Supporting Information

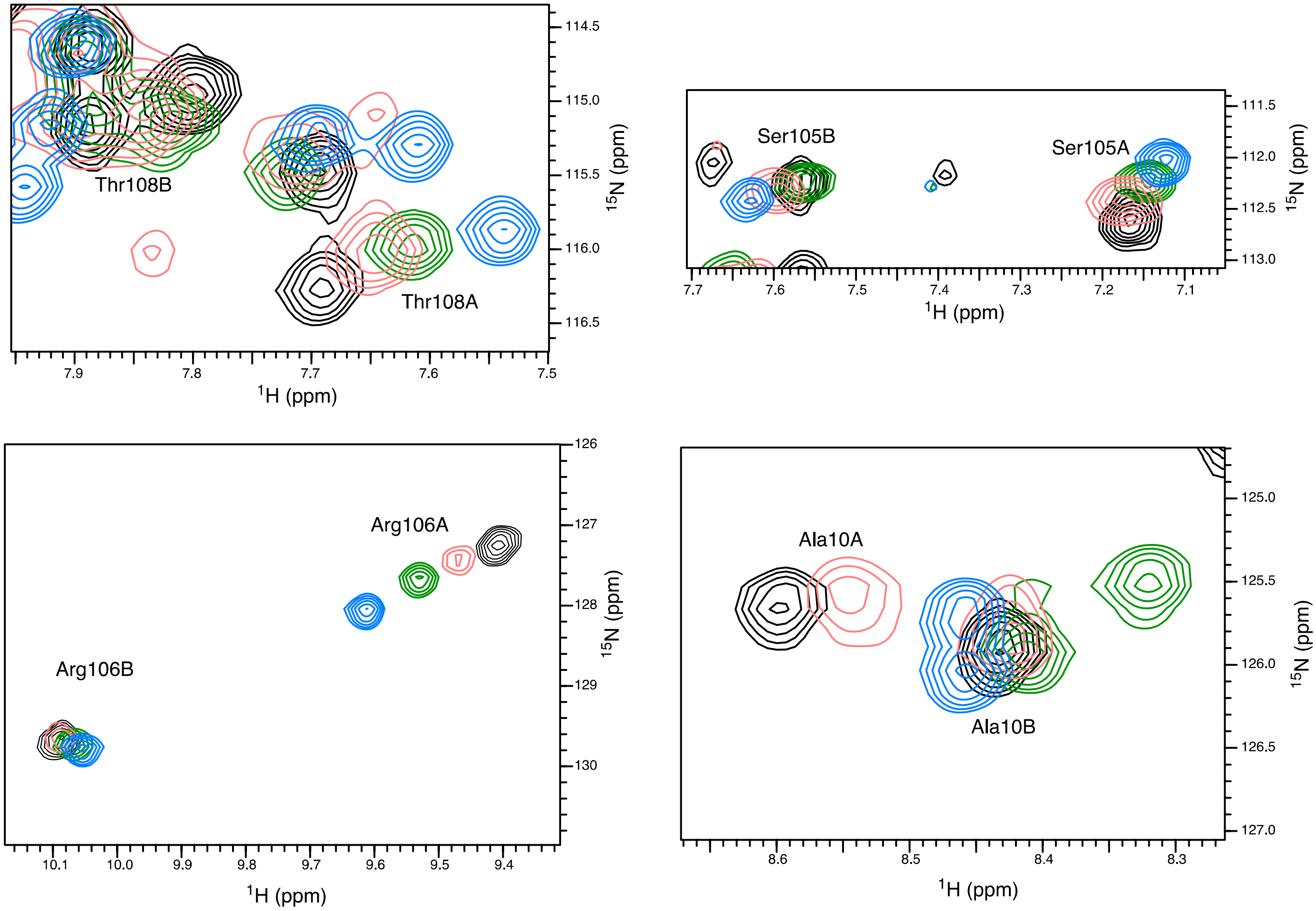

Figure S1 – *Effect of bound drug identity on C-terminal tail.* Spectra of selected residues from EmrE bound to tetraphenylphosphonium (TPP$^+$, **black**), ethyltriphenylphosphonium (EtTPP$^+$, **pink**), 2,5-diethoxyphenyltriphenylphosphonium (DPhTPP$^+$, **blue**), or methylbenzyltriphenylphosphonium (MBTPP$^+$, **green**) at pH 7. Residues in the monomer A C-terminal tail (S105, R106, T108) display significant chemical shift differences when EmrE is bound to different drugs, as we previously reported in (1). These chemical shift perturbations are comparable to those observed for residues in the active site, such as A10 which is located only one turn away from E14 in TM1.

