of EmrE to confer resistance to toxic substrates, the well-established primary function of this transporter. Growth assays of MG1655-*ΔemrE E. coli* cells expressing WT-, Δ107- or E14Q-EmrE show that all strains grow well in the absence of toxic compounds (Fig. 2B). In these assays, uninduced leaky expression from a low-copy number plasmid (p15 origin) with a pTrc promoter keeps transporter expression relatively low (11). In the presence of ethidium bromide, a substrate commonly used to assess the activity of EmrE and other multidrug efflux pumps, functional transporter is required for survival (Fig. 2C). This confirms the known resistance activity of WT-EmrE, that E14Q-EmrE is non-functional, and that Δ107-EmrE is properly expressed to the inner membrane, folded and functionally able to confer resistance to toxic compounds in a manner comparable to the WT transporter.

Truncation of the C-terminal tail enhances proton leak through EmrE

There is a small but reproducible growth defect for cells expressing Δ107-EmrE in the absence of exogenous substrate (Fig. 2A, B). This defect becomes apparent around 5 hours, the point at which available fermentable sugars in LB media are depleted, increasing dependence on the proton motive force (PMF) for energy production (22). A similar time dependent growth inhibition is observed for WT-EmrE in the presence of harmane, and we have previously shown that this substrate triggers uncoupled proton leak through EmrE (11) Thus, while Δ107-EmrE competently performs the proton-coupled drug antiport necessary to confer resistance to toxic substrates, it is detrimental to *E. coli* in the absence of known small molecule substrate in a manner suggestive of proton leak.

To directly test this hypothesis, we measured proton leak in proteoliposomes. The pH-sensitive dye pyranine was encapsulated inside proteoliposomes at pH 6.5, and the liposomes were then diluted 100-fold into pH 7.5 buffer. If protons leak out of the liposome (down the proton concentration gradient), the internal pH will rise and pyranine fluorescence will increase. We compared the fluorescence, normalized to time zero, of proteoliposomes with WT, Δ107, or E14Q-EmrE. WT-EmrE proteoliposomes show a gradual increase in internal pH over time (Fig. 3A, solid black), which is faster than the pH change for E14Q-EmrE proteoliposomes (Fig. 3A, solid gray). This is consistent with a small amount of proton leak through WT-EmrE and a role for E14 in mediating leak. Δ107-EmrE proteoliposomes show a much faster rise in internal pH in this assay (Fig. 3A, solid red). Repeating the experiment with the protonophore CCCP in the external buffer results in rapid proton leak for all proteoliposome samples (Fig. 3A, dotted lines), and the results match the timescale and amplitude of proton leak observed for the Δ107-EmrE proteoliposomes. Thus, C-terminal tail truncation causes rapid proton leak through EmrE.

Initiation of the assay by dilution complicates measurement of the fluorescence baseline, obscuring rapid changes in the few seconds between dilution and initial fluorescence read. We repeated the assay with side-by-side dilution of proteoliposomes (internal pH 6.5) into pH 6.5 buffer (baseline), or pH 7.5 buffer (transport). This baseline normalization reveals a rapid internal pH change (Fig. 3B). The empty liposome control is flat, showing that the liposomes do not leak without EmrE. However, the signal is non-zero (Fig. 3B, black arrow), likely due to residual exterior pyranine. The time 0 fluorescence of WT-, E14Q- and Δ107-EmrE proteoliposomes are much higher (Fig. 3B, blue arrow), indicating an additional rapid change in the internal pH in the presence of protein. While the encapsulated pyranine is protected from the direct impact of external pH change, proton transport (leak) through EmrE will change internal pH. The transmembrane pH gradient may also affect EmrE itself, altering the pKa of key residues (E14, H110) and causing rapid release of protons inside the liposome, or cause a rapid burst-phase of leak as the transporter transitions from a symmetric-pH conformation to an asymmetric-pH conformation.

To further assess the initial rapid pH change, we used solid supported membrane electrophysiology (SSME) to measure ΔpH-driven current in proteoliposomes since this technique provides a continuous readout as a gradient is applied. Many proteoliposomes can be adsorbed onto the gold-coated sensor, enabling highly sensitive detection of electrogenic transport. The same lipid to protein ratio was used as in the pyranine assay, and reported values are an average of 3 independently prepared sensors per mutant

to account for variability in liposome adsorption onto sensors. The liposomes are first equilibrated with external buffer identical to the interior, and then a different external buffer is rapidly washed over the liposomes to create a transmembrane pH gradient while recording is in progress. The capacitive current is measured (Fig. 3C,Fig. S1 left, Fig. S2 A,B) and integrated to yield the total transported charge in response to the applied gradients (Fig. 3D, Fig. S1 right, Fig. S2 C). In the absence of drug, empty liposomes have minimal charge movement as expected for minimal proton leak. However, both WT and Δ107-EmrE have measurable current in the presence of a pH gradient, with increased net charge movement at higher absolute pH and Δ107-EmrE has consistently higher leak under all pH conditions (Fig. S1). The consistency of the SSME and pyranine assay results establishes the validity of the assay for comparing WT- and Δ107-EmrE proton leak, and the ability to perform multiple assays with the same proteoliposome sensors to compare flux at different absolute pH or gradient magnitudes (Fig. S1, S2).

In symmetric pH environments, the only amino acid side chains with pKa values near neutral pH are E14 (pKa 6.8 ± 0.1 and 8.5 ± 0.2 at 25 °C) and H110 (6.98 ± 0.01, 7.05 ± 0.02). Δ107-EmrE is lacking Histidine (H110), so any proton binding/release from H110 that contribute to the capacitive current will be absent in Δ107-EmrE, but the net charge transport (Fig. 3D, Fig. S1 right, Fig. S2 C,D) is greater for Δ107-EmrE than for WT-EmrE. This rules out a simple model where proton binding/release from H110 accounts for the fast proton flux. Any conformational change involving the C-terminal tail that contributes to the capacitive current should also decrease as average pH increases and H110 protonation and net charge decrease, however the opposite pH dependence is observed for WT-EmrE, indicating that this is unlikely to be a major contributor to the SSME current. A model where movement of the C-terminal tail regulates access to E14 and C-terminal truncation alters this gating process would be consistent with the observed currents and their pH dependence (Fig. 3C-E, Fig. S1, Fig. S2), as well as the longer-timescale change in intra-liposome pH (Fig. 3B). We note that the one second SSME traces do not reach steady state (Fig. 3D, Fig. S1 right, Fig. S2 C), but do provide insight into the dead time of the pyranine assay (Fig. 3B) and match the relative magnitude of the observed burst phase. Altogether, this data supports a role for the C-terminal tail as a secondary gate that minimizes proton leak through WT-EmrE in the absence of substrate.

The pH-dependent rate of alternating access in Δ107-EmrE is distinct from WT-EmrE

We next used solution NMR to assess the impact of C-terminal tail truncation on the structure and dynamics of EmrE, since this will impact gating and transport. Due to the asymmetric structure of EmrE, the two subunits have unique chemical shifts. As EmrE undergoes alternating access, the two subunits swap conformations resulting in exchange between AB and BA dimer topology (23). The rate of the alternating access exchange process affects the NMR line shape, resulting in distinct sets of peaks for each subunit when exchange is slow, line broadening as the rate increases, and eventually coalescence into a single set of peaks at the average chemical shift when exchange is fast. Thus, the appearance of simple 2D ^{1}H-^{15}N TROSY-HSQC spectra can provide significant insight into both the structure and dynamics of the transporter under different conditions.

In the absence of drug, Δ107-EmrE is in fast-intermediate exchange at both pH values, with only one set of peaks and significant line broadening (Fig. 4A). However, the spectra are distinct, with slightly more line broadening at high pH. This indicates that protonation of E14 still affects the overall structure of Δ107-EmrE and alternating access is slightly slower at high pH. WT-EmrE has similar fast exchange behavior at low pH, but increasing pH results in a significantly slower rate of alternating access (Fig. S3) (23).

Upon addition of a tight binding drug-substrate, tetraphenylphosphonium (TPP^+) to Δ107-EmrE at low pH, two distinct peaks become visible for each residue (Fig. 4B, red). This confirms that the poor spectral quality of the substrate-free spectrum was due to protein motion and not degradation or aggregation, and that alternating access is significantly slower with TPP^+ bound. The similarity of this spectrum of Δ107-EmrE bound to TPP^+ at low pH with the spectrum of WT-EmrE under the same conditions (low pH, TPP^+ bound) also provides additional evidence that the general structure of Δ107-EmrE remains intact and the binding site has undergone minimal perturbation upon truncation of the last 4 amino acids (Fig. S3A).

At high pH, TPP^+-bound Δ107-EmrE shows significant line broadening and partial coalescence of the two distinct sets of peaks in the NMR spectrum, indicating that the rate of alternating access is faster (Fig. 4B, blue). We quantitatively measured the rate of alternating access as a function of pH for TPP^+-bound Δ107-EmrE with ^{1}H-^{15}N TROSY-ZZ-exchange NMR experiments (24) . In this experiment, a delay

is inserted in the pulse sequence between recording the ^{15}N and ^{1}H chemical shifts, such that a conformational exchange during this delay will result in the appearance of cross-peaks with the ^{15}N chemical shift of the original state and ^{1}H chemical shift of the final state (Fig. 4C, D). By comparing the intensity of these cross-peaks relative to the auto-peaks as a function of the delay time, we can determine the rate of alternating access (25) (Fig. 4E, S4). The alternating access rate for TPP^+-bound Δ107-EmrE is $4 \pm 1\ s^{-1}$ at low pH, and $17 \pm 3\ s^{-1}$ at high pH. There is greater scatter in the peak intensity ratio at high pH due to enhanced exchange with water for residues on the open face of the transporter, which reduces the peak intensity. However, there is no overlap between low pH and high pH, clearly demonstrating a significant change in alternating access rate for ∆107-EmrE with pH. TPP^+-bound WT EmrE has the same rate of alternating access as ∆107-EmrE at low pH, but does not vary significantly with pH (13). Thus, truncation of the C-terminal tail alters the pH-dependence of alternating access for ∆107-EmrE in both the absence and presence of drug-substrates, supporting a role for this region in regulating the pH-dependent conformational dynamics of EmrE.

The C-terminus controls a water wire into the primary binding site

To further investigate how the C-terminal tail of EmrE may interact with other regions of EmrE and gate access into the transport pore, we carried out MD simulations of substrate-free WT-EmrE and ∆107-EmrE in a DMPC lipid bilayer. The protonation states were set to simulate a pH between 7.0 and 8.0, where only $E14^A$ with the higher pKa is protonated. We used an NMR structure determined with TPP^+ (PDB: 7JK8) as the initial structural model (26). Since this NMR structure does not include the C-terminal tail, we modeled it with CHARMM-GUI (27). The systems were first equilibrated at constant temperature 310 K and 1 bar pressure for 400 ns, while position restraints were gradually released. After this equilibration, the RMSD of the protein compared with the initial structure plateaued. The production simulations were run for another 1000 ns and MD trajectories were output every 0.1 ns. With the same starting structure, we ran three parallel replicas to ensure consistency.