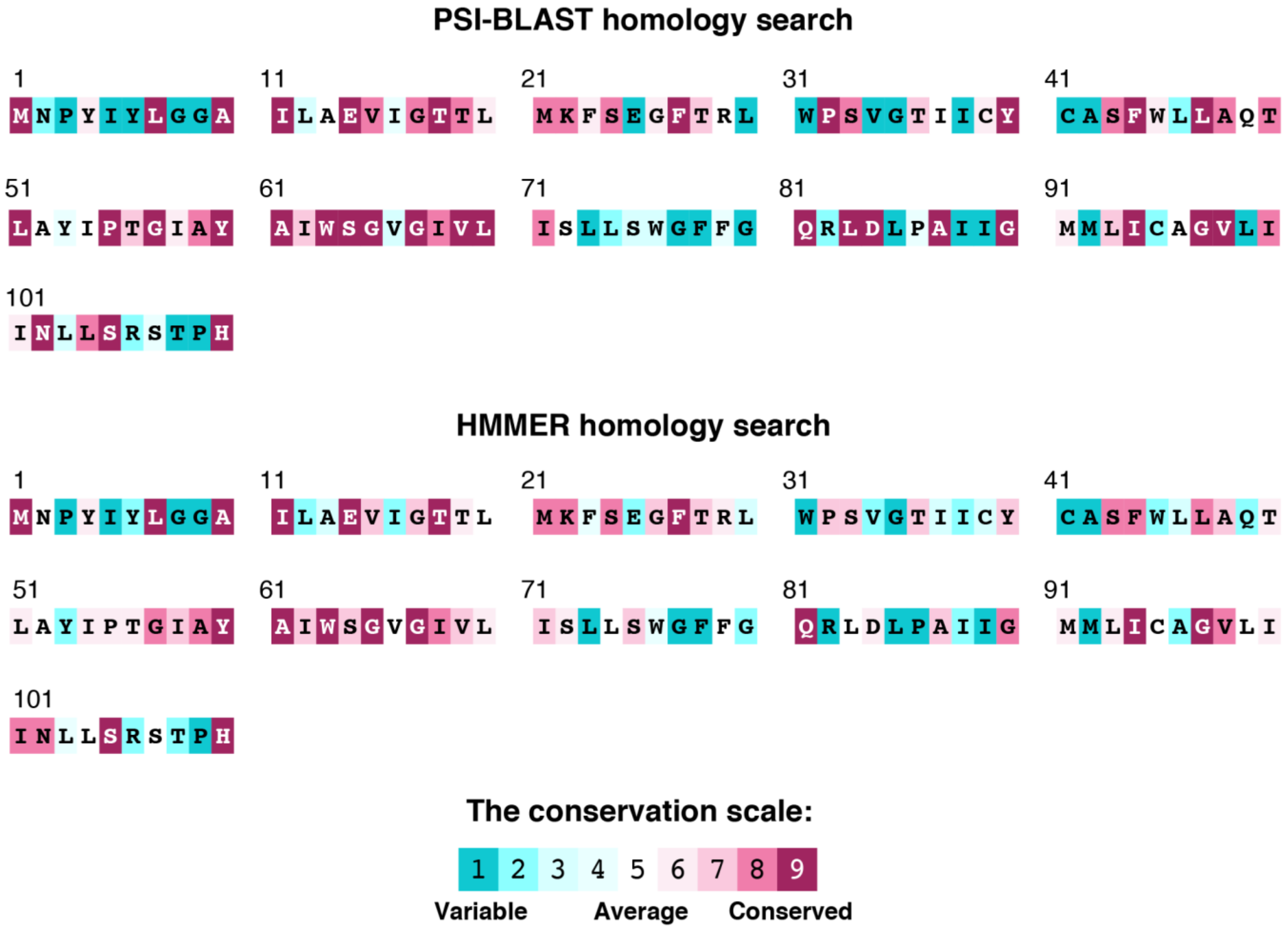

Figure S2 – *Sequence conservation for EmrE*. ConSurf analysis (2–5) was performed on the EmrE amino acid sequence using either five iterations of PSI-BLAST (6) or one iteration of HMMER (7) to search for homologues. Homologous sequences were clustered using CD-HIT with a 90% identity threshold, resulting in 477 results from the PSI-BLAST search and 5979 results from the HMMER search. 300 representative sequences from each search were aligned using MAFFT (8), and calculated conservation scores are shown.

Both homology searches display high conservation of residues that are known to be functionally important, including E14. Additionally, in both searches, H110 was shown to be maximally conserved. However, the PSI-BLAST alignment displayed significantly greater conservation overall. Recent work has shown that the majority of transporters within the SMR family are highly selective guanidinium transporters, and that EmrE is a member of a smaller clade of promiscuous toxin efflux proteins within this family (9). 10% of the 300 representative sequences from the HMMER search were annotated as SugE, the prototypical guanidinium exporter, while SugE was not present at all in the results from the PSI-BLAST search, indicating that the PSI-BLAST search returned sequences primarily within the EmrE-like toxin efflux clade while the larger HMMER dataset included sequences across the family. Notably, residues with higher conservation in the PSI-BLAST search include aromatic residues Y40 and F44 known to be important for binding and transport of aromatic cations that are substrates of EmrE (10–13). D84 in the TM3-4 loop is also much more highly conserved within the EmrE-like clade than in the family overall.

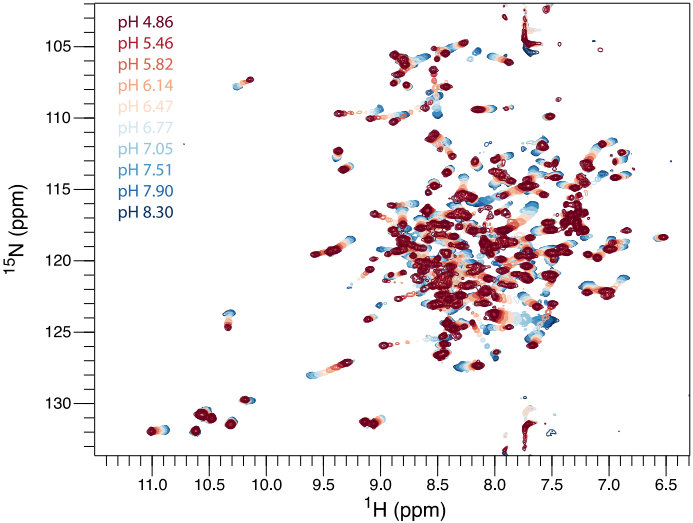

Figure S3 – *TROSY-HSQC pH titration of TPP⁺-bound EmrE at 25 °C.* All the peaks change position with pH, reflecting fast proton on-/off- rates and the global conformational change that occurs upon protonation of E14 in EmrE, as we have previously shown at 45 °C (14).