A critical prerequisite of proton transport is a water wire, either transient or long-lasting, that allows protons to transport through the Grotthuss hopping mechanism (28). In the ∆107-EmrE MD simulations, we identified a water chain (aka "water wire") not seen in WT simulations that connects E14 at the primary binding site to bulk water. It enters the protein from the open side near $R106^A$, passing through the triad of $A61^A$, $I68^B$, and $I71^B$, and then comes into the primary binding site (Fig. 5C). To quantitatively understand the connectivity of this water wire over time, we calculated the length for the shortest water path *S* for each frame in the trajectory with graph theory. This length is defined such that smaller values for *S* reflect better connectivity of the water molecules, as described in more detail in the methods section and (29). The logarithm of the shortest path, log(*S*), is plotted for all simulation systems (Fig. 5A). For ∆107-EmrE, there is a leak state characterized by a much smaller log(*S*) where the water wire is very well connected, while for WT-EmrE log(*S*) is consistently large. The existence of this water chain in ∆107-EmrE is consistent with the enhanced proton leak observed experimentally. This newly found water wire starts very close to the C-terminus and is distinct from the ligand entry path (30).

Structural basis for the C-terminus gating

In the initial structure of WT-EmrE, the C-terminus is floating in bulk water. After equilibration, we observed the tail coming closer and interacting with the protein in all three replicas. The tail has two notable interactions with other parts of the protein; the first of which is a salt bridge between D84 and R106. The second involves the carbonyl group of the C-terminus, which forms hydrogen bonds with T56, and occasionally forms a salt bridge with K22 (Fig. S5). When these interactions occur, the tail moves near the water chain. Examining characteristic snapshots from the simulation trajectories illustrates the effect of the C-terminal tail on water wire formation. Figure 5B shows the worst-hydrated snapshot of WT-EmrE determined by the highest *S*. Figure 5C shows the best-hydrated snapshot of ∆107-EmrE determined by the lowest *S*. Despite the overall dryness in the channel of WT-EmrE, there were a few rare moments where a transient water wire formed, characterized by a sudden drop in water path length, and Figure 5D shows the best-hydrated snapshot from that simulation. In WT-EmrE, the water wire is broken at the triad of $A61^A$, $I68^B$, and $I71^B$ (Fig. 5B, 6D), suggesting that these three hydrophobic residues may act as a bottleneck for the water wire.

To test this hypothesis with statistically meaningful results, we calculated the distances between the closest pairs of side-chain hydrogens among these three residues for the whole trajectory. For WT-EmrE, the minimum hydrogen-hydrogen distances of $A61^A$-$I68^B$ and $A61^A$-$I71^B$ are 2.9 ± 0.6 Å and 2.8 ±

0.6 Å respectively, and for Δ107-EmrE, these distances increased to 5.9 ± 1.0 Å and 5.4 ± 1.2 Å (Fig. 6A). The $I68^B$-$I71^B$ distance does not differ significantly. These increased distances suggest that $A61^A$ is moving away in the Δ107-variant, opening a pore for the water wire to form. For comparison, the diameter of a water molecule is measured to be around 2.7 Å(31). This means a triangle larger than 5.4 Å in size is required for a water molecule to fit inside. Additionally, the most hydrated snapshot in WT-EmrE showed $A61^A$ takes a conformation more similar to Δ107-EmrE and very different from dry WT-EmrE. This confirms the role of $A61^A$ rotation in controlling the water wire formation.

The potential of mean force of proton transport supports a hydrophobic bottleneck

The potential of mean force (PMF) for explicit proton transport (see Methods) shows the free energy change of the system as a function of a particular collective variable (CV) or reaction coordinate. It can provide additional information beyond structural snapshots of a reaction, which are only incomplete samples of the ensemble. For example, the existence of a water chain does not necessarily mean good proton conductance, but by contrast, the PMF for explicit proton transport (including Grotthuss proton shuttling) can provide key thermodynamic and kinetic information about the transport process (29, 32). However, free energy sampling involving proton transport is also intrinsically complicated. First, proton transport involves chemical bond breaking and formation is beyond the capability of classical MD, so expensive quantum mechanics/molecular mechanics (QM/MM) may be required. Second, when an excess proton is solvated in the water, the net positive excess charge defect arising from the presence of the excess proton can be delocalized in the water network. It is not possible to define which proton is exactly the “excess proton” as Grotthuss shuttling dynamically rearranges these definitions. To address these issues, we have developed a method called Multiscale Reactive MD (MS-RMD) (33). It can model bond forming and breaking involving excess proton shuttling at a computational cost near classical MD. In this approach, one can also conveniently define a “center of excess charge” (CEC) to describe the location of the excess positive charge defect.

We carried out umbrella sampling with MS-RMD that describes the proton transport from $E14^B$ to the bulk water in WT-EmrE. The CV “x” is defined similarly as in reference (34) as the distance between the glutamate oxygen to the CEC, mapped along a vector that aligns with the proton transport direction (see Methods). The resulting PMF (Fig. 6B) indicates a deep well near x = 0 Å, where the proton is on the glutamate, and a transition state near *x* = 10.0 Å. A conformational snapshot from the transition state shows the CEC (Fig. 6C, orange sphere) is in close proximity to the pore defined by I68 and A61 (Fig. 6D). This result supports the hypothesis that the bottleneck of the proton transport is this hydrophobic gate.

Re-assessing protonation state by NMR

Prior NMR pH titrations of WT-EmrE in the absence of drug-substrate revealed that the two essential E14 residues in the asymmetric homodimer have distinct pKa values (23), reflecting their unique structural environments. Upon binding TPP^+, one of the E14 residues is protected from protonation and no longer titrates, while the other ($E14^A$) retains a pKa of 6.8 ± 0.1, similar to the drug-free state (23). The only other titratable residue previously identified in WT-EmrE is the C-terminal histidine (12), which also has a pKa near neutral pH in WT-EmrE. Since H110 was removed by truncation in Δ107-EmrE, we expected only one protonation event in NMR pH titrations of TPP^+-bound Δ107-EmrE (Fig. 7A), corresponding to the one E14 residue that remains titratable when substrate is bound. A single protonation event would normally result in linear change in peak position from the chemical shift of the protonated state to the chemical shift of the deprotonated state over the course of the titration. This is because proton on-/off- is almost always in the fast-exchange limit for NMR, resulting in observation of a single peak at the population-weighted average chemical shift at each titration point. However, several peaks exhibit distinctly curved titration paths with transitions in two different pH ranges (Fig. S6) for TPP^+-bound Δ107-EmrE. Plotting the chemical shift of well-resolved peaks as a function of pH yields titration curves that can be fit using standard pKa equations. In this case, the data is well fit with a global 2 pKa model yielding apparent pKa values of 5.6 ± 0.2 and 7.1± 0.2. The higher of these two pKa values is close to the $E14^A$ pKa in TPP^+-bound WT-EmrE and the residues most sensitive to this protonation event are found lining the transport pore near $E14^A$, supporting assignment of this pKa to $E14^A$ (Fig. 7D).

This leaves the lower pKa unaccounted for. Possibilities include other acidic residues in the loops, E25 or D84, or the C-terminal carboxylate itself. The residues most sensitive to this lower pKa include R106 and the tail of subunit A, the TM1-TM2 loop of subunit B and TM3-TM4 loop of monomer B, all of which are

on the same "open" face of EmrE (Fig. 7C). Residues E25 and D84 are located in these loops, have previously been suggested to be part of a secondary binding site for substrates like TPP^+ (14). There is no evidence that these residues titrate in this pH range in WT-EmrE or in other mutants for which we have carried out NMR pH titrations (9, 23, 35), making it unlikely that E25 or D84 titrate in this pH range in the full-length transporter. However, truncation of the C-terminus in Δ107-EmrE could alter the structure, environment and pKa of these residues. Indeed, the hydrogen bond between R106 and D84 observed in the WT-EmrE simulations (Fig. 5B, D) is broken when the tail is truncated in Δ107-EmrE (Fig. 5C), and D84 and R106 have some of the largest pH-dependent chemical shift changes. There are relatively few experimental reports of the pKa of the terminal carboxylate in proteins, but it has been reported to have a pKa as high as 5.9 for the partially buried C-terminus of subunit *c* of F_1F_0 ATP Synthase (36). The MD simulations show the C-terminal carboxylate in WT-EmrE hydrogen bonds with T56 (TM 2-3 loop) and K22 (TM 1-2 loop) on the open face of the transporter (Fig. 5B, D), but the C-terminus in Δ107-EmrE no longer interacts with these residues (Fig. 5C). Examining the residues that sense this lower pKa shows that the TM3-TM4 and TM1-2 loops on the open face have larger chemical shift changes associated with the low pKa protonation event than those loops on the closed face of EmrE, while the C-terminal tail residues in both subunits detect the lower pKa. Comparison with the MD simulations shows that the residues involved in the hydrogen bond networks anchoring the C-terminal tail over the pore align well with the full list of residues that are sensitive to the lower pKa in the NMR titrations, including $T56_A$ (TM2-TM3 loop on the "open" face), V15-G17, I37, Y40, V69, and S72-L73. Thus, although this lower pKa is likely an artifact of tail truncation, this data experimentally supports the importance of the interactions between the tail and the rest of EmrE identified in the MD simulations as important for disrupting the water wire and occluding the E14 binding site.

Proton leak through Δ107-EmrE does not synergize with harmane

The well-established function of EmrE is proton-coupled antiport of toxic substrates, leading to toxin efflux and drug resistance. Recently, we discovered that some substrates, such as harmane, instead trigger uncoupled proton uniport, leading to ΔpH dissipation and defects in NADH production and growth in *E. coli* (11), essentially causing susceptibility rather than resistance. We suspected the enhanced proton leak observed through Δ107-EmrE and this harmane-triggered proton leak might have common elements in their underlying mechanism. Using SSME, we first compared the inherent proton leak through WT- and Δ107-EmrE in the absence of substrate. For the same ΔpH driving force, Δ107-EmrE has ≈3-fold greater proton leak than WT-EmrE (Fig. 8A, D, S7, S8). However, upon addition of 16 μM harmane, there is a large increase in proton leak through WT-EmrE and a small increase in leak through Δ107-EmrE, such that this substrate triggers identical total leak through either transporter (Fig. 8B,D, S7, S8). This SSME-detected proton leak increases with harmane concentration and is saturable in both WT- and Δ107-EmrE (Fig. 8C, S7, S8). If harmane acts as an allosteric regulator of the transporter that can unlock the secondary gate, then a saturating amount of harmane will result in the maximal signal for the WT transporter as observed. This was also confirmed in a pH-detected liposomal assay where addition of harmane decreases the magnitude of the pH change upon addition of CCCP to WT-EmrE containing proteoliposomes relative to empty liposomes (Fig. S9). In Δ107-EmrE, C-terminal truncation removes the majority of the secondary gate and key residues in the allosteric site, rendering proton leak comparably independent to harmane (Fig. 8, S7, S8).