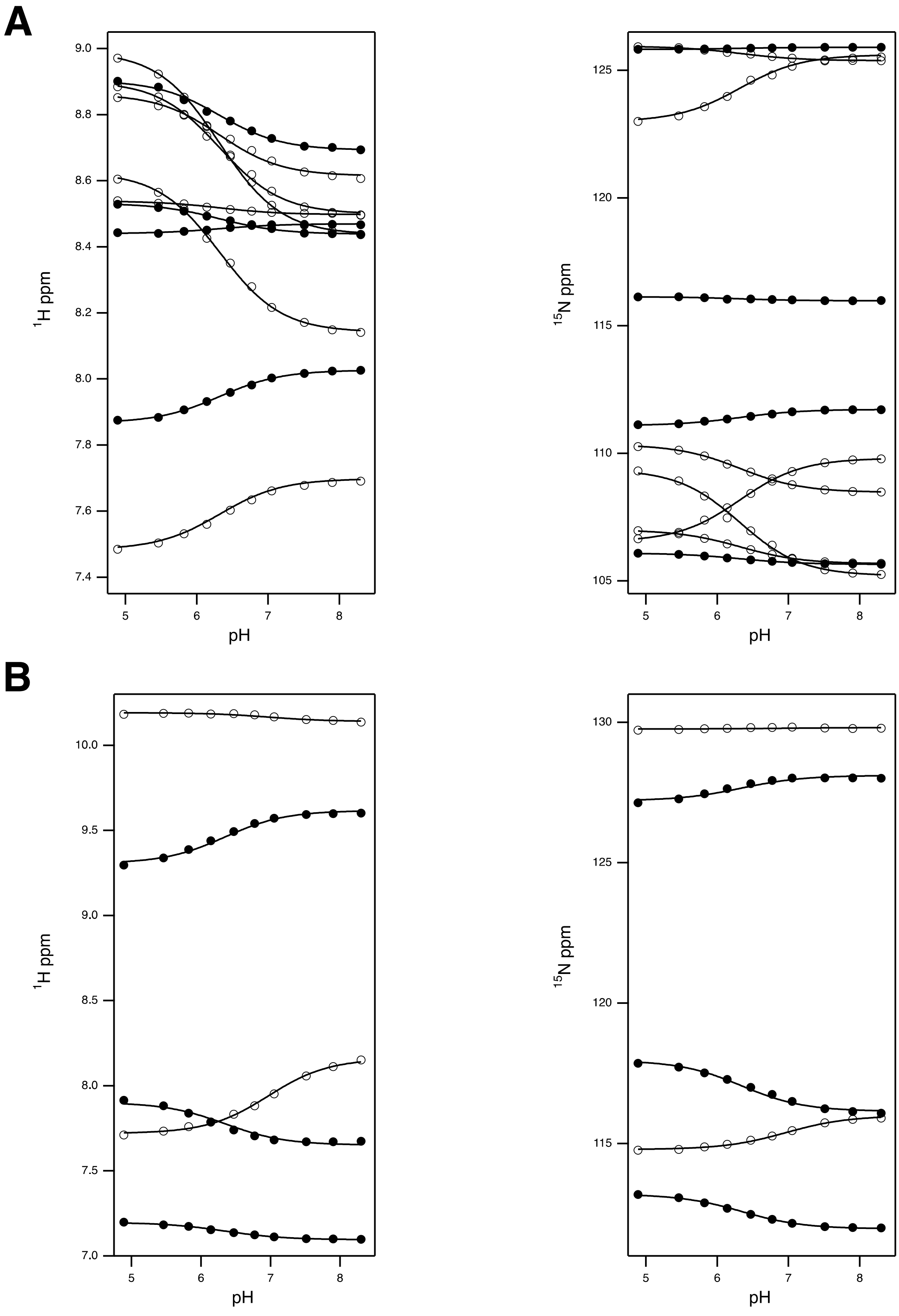

Figure S4 – *pKa determination for residues in TPP$^+$-bound EmrE from TROSY-HSQC NMR pH titration.* A) Global fit of backbone amide chemical shifts for selected core residues for monomer A (filled symbols, residues 10, 17, 44, 65) or monomer B (open symbols, residues 8, 9, 10, 17, 65, 69). B) Global fit of C-terminal tail residues in monomer A (filled symbols, residues 105,106,108) fit to a pK$_a$ of 6.3, indicating strong coupling to the binding site, while chemical shifts of C-terminal tail residues in monomer B (open symbols, residues 106,108) fit to a pK$_a$ of 7.0, matching the pK$_a$ of the imidazole sidechain of H110^B.

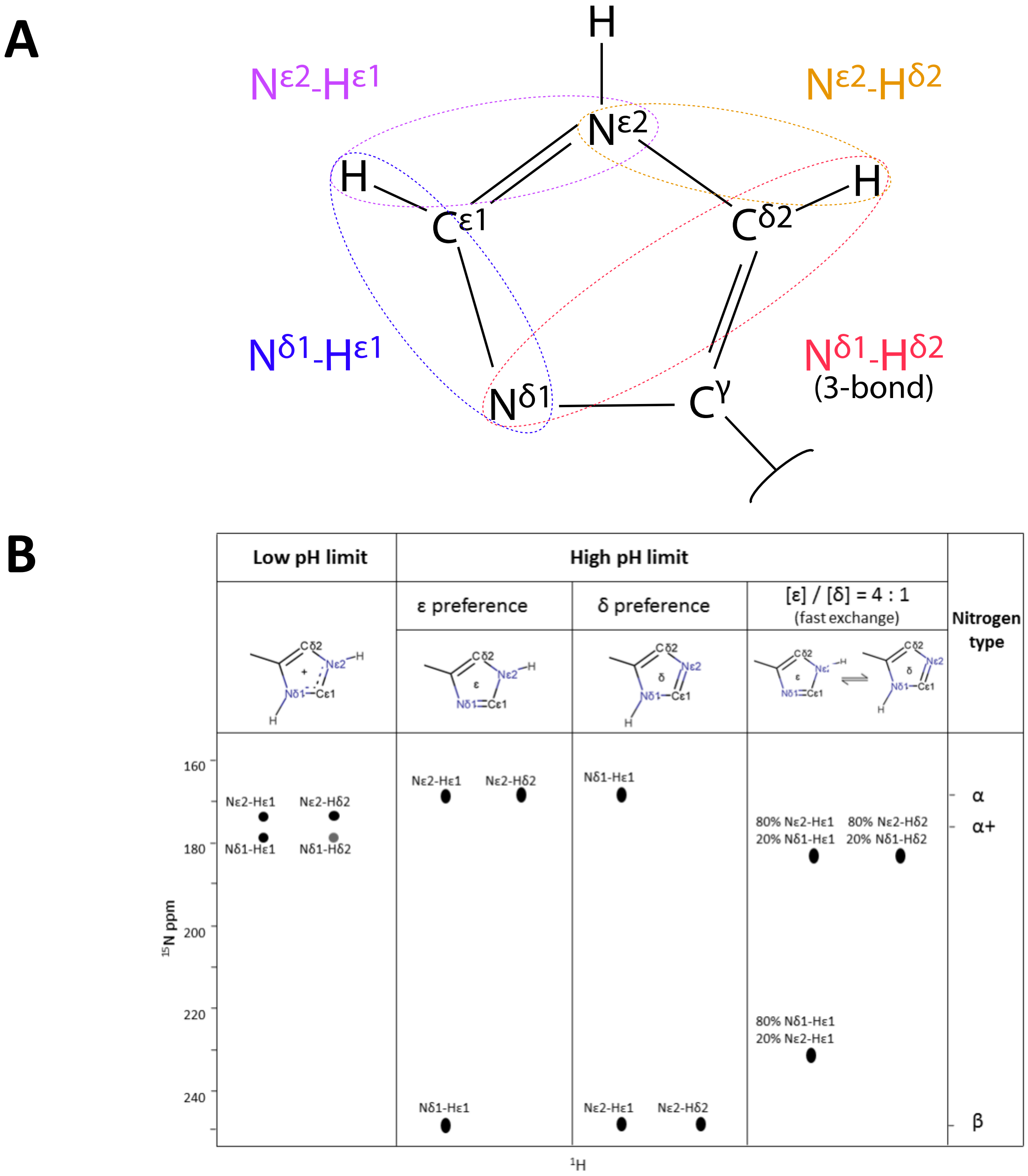

Figure S5 – *Histidine side chain correlations in HMBC spectra.* In this two-dimensional NMR experiment, magnetization is transferred from the $H^{\delta2}$ and $H^{\varepsilon1}$ protons to the $N^{\delta1}$ and $N^{\varepsilon2}$ on the imidazole side chains. This leads to the appearance of four peaks in the HMBC spectra of cationic histidine, representing the correlation of both protons to both nitrogens. The three-bond $N^{\delta1}$–$H^{\delta2}$ correlation (red) weakens for neutral histidine, and often disappears from the spectrum at high pH. The relative position of the three visible peaks at high pH can be used to calculate the tautomer populations of the side chain (4:1 ε/δ for free imidazole).