To test this theory in the native organism, we conducted *in vivo* growth assays with WT- and Δ107-EmrE in the presence of harmane. MG1655-ΔemrE *E. coli* cells constitutively expressing WT-, E14Q-, and Δ107-EmrE from a plasmid were grown in the presence of 25 μM harmane. The cells expressing E14Q grew equally well in the presence or absence of harmane as the mutation of the primary binding site prevents proton binding in the transport pore and abolishes any proton leak (Fig. 8E). In the presence of harmane the difference in growth between Δ107- and WT-EmrE is eliminated, with significant growth defect for both constructs relative to E14Q-EmrE (Fig. 8E-F, Fig. S10). This *in vivo* data exactly matches the *in vitro* SSME and pyranine transport assays, demonstrating the importance of the C-terminal tail in gating and allosteric regulation of EmrE *in vivo*.

Discussion

There is a growing appreciation that uniport, symport and antiport simply represent extremes of a unified transport model that includes all possible binding and conformational states and their transitions

(37). Despite expectation that EmrE would have a simple mechanism and clearly illuminate the minimal requirements for coupled transport (4), it has proven to be surprisingly complex, exposing unexpected features of membrane protein topology and transport mechanism. The free exchange model, an extension of a universal 8-state transport model to include the ability of EmrE to bind two protons at the two E14 residues in the core of the homodimer (9), accounts for most of the available data. It includes all states and transitions observed by NMR and can account for the ability of EmrE to confer resistance to some substrates and susceptibility to other substrates (11). However, this model predicts rapid proton leak through WT-EmrE, while experimental data shows a small proton leak of smaller magnitude and similar timescale to coupled transport. In combination with prior data noting the importance of the C-terminal tail (12), the experimental data and MD simulations presented here support a regulatory role of the C-terminal tail as part of a secondary gate that minimizes proton leak in the absence of substrate and can be opened by binding of a drug-substrate.

Prior NMR data (11, 12, 14) led to the hypothesis that the C-terminal tail acts as a secondary gate occluding the primary E14-defined binding site in the absence of drug-substrate, with drug binding to a peripheral site opening this gate and allowing release of protons from E14. This model can explain the observed coupling of the C-terminal tail with both drug-binding and protonation events at the primary site (12) and the correspondence of proton off-rate and substrate on-rate in prior stopped-flow studies of EmrE (38). Here we combine MD simulations with experimental studies of a tail-truncated mutant, Δ107-EmrE, to more directly test the tail-gating hypothesis and determine whether this model can explain the minimal proton leak observed for WT-EmrE (9) and the newly-discovered harmane-gated proton uniport activity of EmrE (11).

If the tail is important for gating proton access to the binding pocket and preventing proton leak through the WT transporter, then truncation should enhance proton leak through EmrE. This is exactly what we observe, with increased uncoupled proton flux through Δ107-EmrE *in vitro* (Fig. 3) and diminished growth of *E. coli* expressing Δ107-EmrE *in vivo* (Fig 2). Comparing MD simulations of WT and Δ107-EmrE shows that the C-terminus can interact with TM3 to block formation of a water wire, providing a structural hypothesis for how the tail gates access to the primary EmrE binding site at E14 and regulates proton entry and exit from that site, as required for proton leak. Truncation of the C-terminal tail in Δ107-EmrE also removes key residues that are part of a secondary substrate binding site, reduces the sensitivity to harmane-triggered proton leak *in vitro* and *in vivo* (Fig. 8). Identical maximal harmane-triggered proton leak through WT- and Δ107-EmrE further supports the model that substrates bind at a secondary site in the vicinity of the C-terminal tail and releasing this secondary gate to allow proton flux.

The residues identified as important for regulating formation of the water wire, A61, I68, and I71 are all highly conserved. An analysis of 369 EmrE-related SMR sequences (39) shows A61 is fully conserved, while I68 and I71 are highly conserved with valine as the only substitution. D84 is the only fully conserved charged residue other than E14. A more recent analysis of SMR genes within the Joint Genome Institute's Genomic Encyclopedia of Bacteria and Archaea shows A61 and K22 are highly conserved across the SMR family, while I68, I71, T56 and D84 are conserved within the Qac subfamily (40). A61 and K22 are nearly as well conserved as the GXG motif in TM3 known to act as a fulcrum for conformational exchange between open-in and open-out conformations or the G97 in TM4 that is important for dimerization. A61C is not reactive with NEM (I68 and I71 not tested) (41), consistent with the closed hydrophobic gate observed in the MD simulations. Although drug binding and transport do not report on hydrophobic gating as directly, A61C has impaired resistance to acriflavine and methyl viologen (41) while A61L has impaired growth on ethidium (35). I68W, I68C and I71W impair growth on ethidium; I68A and I71G impair resistance to methyl viologen; and I61C, I68W, I68C, I71W, I71C have impaired TPP^+ binding (35, 42, 43). In addition, K22C, T56C and D84C reduce TPP binding (43); K22C reduces ethidium resistance (43); and D84C shows reduced resistance to ethidium and methyl viologen (44). In methyl viologen uptake assays, substitution of like charge at E25D and R82K resulted in transport comparable to WT, while K22R, D84E, and R106K had impaired uptake indicating a more specific requirement for these positions (44). Chemical shift perturbations upon harmane binding also highlight D84 and R106 (11). The secondary gating model and MD simulations presented here provide a rational for the functional significance of these residues observed in the prior work.

Active transport requires that a transporter is only ever open to one side of the membrane. This is generally thought to require formation of an occluded state where the substrate binding site is closed off from both sides of the membrane as the transporter transitions from the conformation open to one side of the membrane to the conformation open to the other side in order to avoid even transient formation of a

channel. Often a single gate is thought to control access to the transport pore, but sometimes multiple gates regulate a more complex transport cycle (45, 46). This is clearly seen in elevator mechanism transporters such as Glt_{Ph}. A mobile core domain contains the substrate binding site and moves up and down relative to the more rigid scaffold domain, effectively transitioning between inward-occluded and outward-occluded conformations. From either of these endpoint occluded states, a small hairpin domain can open to expose the binding pocket for substrate entry or exit. This hairpin gate must close to allow the sliding elevator movement and subsequent gate opening on the other side of the membrane (47). Studies of Glt_{Ph} have highlighted the evolutionary benefit of a kinetically controlled transport mechanism and the role of allosteric regulation in opening and closing the gate (48, 49). UapA, the xanthine-uric acid/H+ symporter from the Nucleobase-ascorbate transporter (NAT) family operates through a similar elevator mechanism, and residues outside of the primary binding site have also been shown to regulate substrate affinity, specificity, and transport dynamics, supporting a role for allosteric regulation of the transport cycle (50–54). In the Major Facilitator Superfamily (MFS) sugar transporters, multiple occluded state structures have been identified, suggesting that multiple gates may regulate function of these transporters as well, and may explain the ability of some transporters in this family to switch between proton-coupled sugar symport and uncoupled proton uniport (55). Here we show that even very small transporters, such as EmrE can have complex mechanisms of gating and transport regulation. This has broad relevance because phylogenetic analysis suggests that SMRs may have been progenitors of the MFS, Bacterial/Archaeal transporters (BAT), and drug-metabolite transporter (DMT) superfamilies as a whole (56, 57).

Materials and Methods

Microplate Growth Assays

Each EmrE construct was cloned into the pWB vector(11), a low copy number plasmid vector with a p15A origin and pTrc promoter, and transformed into *ΔemrE*- MG1655 *E. coli.* For experiments, LB plates were streaked and grown overnight at 37°C. In the morning, single colonies were picked to inoculate liquid LB cultures at 37°C. Once liquid cultures reached log phase growth, they were diluted back to an OD600 of 0.2 and further diluted 20-fold into microplates with LB media containing the indicated amount of substrate. Growth in microplates at 37°C was monitored for 15 hours using a TECAN Spark or BMG-Labtech microplate reader at OD700 (Ethidium) or OD600. Reported growth curves and final ODs are mean values of two biological replicates containing technical triplicates, with errors calculated using the standard deviation of the mean.

EmrE expression and purification

Protein expression utilized BL21 (Gold) DE3 *E. coli* transformed with a pET15b plasmid containing the respective EmrE construct, with cells grown in M9 minimal media. Protein was solubilized in decyl maltoside (DM) detergent and purified using immobilized nickel chromatography and size exclusion chromatography as previously described (13).

For pyranine fluorescence assays:

BL21 Gold (DE3) *E. coli* cells transformed with pET15b-EmrE, pET15b-E14QEmrE, or pET15-Δ107EmrE were grown in M9 minimal media to an OD600 of 0.9. The bacteria were flash cooled and then induced with 0.33M IPTG overnight at 17 °C. The *E. coli* cells were collected with centrifugation, lysed, and the membrane fraction solubilized with 40 mM DM. Purification was via Ni-NTA chromatography followed by size exclusion chromatography with a Superdex 200 column, with 10 mM decyl maltoside in all buffers (DM, Anatrace, Maumee, OH) as described (13). Protein concentrations were determined using absorbance at 280 nm with an extinction coefficient of 38,400 L/mol cm (55). Fractions containing EmrE in DM were reconstituted into a 3:1 mixture of 1-palmitoyl-2-oleoyl-glycero-3- phosphocholine (POPC, Avanti Polar Lipids, Alabaster, AL) and 1-palmitoyl-2- oleoyl-glycero-3-phosphoglycerol (POPG, Avanti Polar Lipids, Alabaster, AL) liposomes as follows. POPC and POPG in chloroform were dried under nitrogen, washed 3X with pentane to remove residual chloroform, and lyophilized overnight. Dry lipids were hydrated for 1 hr in 100 mM MOPS, 20 mM NaCl, and 1 mM pyranine, pH 6.5, sonicated for 1 minute before 0.5% octyl-glucoside was added. The mixture was sonicated for another 30 seconds and allowed to permeabilize for 15 min at room temperature. Hydrated lipids were mixed with EmrE in DM at a 400:1 lipid:protomer mol:mol

ratio (final lipid concentration 12 mg/ml) and allowed to equilibrate for 20 min. Detergent was removed by Biobeads as previously described (50). Proteoliposomes were extruded 11 times through a 0.2 µm filter (Avanti Polar Lipids, Alabaster, AL) and dialyzed overnight to remove residual pyranine. Proteoliposomes were then concentrated down 10 fold to allow for a final protein concentration of 2 µM upon dilution.