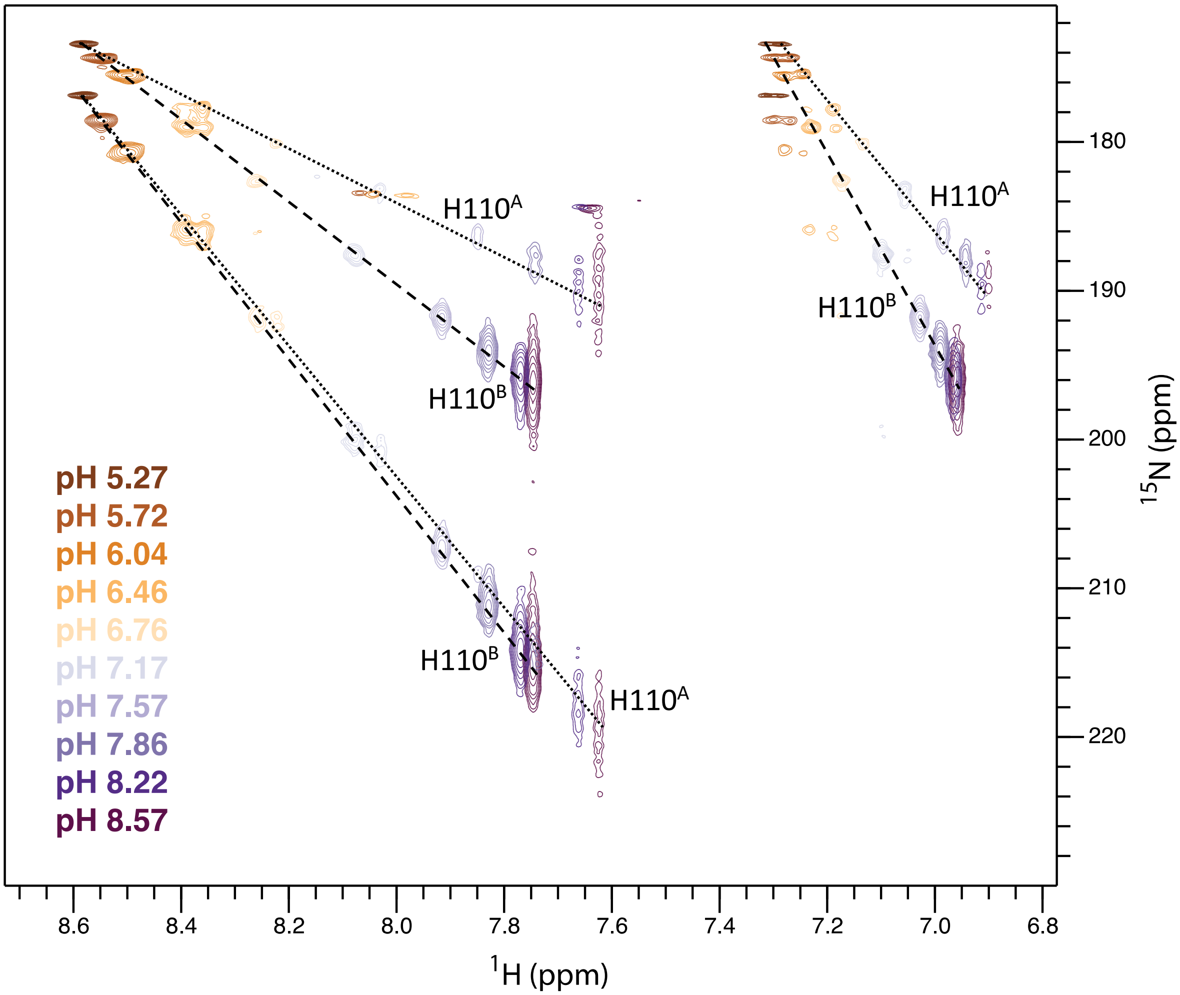

Figure S6 – *Full HMBC NMR-monitored pH titration of H110 for drug-free EmrE.* Spectra of 0.8 mM EmrE in isotropic bicelles at 25°C.

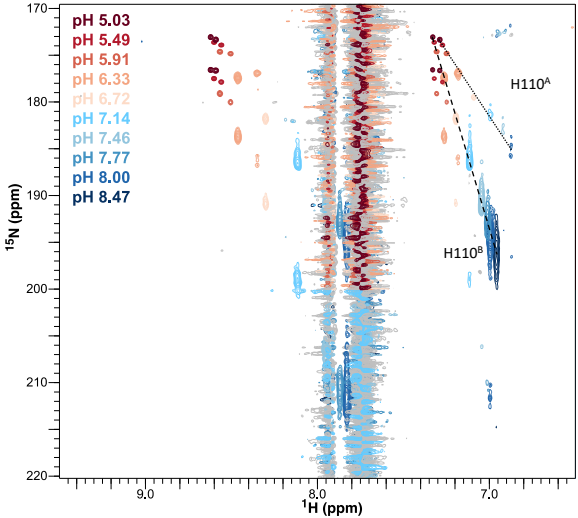

Figure S7 – *HMBC NMR-monitored pH titration of H110 for TPP$^+$-bound EmrE.* 0.8 mM EmrE in isotropic bicelles saturated with 5 mM TPP$^+$ at 25°C. Peaks from monomer A and monomer B are resolved at low and neutral pH. Peaks corresponding to H110^A are greatly broadened in the nitrogen dimension at high pH but can be seen in proton 1D slices (data not shown). Streaking at 7.7 and 7.9 ppm in the proton dimension are due to the large amounts of TPP$^+$ in the sample.

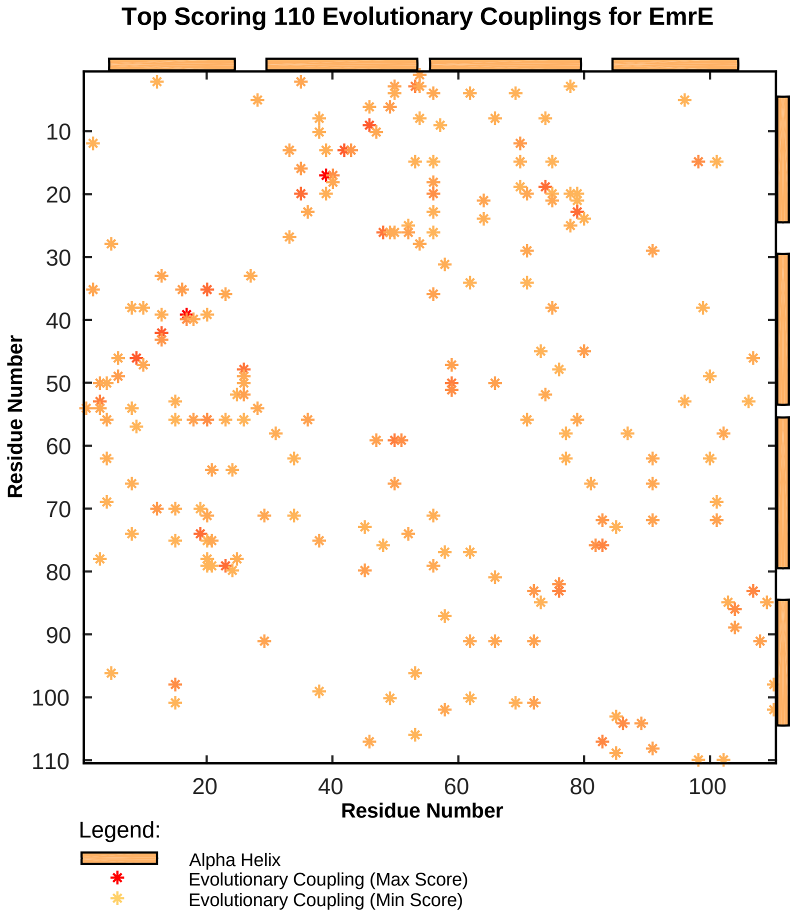

Figure S8 – *Evolutionary couplings for EmrE*. H110 is highly conserved, suggesting it may serve some functional role. To further probe the possibility that the C-terminal tail is coupled to the drug binding pocket, we constructed an alignment of 1408 proteins from the SMR family, each with less than 50% identity with EmrE, and generated an evolutionary coupling map using EVfold. To retrieve EmrE homologues, seven-iterations of PSI-BLAST were performed on June 12, 2018 using the BLOSUM45 scoring matrix (6). For the first five iterations, 1000 sequences were retrieved, while 4000 sequences were retrieved for the final two iterations. In between each iteration, sequences with greater than 90% identity to EmrE were removed from the Position Specific Score Matrix (PSSM). Sequences with less than 50% identity were aligned using Clustal Omega (15). Residue positions not found in EmrE were removed from the alignment, and the final alignment of 1408 sequences were uploaded to the EVFold server to perform the evolutionary coupling analysis (16, 17).

The map revealed many couplings that were consistent with the known structure and function of EmrE, such as extensive couplings from TM1 to TM2 and TM3 lining the drug-binding pocket of EmrE. However, the pattern of the couplings is less well defined than for many other helical membrane proteins, possibly because of the highly dynamic nature of EmrE. An inherent drawback with evolutionary coupling analysis is that highly conserved residues that are important for function lack sufficient variation for robust analysis and thus display fewer couplings than moderately conserved residues. H110 is highly conserved, so it only displays two weak couplings, both to TM4. Nevertheless, other residues in the C-terminal tail display strong couplings to several parts of the protein, including TM2 and TM3. Of particular note are couplings from the C-terminal tail to the TM3-4 loop, suggesting a conserved interaction between these regions.