For SSME transport assays:
BL21 Gold (DE3) *E. coli* cells transformed with pET15b-EmrE or pET15- Δ107EmrE were grown in M9 minimal media to an OD600 of 0.9. The bacteria were flash cooled and then induced with 0.33M IPTG overnight at 17 °C. The *E. coli* cells were collected with centrifugation, lysed, and the membrane fraction solubilized with 40 mM DM. Purification was via Ni-NTA chromatography followed by size exclusion chromatography with a Superdex 200 column, with 10 mM decyl maltoside in all buffers (DM, Anatrace, Maumee, OH) as described13. Protein concentrations were determined using absorbance at 280 nm with an extinction coefficient of 38,400 L/mol cm (55). Fractions containing EmrE in DM were reconstituted into 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC, Avanti Polar Lipids, Alabaster, AL) liposomes as follows. POPC in chloroform was dried under nitrogen, washed 3X with pentane, and lyophilized overnight to remove residual chloroform. Dry lipids were hydrated for 1 hr in 50 mM MES, 50 mM MOPS, 50 mM bicine, 100 mM NaCl, and 2 mM MgCl2, pH 7, and permeabilized with 0.5% octyl-glucoside for 15 min at room temperature. Hydrated lipids were mixed with EmrE in DM at a 400:1 lipid:protomer mol:mol ratio (final lipid concentration 2.5 mg/ml) and allowed to equilibrate for 20 min. Detergent was removed by Biobeads as previously described(59). Proteoliposomes were extruded 11 times through a 0.2µm filter (Avanti Polar Lipids, Alabaster, AL) and flash frozen in aliquots stored at -80°C until needed for experiments.

For NMR:
Samples for 2D 1H-15N TROSY experiments, growth was carried out in perdeuterated M9 with 15N-NH4Cl as the sole nitrogen source, 2H-glucose as the sole carbon source, and 0.5g/L 2H,15N isogro. For D107-EmrE NMR assignment experiments, growth was carried out in perdeuterated M9 with 1 g 15NH4Cl, 0.75 g 2H,13C-glucose, and 0.5 g CND-Isogro per liter. Cells were harvested and EmrE purified in DM as described above. S200 fractions containing EmrE with 10 mM DM were reconstituted into DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine, Avanti Polar Lipids, Alabaster, AL) at 75:1 lipid:EmrE monomer mole ratio following the protocol in (50) using Biobeads (Biorad Laboratories, Hercules, CA) to remove detergent. EmrE proteoliposomes were collected by ultracentrifugation (100,000 g, 2 hr, 6 °C) and resuspended in NMR buffer with DHPC (1,2-dihexanoyl-*sn*- glycero-3-phosphocholine, Avanti Polar Lipids, Alabaster, AL) and freeze-thawed 3 times to create q=0.33 DMPC/DHPC bicelles (56) (q value confirmed with 1D proton NMR). Final NMR samples contained 0.7-1.0 mM EmrE monomer, 10% D2O, 0.05% NaN3, 2 mM TCEP (*tris*(2-carboxyethyl)phosphine), 2 mM EDTA (Ethylenediaminetetraacetic acid), and 2 mM DSS (4,4-dimethyl-4-silapentane-1- sulfonic acid)(59).

2.4.3 NMR spectroscopy
Triple resonance backbone walk experiments were acquired for backbone assignment of TPP+-bound Δ107-EmrE at pH 5.5 and 45°C using a sample with 1.25 mM 2H,15N,13C Δ107-EmrE and 16 mM TPP+. TROSY-HNCA, TROSY HNcoCA, TROSY-HNCACB experiments were acquired on a 900 MHz Bruker Avance III NMR spectrometer equipped with a TCI cryoprobe and TROSY HNCO, TROSY-HNcaCO experiments were acquired on a 750 MHz Bruker Avance III NMR spectrometer equipped with a TCI cryoprobe. Amide assignments were transferred to other pH values using pH titrations. 2D TROSY-HSQC and TROSY- selected ZZ-exchange spectra of D107-EmrE at pH 5.5 or pH 8.5, and TPP+-bound D107-EmrE at pH 5.5 or pH 7.7, were acquired on an 800 MHz Varian VNMRS DD spectrometer equipped with a 5 mm cryoprobe at 45°C using samples with 0.7-1 mM 2H,15N D107-EmrE using standard pulse sequences with gradient coherence selection. 70% of the backbone resonances of TPP+-bound Δ107-EmrE were assigned at pH 5.5 by combining standard triple resonance experiments (TROSY- HNCA, TROSY-HNCACB, TROSY-HNCO, TROSY-HN(CO)CA) with ZZ- exchange data. For NMR pH titrations, identical samples were prepared at the extreme pH values and the two samples were gradually mixed to create intermediate pH values, ensuring constant protein, lipid, and salt concentrations across the titration. To analyze the ZZ-exchange experiments, peak intensities were fit using the nlinls function in nmrPipe to accurate extract peak parameters. Residues for analysis were chosen that had all four peaks (two auto peaks, *I*AA and *I*BB, and two exchange cross-peaks, *I*AB and *I*BA) resolved in the 2D planes. Exchange with water reduces the peak intensity of the auto and cross-peak from the open face of the transporter at

high pH, resulting in greater scatter for the high pH data. The peak intensity ratio was calculated using the method developed by (25) the Palmer lab:

$$Peak\ intensity\ ratio = \frac{I_{AB}I_{BA}}{I_{BB}I_{AA} - I_{AB}I_{BA}} = k^2t^2 \qquad \text{(Eq. 1)}.$$

Calculation of this peak ratio cancels out initial peak intensity and intrinsic relaxation rates to first order, and depends on the mixing time (*t*) of the ZZ- exchange experiment in a simplified manner as shown in the equation above. Since the forward and reverse rate constants are identical for EmrE in bicelles (13), there is only a single rate constant for alternating access, *k*.

Pyranine Fluorescence Assays

All data were acquired an a TECAN spark instrument. The excitation wavelength was 465 nm (35 nm bandwidth) and the emission wavelength was 530 (25 nm bandwidth). The excitation spectrum maximum of pyranine shifts from 400 nm to 450 nm as pH increases, so with a constant 465 nm excitation wavelength, the observed fluorescence signal will increase as pH increases. The number of flashes was set to 30 to reduce well to well measurement time. To minimize instrument integration time, replicates were allowed to equilibrate for the full 30 minutes, and an average of the Z-position and gain recorded by the instrument were used as manual input for the reported assays. Liposome stocks with an internal buffer concentration of 100 mM MOPS, 20 mM NaCl, and 1 mM pyranine, pH 6.5aliquots were first pipetted into the plate which was then input into the instrument and the assay was started to perform instrument checks at which point the instrument was paused. The plate was ejected and 198 µL of 100 mM MOPS, 20 mM NaCl pH 7.5 buffer was pippeted into the well containing the liposomes and returned into the instrument to begin recording as soon as possible. Conditions with CCCP contained 1 µL of CCCP at 200 ug/mL on the opposite side of the well for a final concentration of 1 µg/µL. No gradient conditions were diluted into 198 µL of 100 mM MOPS, 20 mM NaCl pH 6.5 buffer. Reported data are average values of 3 replicate wells recorded for 30 minutes each to minimize well to well measuring times, with error bars representing the standard deviation of the mean.

Solid Supported Membrane Electrophysiology Transport Assays

All SSME data were acquired an a Nanion SURFE2R N1 instrument. Liposome aliquots were thawed, diluted 4-fold, and briefly sonicated. 10 µL of liposomes were added to prepare 3 mm sensors according to a standard protocol (60). For comparison of different mutants, sensors were prepared side-by-side for all variants (including all replicates) on the same day using a single batch of sensors to ensure maximum similarity in proteoliposome loading onto the sensor. While results obtained with different batches of sensors prepared on different days show similar results in terms of relative leak between variants, the absolute value varies from batch to batch and day to day. Thus, while ∆107-EmrE was always leakier than WT-EmrE, the absolute flux through the WT- or ∆107- transporter varied between batches of sensors prepared. Data was not averaged or compared across different batches of sensors. Equivalence of the SSME data and pyranine assay demonstrates the success of this approach. Prior to experiments, sensor capacitance and conductance values were obtained to ensure sensor quality. For all experiments, both internal and external buffers contained 50 mM MES, 50 mM MOPS, 50 mM bicine, 100 mM NaCl, and 2 mM MgCl2, with the pH and drug concentration as indicated for each data set. For data acquisition, sensors were equilibrated with internal buffer and transport was initiated by perfusion of the external buffer before re-equilibration with the internal buffer. Signals were obtained by integrating the current during perfusion of the external buffer, with the final 100 ms of the initial buffer equilibration used as the baseline. Reported data are average values of data recorded from at least three separate sensors, with error bars representing the standard deviation of the mean.

pH-Detected Liposomal Transport Assays

Liposomal transport assays were performed as previously described (9). Briefly, 1 mL aliquots with internal buffer (50 mM MOPS pH 7, 100 mM KCl) were thawed and extruded the day of the experiment as described above. The samples were run over 2 PD-10 spin columns (Cytiva) equilibrated in external buffer (50 µM MES pH 6 with 1 mM KCl and 99 mM NaCl) following the manufacturer's spin protocol. Samples were then diluted to 1.5 mL in external buffer. Eluted

samples were added to 2 mL cuvettes with a stir bar and a microelectrode was inserted and allowed to equilibrate. The pH was monitored in real time by a WINDAQ DI-710 from DataQ at a rate of 100 per second. Aliquots of valinomycin and CCCP at 1 mg/ml in 100% DMSO were thawed and diluted by half in external buffer to better match the pH. During the recordings valinomycin was added to a final concentration of 1 μg/mL to create a $\Delta\Psi$, Harmane to a concentration of 100 μM, CCCP to a concentration of 1 μg/mL as a control, and 50 nmol of HCl was added for quantification.

Molecular Dynamics

All MD simulations were conducted with GROMACS 2020.4(61). Simulation inputs were generated by CHARMM-GUI membrane bilayer builder (27). The protein was solvated by 162 DMPC molecules, and 40 mM NaCl was added to the water to neutralize the system. The system was coupled to a Nose-Hoover thermostat (62, 63) and a Parrinello-Rahman barostat (64), at 310 K and 1 bar, respectively. The system was minimized, then equilibrated with position constraints gradually releasing, as by the default setting of CHARMM-GUI. Then, the system was further equilibrated for 400 ns without constraint, where the RMSD plateaued, and the box sizes were stable. For the WT D107-EmrE simulation, we observed lipids penetrating the protein in this 400 ns equilibration, so we added another 100 ns simulation with backbone constraints to further equilibrate the membrane before releasing these constraints again. Figures were rendered with ChimeraX. The hydrogen distances were analyzed with PLUMED (65), by computing the softmin of all hydrogen-hydrogen distances between two sidechains with β=500.

Water path length calculation

The water path length calculation was implemented in an in-house modification of PLUMED. The algorithm can be briefly described as follows: Each water oxygen is considered a node in a graph, and the distance for each edge connecting two node is determined by a function that is close to 1 when the oxygen-oxygen distance is smaller than a $r0$ and grows rapidly when it is larger than $r0$. This $r0$ is set to be 3 Å, which is the typical distance between the oxygens of hydrogen- bonded water. A more detailed discussion can be found in reference (29). Then, for each frame in the trajectory, the shortest path is found for the graph. We used the two oxygens of E14B at the starting point and the midpoint of Cα of R102A and G57A as the destination.

Umbrella sampling with MS-RMD

The simulations were run with the LAMMPS MD engine and the umbrella sampling was carried out as implemented in PLUMED (65, 66). The codes were co-compiled with a plug-in developed in house called RAPTOR to model proton transport reactions (33). The source code of RAPTOR is available at https://github.com/uchicago-voth/raptor. The starting structure was taken from the classical MD simulation of WT-EmrE. A water molecule was protonated at the mouth of the channel and steered MD was then used to create initial configurations at different collective variable (CV) values. A total of 43 umbrella windows spanning from CV = 0.0 Å to 15.0 Å were used, with a varying restraint force constant of 80 kcal/mol/Å2 to 15 kcal/mol/Å2. The CV is defined as

$$x = d_{OC} \cdot e_{PT} \qquad \text{(Eq. 2)}$$

where d_{OC} is a vector pointing from the closer glutamate oxygen to the CEC, and e_{PT} is a unit vector of the direction of proton transport. Each umbrella window was equilibrated for 1 ns and then the production run was for 2 ns. The PMF was reconstructed with the weighted histogram analysis method (WHAM) (67). In these simulations, the Cα of the residues at least 10 Å away from the path and those in TM1-2 were restrained to its initial coordinate with a 2.4 kcal/mol/Å2 harmonic potential to ensure the bias force does not unrealistically distort the protein conformation.

Acknowledgments

This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH grant R24GM141526 (NIGMS). Research reported in this publication was supported by the National Institute of General Medical Sciences of the NIH through grant R01GM053148 (to GAV) and R01GM095839, R35GM141748 (to KHW). Computational resources were provided by the Research

Computing Center (RCC) at the University of Chicago. M. Brousseau was supported in part by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number T32GM008505 (Chemistry–Biology Interface Training Program). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. J. M. Munita, C. A. Arias, Mechanisms of Antibiotic Resistance. *Compend Contin Educ Vet* **23**, 464–472 (2001).

2. M. H. Brown, R. A. Skurray, Staphylococcal Multidrug Efflux Protein QacA. *J Mol Microbiol Biotechnol* **3**, 163–170 (2001).

3. M. Pérez-Varela, J. Corral, J. Aranda, J. Barbé, Roles of Efflux Pumps from Different Superfamilies in the Surface-Associated Motility and Virulence of Acinetobacter baumannii ATCC 17978. *Antimicrob Agents Chemother* **63**, 1–11 (2019).

4. S. Schuldiner, EmrE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochim Biophys Acta Proteins Proteom* **1794**, 748–762 (2009).

5. H. Yerushalmi, S. Schuldiner, A model for coupling of H+ and substrate fluxes based on "time-sharing" of a common binding site. *Biochemistry* **39**, 14711–14719 (2000).

6. L. R. Forrest, R. Krämer, C. Ziegler, The structural basis of secondary active transport mechanisms. *Biochim Biophys Acta - Bioenerg* **1807**, 167–188 (2011).

7. O. Boudker, G. Verdon, Structural perspectives on secondary active transporters. *Trends Pharmacol Sci* **31**, 418–426 (2010).

8. D. Rotem, S. Schuldiner, EmrE, a multidrug transporter from Escherichia coli, transports monovalent and divalent substrates with the same stoichiometry. *J Biol Chem* **279**, 48787–48793 (2004).

9. A. E. Robinson, N. E. Thomas, E. A. Morrison, B. M. Balthazor, K. A. Henzler-Wildman, New free-exchange model of EmrE transport. *Proc Natl Acad Sci USA* **114**, E10083–E10091 (2017).

10. G. A. Hussey, N. E. Thomas, K. A. Henzler-Wildman, Highly coupled transport can be achieved in free-exchange transport models. *J Gen Physiol* **152**, e201912437 (2020).

11. P. J. Spreacker, *et al.*, Activating alternative transport modes in a multidrug resistance efflux pump to confer chemical susceptibility. *Nat Commun* **13**, 7655 (2022).

12. N. E. Thomas, *et al.*, The C terminus of the bacterial multidrug transporter EmrE couples drug binding to proton release. *J Biol Chem* **293**, 19137–19147 (2018).

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

14. C. Glaubitz, *et al.*, ^{31}P-CPMAS NMR studies on TPP^+ bound to the ion-coupled multidrug transport protein EmrE. *FEBS Lett* **480**, 127–131 (2000).

15. S. J. Fleishman, *et al.*, Quasi-symmetry in the Cryo-EM Structure of EmrE Provides the Key to Modeling its Transmembrane Domain. *J Mol Biol* **364**, 54–67 (2006).

16. Y. J. Chen, *et al.*, X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci USA* **104**, 18999–19004 (2007).

17. A. A. Kermani, *et al.*, Crystal structures of bacterial Small Multidrug Resistance transporter EmrE in complex with structurally diverse substrates. *Elife* **11**, 76766 (2022).

18. A. A. Shcherbakov, P. J. Spreacker, A. J. Dregni, K. A. Henzler-Wildman, M. Hong, High-pH structure of EmrE reveals the mechanism of proton-coupled substrate transport. *Nat Commun* **13**, 991 (2022).

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

20. J. Li, *et al.*, Dynamics underlie the drug recognition mechanism by the efflux transporter EmrE. *Nat Commun* **15**, 4537 (2024).

21. J. V. Vermaas, S. B. Rempe, E. Tajkhorshid, Electrostatic lock in the transport cycle of the multidrug resistance transporter EmrE. *Proc Natl Acad Sci USA* **115**, E7502–E7511 (2018).

22. M. V. Baev, D. Baev, A. J. Radek, J. W. Campbell, Growth of Escherichia coli MG1655 on LB medium: Monitoring utilization of sugars, alcohols, and organic acids with transcriptional microarrays. *Appl Microbiol Biotechnol* **71**, 310–316 (2006).

23. E. A. Morrison, A. E. Robinson, Y. Liu, K. A. Henzler-Wildman, Asymmetric protonation of EmrE. *J Gen Physiol* **146**, 445–461 (2015).

24. Y. Li, A. G. Palmer, TROSY-selected ZZ-exchange experiment for characterizing slow chemical exchange in large proteins. *J Biomol NMR* **45**, 357–360 (2009).

25. V. Z. Miloushev, A. G. Palmer, R1ρ relaxation for two-site chemical exchange: General approximations and some exact solutions. *J Magn Reson* **177**, 221–227 (2005).

26. A. A. Shcherbakov, *et al.*, Structure and dynamics of the drug-bound bacterial transporter EmrE in lipid bilayers. *Nat Commun* **12**, 1–13 (2021).

27. J. Lee, *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).

28. N. Agmon, The Grotthuss mechanism. *Chem Phys Lett* **244**, 456–462 (1995).

29. C. Li, G. A. Voth, A quantitative paradigm for water-assisted proton transport through proteins and other confined spaces. *Proc Natl Acad Sci USA* **118**, e2113141118 (2021).

30. J. Jurasz, M. Bagiński, J. Czub, M. Wieczór, Molecular mechanism of proton-coupled ligand translocation by the bacterial efflux pump EmrE. *PLoS Comput Biol* **17**, e1009454 (2021).

31. P. Schatzberg, Molecular diameter of water from solubility and diffusion measurements. *J Phys Chem* **71**, 4569–4570 (1967).

32. B. Ilan, E. Tajkhorshid, K. Schulten, G. A. Voth, The mechanism of proton exclusion in aquaporin channels. *Proteins: Struct, Funct, and Bioinf* **55**, 223–228 (2004).

33. S. Kaiser, *et al.*, Molecular dynamics simulation of complex reactivity with the Rapid Approach for Proton Transport and Other Reactions (RAPTOR) software package. *ChemRxiv*. https://doi.org/10.26434/chemrxiv-2024-x6031.

34. C. Li, G. A. Voth, Accurate and Transferable Reactive Molecular Dynamics Models from Constrained Density Functional Theory. *J Phys Chem B* **125**, 10471–10480 (2021).

35. C. Wu, *et al.*, Identification of an Alternating-Access Dynamics Mutant of EmrE with Impaired Transport. *J Mol Biol* **431**, 2777–2789 (2019).

36. G. R. Grimsley, J. M. Scholtz, C. N. Pace, A summary of the measured pK values of the ionizable groups in folded proteins. *Prot Sci* **18**, 247–251 (2009).

37. O. Beckstein, F. Naughton, General principles of secondary active transporter function. *Biophys Rev* **3**, 011307 (2022).

38. Y. Adam, N. Tayer, D. Rotem, G. Schreiber, S. Schuldiner, The fast release of sticky protons: Kinetics of substrate binding and proton release in a multidrug transporter. *Proc Natl Acad Sci USA* **104**, 17989-17994 (2007).

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

40. O. E. Burata, T. J. Yeh, C. B. Macdonald, R. B. Stockbridge, Still rocking in the structural era: A molecular overview of the small multidrug resistance (SMR) transporter family. *J Biol Chem* **298**, 102482 (2022).

41. S. S. Mordoch, D. Granot, M. Lebendiker, S. Schuldiner, "Scanning Cysteine Accessibility of EmrE, an H-coupled Multidrug Transporter from Escherichia coli, Reveals a Hydrophobic Pathway for Solutes. *J Biol Chem* **274**, 19480-19486 (1999).

42. P. Lloris-Garcerá, *et al.*, In vivo trp scanning of the small multidrug resistance protein EmrE confirms 3D structure models. *J Mol Biol* **425**, 4642–4651 (2013).

43. S. T. Amadi, H. A. Koteiche, S. Mishra, H. S. Mchaourab, Structure, dynamics, and substrate-induced conformational changes of the multidrug transporter EmrE in liposomes. *J Biol Chem* **285**, 26710–26718 (2010).

44. H. Yerushalmi, S. Schuldiner, An Essential Glutamyl Residue in EmrE, a Multidrug Antiporter from Escherichia coli. *J Biol Chem* **275**, 5264–5269 (2000).

45. G. Diallinas, Understanding transporter specificity and the discrete appearance of channel-like gating domains in transporters. *Front Pharmacol* **5**, 105861 (2014).

46. G. Rudnick, Cytoplasmic permeation pathway of neurotransmitter transporters. *Biochemistry* **50**, 7462–7475 (2011).

47. N. Reyes, C. Ginter, O. Boudker, Transport mechanism of a bacterial homologue of glutamate transporters. *Nature* **462**, 880–885 (2009).

48. E. A. Riederer, F. I. Valiyaveetil, Investigation of the allosteric coupling mechanism in a glutamate transporter homolog via unnatural amino acid mutagenesis. *Proc Natl Acad Sci USA* **116**, 15939–15946 (2019).

49. S. Oh, O. Boudker, Kinetic mechanism of coupled binding in sodium-aspartate symporter GltPh. *Elife* (2018). https://doi.org/10.7554/eLife.37291.001.

50. V. Kosti, I. Papageorgiou, G. Diallinas, Dynamic elements at both cytoplasmically and extracellularly facing sides of the UapA transporter selectively control the accessibility of substrates to their translocation pathway. *J Mol Biol* **397**, 1132–1143 (2010).

51. A. Vlanti, S. Amillis, M. Koukaki, G. Diallinas, A novel-type substrate-selectivity filter and ER-exit determinants in the UapA purine transporter. *J Mol Biol* **357**, 808–819 (2006).

52. M. Koukaki, *et al.*, The Nucleobase-ascorbate transporter (NAT) signature motif in UapA defines the function of the purine translocation pathway. *J Mol Biol* **350**, 499–513 (2005).

53. I. Papageorgiou, *et al.*, Specific Interdomain Synergy in the UapA Transporter Determines Its Unique Specificity for Uric Acid among NAT Carriers. *J Mol Biol* **382**, 1121–1135 (2008).

54. G. Diallinas, Allopurinol and xanthine use different translocation mechanisms and trajectories in the fungal UapA transporter. *Biochimie* **95**, 1755–1764 (2013).

55. M. G. Madej, L. Sun, N. Yan, H. R. Kaback, Functional architecture of MFS D-glucose transporters. *Proc Natl Acad Sci USA* **111**, E719-E727 (2014).

56. D. C. Bay, R. J. Turner, Diversity and evolution of the small multidrug resistance protein family. *BMC Evol Biol* **9**, 140 (2009).

57. D. L. Jack, N. M. Yang, M. H. Saier, The drug/metabolite transporter superfamily. *Eur J Biochem* **268**, 3620–3639 (2001).

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

59. N. E. Thomas, W. Feng, K. A. Henzler-Wildman, A solid-supported membrane electrophysiology assay for efficient characterization of ion-coupled transport. *J Biol Chem* **297**, 101220 (2021).

60. M. J. Abraham, *et al.*, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

61. The PLUMED consortium, Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* **16**, 670–673 (2019).

Figures

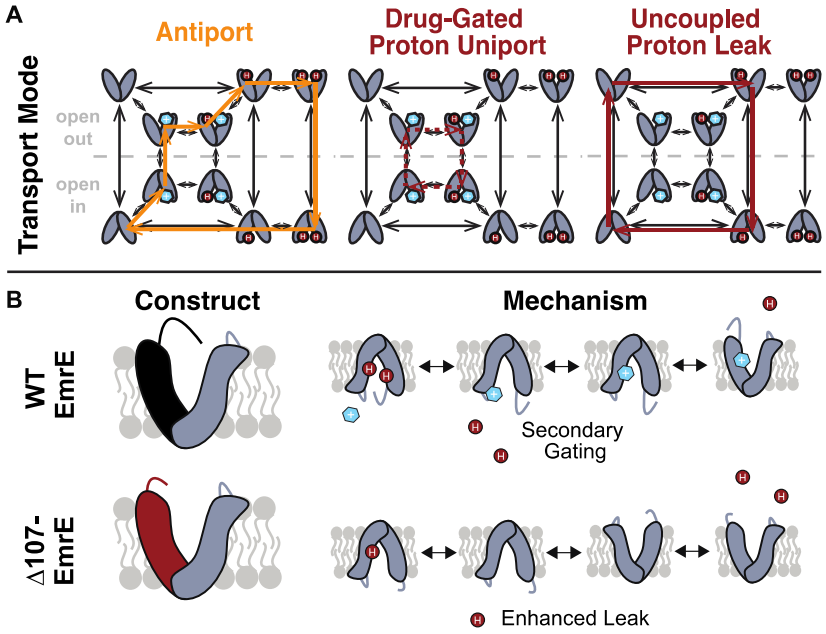

Figure 1. Model of coupled antiport and uncoupled proton leak through EmrE. (A) All of the drug- and proton-bound states that are reasonably populated at near physiological temperature and pH and the transitions between these states observed by NMR leads to a model for EmrE transport that allows for both coupled antiport (orange) and proton leak (red solid line). (B) In WT-EmrE, the C-terminal tail on the open face acts as a secondary gate (top), minimizing proton leak in the absence of substrate. Truncation of EmrE in Δ107-EmrE removes this gate (bottom). (C) The Drug binding to a secondary binding site near the tail opens the gate (top), allowing proton exit from the primary binding site near E14, and drug to progress to the primary binding site at E14. This leads to either coupled antiport (A, orange) as shown. If the substrate does not rapidly move into the primary binding site, only proton entry/exit occurs upon opening of the secondary gate, resulting in drug-gated proton leak (A, red dashed line). Truncation of the C-terminal tail in Δ107-EmrE (bottom) allows uncoupled proton leak in the absence of substrate (A, red solid line).

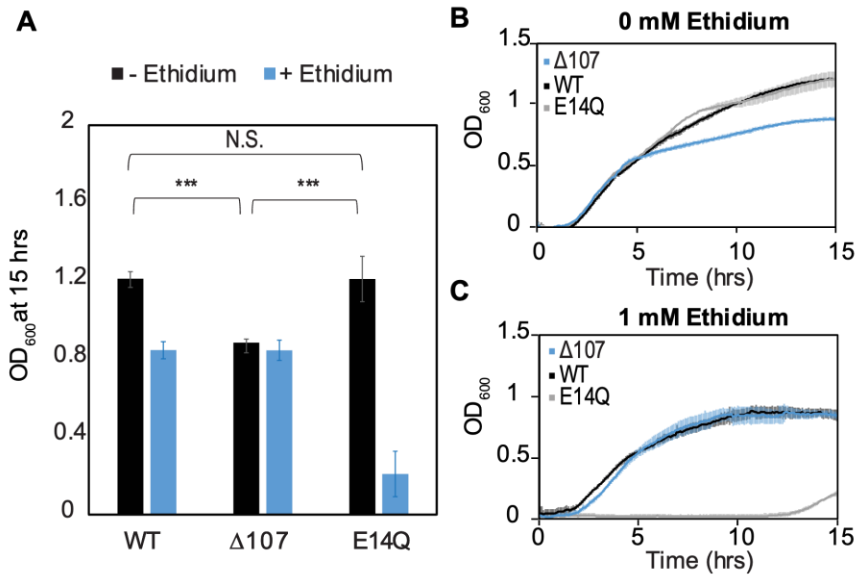

Figure 2. C-terminal tail truncation does not impair the ability of EmrE to confer resistance to toxic compounds. (A-C) WT-, E14Q-, or Δ107-EmrE was heterologously expressed in MG1655-Δ*emre E. coli* using a plasmid with p15 origin and pTrc promoter without induction to minimize any growth defect due to expression. *In vivo* growth assays were monitored by OD_{700} to allow consistent monitoring in the absence (B) or presence of (C) ethidium bromide. Growth at 15 hours (A) shows identical growth for WT-EmrE and Δ107-EmrE in the presence of ethidium, while E14Q-EmrE is severely impaired (A,B). There is a 20% reduction in growth for Δ107-EmrE relative to WT-EmrE or non-functional EmrE ($p < 0.001$), but this does not prevent the mutant from transporting ethidium out of the cell and thus conferring resistance (A,C). The error bars show the standard deviation across six replicates (two biological replicates with three technical replicates each). All *p*-values were calculated from a two-sided *t*-test. (*) $p < 0.05$ (**) $p < 0.01$ (***) $p < 0.001$.

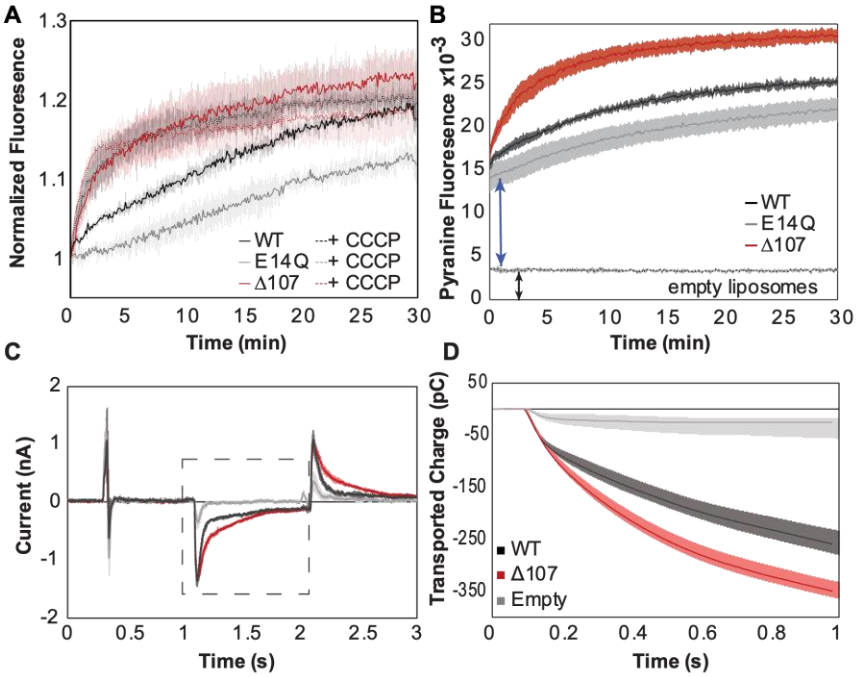

Figure 3. C-terminal tail truncation enhances proton leak. (A-B) Pyranine fluorescence directly reports on proton leak through EmrE. (A) WT (black), ∆107 (red, to distinguish *in vitro* assays from the cellular assays of Fig. 2) or E14Q-EmrE (gray) proteoliposomes with 1 mM internal pyranine and internal pH 6.5 were diluted 100-fold into pH 7.5 buffer (solid lines) or pH 7.5 buffer with CCCP (dashed lines) and fluorescence was normalized to time zero. CCCP is a protonophore, providing a positive control for maximal proton leak under these conditions. (B) Pyranine fluorescence normalized by subtracting the fluorescence of proteolipsomes diluted into pH 6.5 (no gradient, baseline) from the fluorescence of proteoliposomes diluted into pH 7.5 (transport) shows intraliposomal pH change with proteoliposomes in the lag time prior to initial fluorescence read and increased intraliposomal pH change for ∆107-EmrE than WT-EmrE or E14Q-EmrE. (C-E) Solid supported membrane electrophysiology data shows measurable charge movement through WT- and Δ107-EmrE proteoliposomes in the presence of a pH gradient alone, as compared to empty liposomes, with increased charge transport through ∆107-EmrE**.** (C) Current is recorded in real time as a matching pH internal buffer (pH 6.5) is flowed over the liposomes to establish baseline, then a higher pH (pH 7) buffer is flowed over the liposomes to create an outwardly directed proton gradient (dashed box), and finally the initial buffer (pH 6.5) is flowed back over the liposomes to reverse the charge movement and return to baseline. (D) The recorded current during the period of the applied gradient (dashed box, C) is integrated to determine the transported charge during that time. In all cases, Δ107-EmrE shows increased proton leak compared to WT-EmrE and controls. The error bars show the standard deviation across three replicates or sensors.

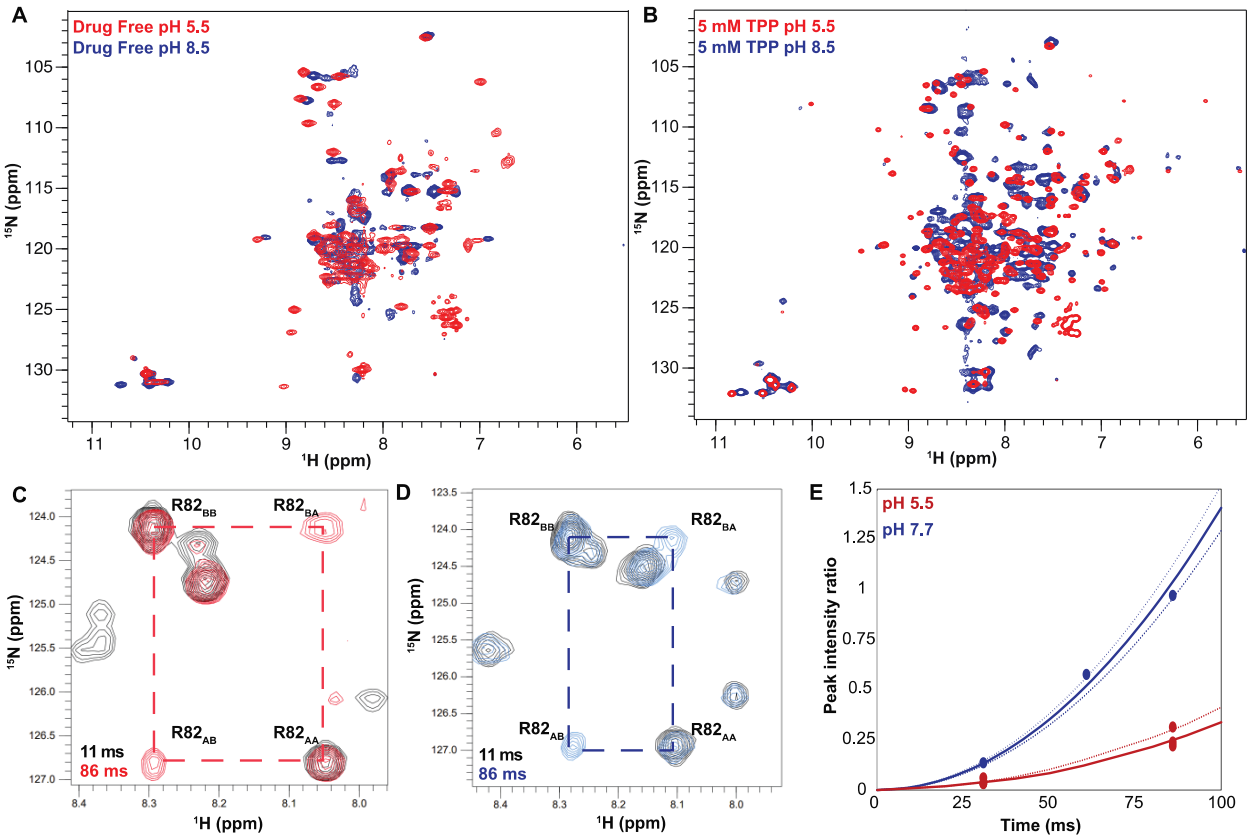

Figure 4. The pH dependence of alternating access in Δ107-EmrE is distinct from WT-EmrE. TROSY-HSQC spectra of Δ107-EmrE in the absence (A) and presence (B) of the tight-binding ligand tetraphenylphosphonium (TPP). While drug binding slows the dynamics of the protein at both low (red) and high (blue) pH, as evident by the better spectral quality in B, in both drug-free and drug-bound Δ107-EmrE the dynamics of the mutant are highly sensitive to the pH conditions. ZZ-Exchange Spectroscopy of Δ107-EmrE bound to of TPP was used to quantify the alternating-access rates at low and high pH. ZZ-exchange spectra with the indicated delays are shown for *(C)* pH 5.5 and *(D)* pH 7.7. *(E)* The composite peak intensity ratios for F78, G80, R82, L83 and R106 fit to an exchange rate of 4±1 s^{-1} at pH 5.5. At pH 7.7, the composite peak intensity ratios for G80, R82, L83 and R106 fit to an exchange rate of 17±3 s^{-1}.

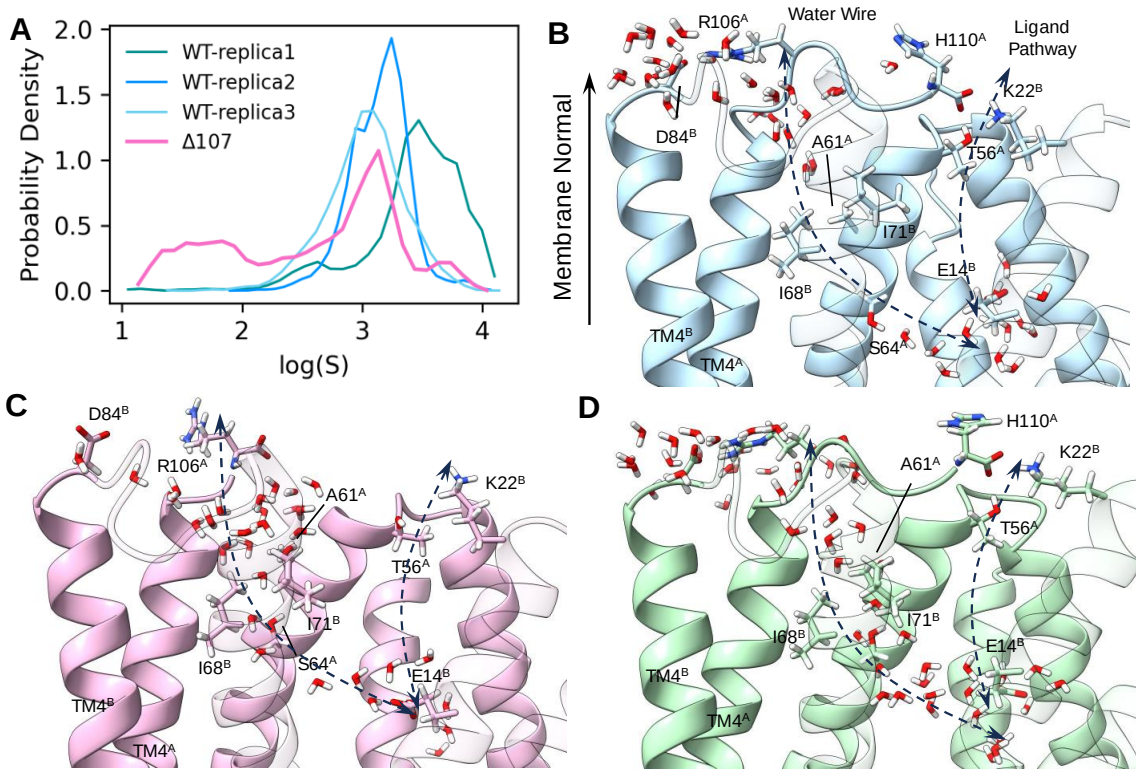

Figure 5. The C-terminus tail caps the water wire from the open side. (A) Logarithm of the minimum water distance log(*S*) histogram. (B-D) The following panels illustrate a few snapshots in the simulation. The membrane normal vector points to the open side of EmrE. Two dashed arrows show the ligand pathway and the water chain respectively. TM1 to TM3 in subunit B is shown transparently to better illustrate the interface between the two subunits. (B) Dry snapshot of WT-EmrE. (C) Wet snapshot of Δ107-EmrE. (D) A rare event snapshot when WT-EmrE is hydrated. The color codes are the same as in Figure 6. Yellow stars highlight the backbone of the C-terminal residue (R106 or H110) and the yellow arrowhead (B, D) highlights the backbone of R106 in the full length construct to illustrate where the tail would terminate in Δ107.

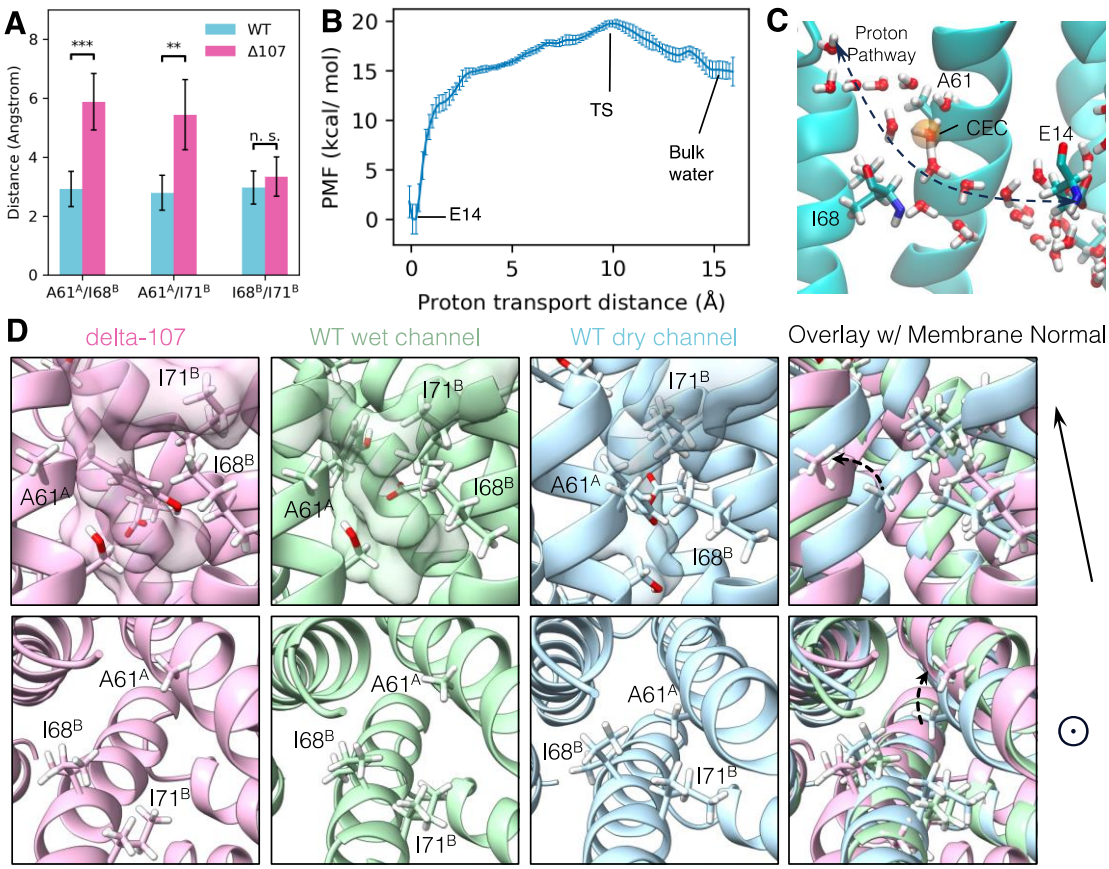

Figure 6. The structural basis of the C-terminus gating. (A) The minimum distance between side chain hydrogens for A61^{A}, I68^{B}, and I71^{B}. In WT-EmrE, the sidechain of A61^{A} is significantly closer to I68^{B} and I71^{B}, while the distance between I68^{B} and I71^{B} does not change significantly. The error bars show the standard deviation along the trajectory. All *p*-values were calculated from a two-sided *t*-test, (**) $p < 0.01$ (***) $p < 0.001$ (B) The proton transport potential of mean force (PMF), as a function of the distance between the center of the excess charge (CEC) and the donor (E14) on the direction of transport (See Eq.1 in Methods). (C) A snapshot of the transition state. The orange sphere is the proton CEC. (D) Conformations of the A61^{A}, I68^{B}, I71^{B} triad from two different angles. The membrane normal shown at the right points to the open side of EmrE. The upper panels are from a side view, and the lower panels are looking top-down into the primary binding site from the open side. The transparent surface in the upper panels shows the water wire. Unlabeled residues shown as stick representation are E14^{B}, Y60^{A}, and S64^{A}.

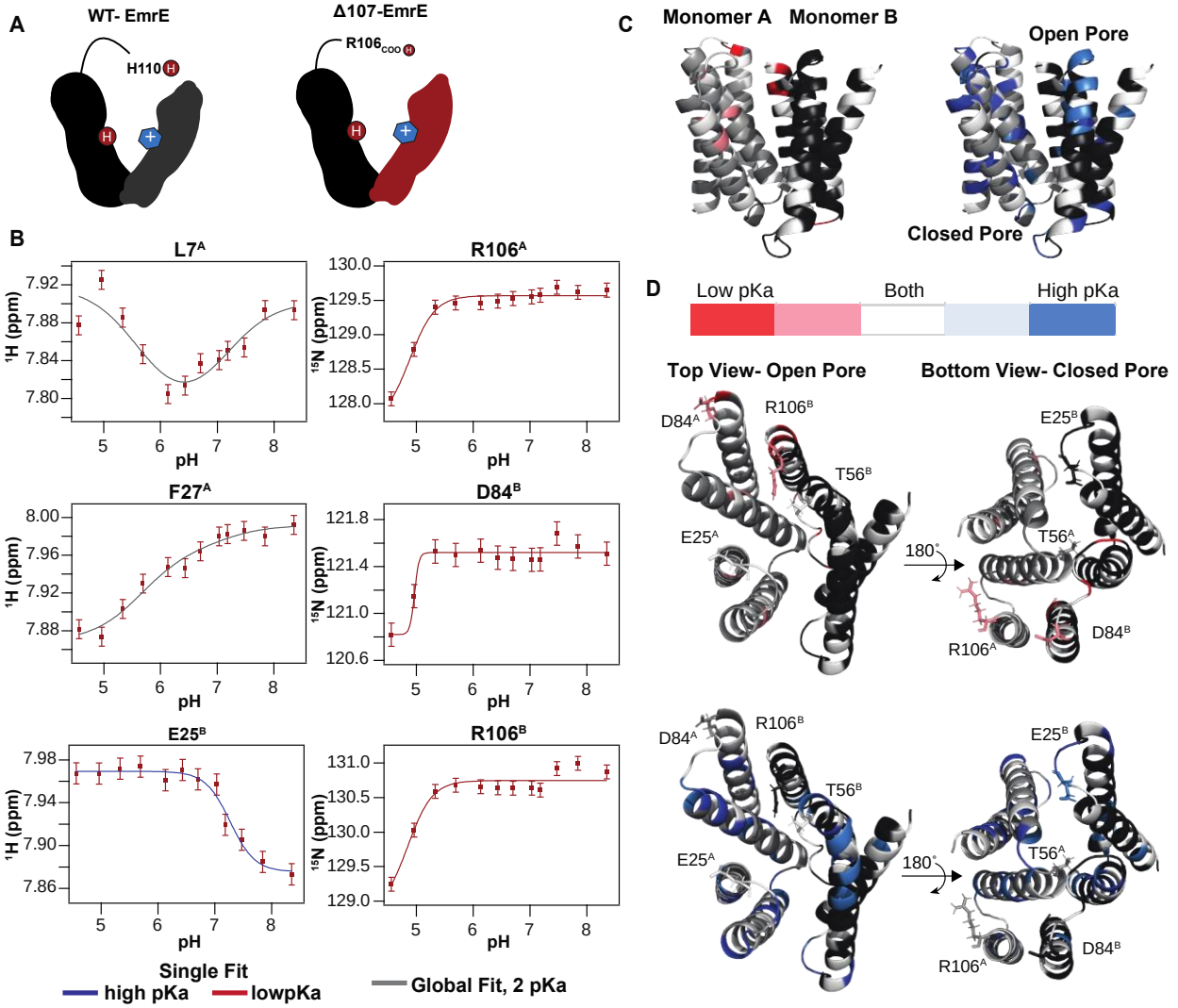

Figure 7. pH titration of TPP⁺-bound Δ107-EmrE supports the possibility of secondary gating. (A) In WT-EmrE bound to TPP⁺, one E14 residue and one H110 residue, are the only titratable sites (dark red circles labeled H⁺). In Δ107-EmrE, H110 is not present, suggesting that only one titratable group should remain (E14). (B) The proton and nitrogen chemical shifts for individual residues of TPP⁺-bound Δ107-EmrE were recorded as a function of pH. The resulting titration profiles do not show the expected single-pKa pattern. Some are curved, consistent with multiple pKa values, and others are consistent with a single pKa but at either high or low pH. All of the data can be globally fit to two pKa values, using either a 2-pKa fit (5.6 and 7.1, grey) or single pKa fit at the relevant value (5.6, red; 7.1, blue). (C) Residues sensitive to each pKa value are plotted on the faRM model (21) using the indicated color scale. (D) Residues strongly sensing the lower pKa value cluster around the C-terminus (R106) and 3-4 loop (includes residue D84) on both the open and closed face of the transporter, while the 1-2 loop (includes residue E25) and T56 on the open side of the pore sense both pKa values (Left).

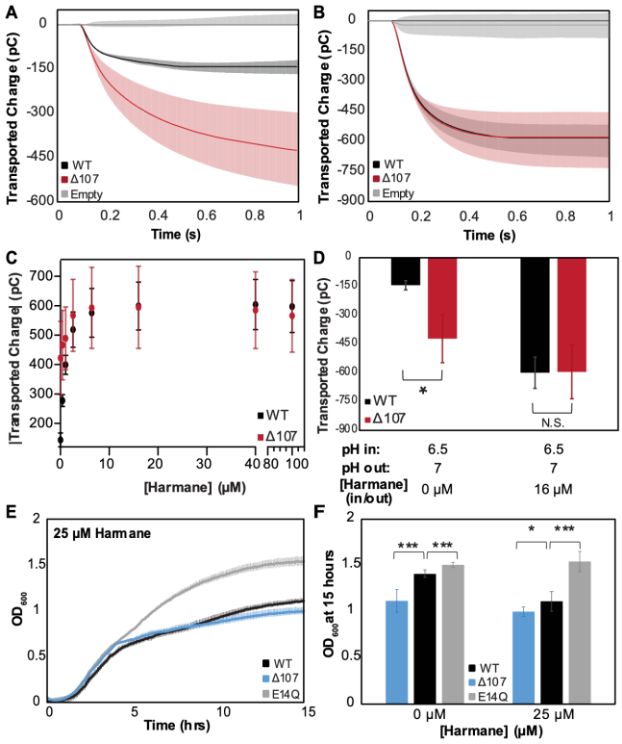

Figure 8. Intrinsic leak in ∆107-EmrE does not synergize with harmane-induced leak. SSME traces of transported charge corresponding to proton leak in the absence (A) and presence (B) of harmane show that 16 µm harmane induces leak in WT-EmrE that is comparable to the leak observed through ∆107-EmrE in the absence of harmane. In the presence of increasing concentrations of harmane (C) the leak signal for WT-EmrE quickly converges to that of ∆107-EmrE. The leak observed for ∆107-EmrE is more variable, displaying larger standard deviations than WT-EmrE proteoliposomes (C). This could be due to greater variability in the unregulated transport activity of ∆107-EmrE compared to harmane-gated leak in WT-EmrE, and the impact of this unregulated behavior on the sensitivity of SSME to variation in the absolute number of proteoliposomes adsorbed on the surface sensor. Growth assays in the absence of substrate (Fig. 1A) show a clear growth defect for *E. coli* expressing ∆107-EmrE compared to WT-EmrE, which is nearly eliminated when cells are grown in the presence of 25 µM harmane (E-F). ∆107-EmrE data is shown in blue for cellular assays and red for *in vitro* assays to readily distinguish the assay type. The error bars show the standard deviation across 3 sensors for SSME or across six replicates for growth assays (two biological replicates with three technical replicates each). All *p*-values were calculated from a two-sided *t*-test (*) $p < 0.05$ (**) $p < 0.01$ (***) $p < 0.001$.