Table S1. ITC Experimental Data

pH	Buffer	Buffer ΔH (kJ/mol)*	[EmrE] (μM)#	[TPP^+] (μM)	ΔH (kJ/mol)	*n*#	K_d^{obs} (nM)	Blank (μJ)
5.5	Cacodylate	-4.72	660	2810	-39 ± 1	0.55	67000	-5.8
5.5	Cacodylate	-4.72	840	4890	-36 ± 1	0.53	78000	-3.8
5.5	MES	14.90	810	4560	-48 ± 2	0.56	80000	-10.5
5.5	MES	14.90	800	4610	-48 ± 2	0.57	73000	-13.7
5.5	Piperazine	32.83	800	4870	-61 ± 2	0.57	58000	-12.5
5.5	Piperazine	32.83	800	4980	-61 ± 2	0.56	62000	-12.6
5.5	Average:					0.56 ± 0.02	70000 ± 9000	
6.5	BES	24.21	84	233	-61 ± 4	0.52	2100	-7.2
6.5	BES	24.21	93	392	-56.4 ± 0.9	0.55	1900	-3.4
6.5	BES	24.21	94	392	-56 ± 1	0.57	1900	-3.3
6.5	Imidazole	36.46	111	254	-60 ± 4	0.48	1700	-12.8
6.5	Imidazole	36.46	92	389	-62 ± 1	0.53	2100	-2.7
6.5	Imidazole	36.46	91	384	-65 ± 2	0.55	2300	-4.1
6.5	KPi	-1.00	93	308	-28.5 ± 0.8	0.57	1800	-1.7
6.5	KPi	-1.00	57	192	-29 ± 1	0.53	1300	-2.6
6.5	KPi	-1.00	94	394	-30.8 ± 0.7	0.54	1900	-1.3
6.5	KPi	-1.00	95	404	-30 ± 1	0.57	1900	-2.7
6.5	Average:					0.54 ± 0.03	1900 ± 300	
7.5	MOPS	21.60	30	133	-40 ± 4	0.42	130	-2.0
7.5	MOPS	21.60	30	117	-40 ± 3	0.45	180	-2.0
7.5	MOPS	21.60	29	117	-40 ± 5	0.48	150	-2.0
7.5	MOPS	21.60	29	119	-42 ± 3	0.49	160	-0.3
7.55	BES	24.21	37	149	-41 ± 2	0.53	210	-2.5
7.55	BES	24.21	37	130	-51 ± 2	0.45	280	-7.6
7.55	Imidazole	36.46	33	141	-47.3 ± 0.7	0.59	270	-1.4
7.55	Imidazole	36.46	34	147	-48.6 ± 0.6	0.57	260	-1.9
7.55	KPi	-1.00	43	307	-15.7 ± 0.6	0.61	120	-1.6
7.55	KPi	-1.00	61	370	-18.7 ± 0.2	0.57	150	-1.1
7.55	Tris	46.27	31	96	-62 ± 2	0.52	250	-0.9
7.55	Tris	46.27	30	97	-63 ± 2	0.48	200	-1.2
7.5	Average:					0.51 ± 0.06	200 ± 60	
8.5	Bicine	26.34	35	141	-31.6 ± 0.8	0.56	40	-2.7
8.5	Bicine	26.34	36	154	-32 ± 1	0.54	40	-2.5
8.5	Bicine	26.34	37	124	-29.3 ± 0.8	0.47	20	0.9
8.5	Bicine	26.34	38	124	-29 ± 1	0.47	20	0.8
8.5	Glycylglycine	43.08	32	124	-38.7 ± 0.7	0.44	30	-0.2
8.5	Glycylglycine	43.08	32	124	-35.1 ± 0.9	0.45	30	-0.1
8.5	Glycylglycine	43.08	30	121	-35.4 ± 0.7	0.40	20	0.0
8.5	Glycylglycine	43.08	32	128	-38 ± 1	0.47	30	-0.9
8.5	Tris	46.27	30	122	-38 ± 2	0.42	50	-2.6
8.5	Tris	46.27	31	133	-43.5 ± 0.7	0.42	50	-1.3
8.5	Average:					0.46 ± 0.05	30 ± 10	

*Buffer data taken from (18) and adjusted for temperature as described in the methods.

#[EmrE] listed is the monomer concentration, *n* is therefore the stoichiometry of TPP^+ bound per EmrE monomer, confirming that 1 TPP^+ binds per dimer.

References

1. Morrison, E. A., and Henzler-Wildman, K. A. (2014) Transported substrate determines exchange rate in the multidrug resistance transporter EmrE. *J. Biol. Chem.* **289**, 6825–36
2. Berezin, C., Glaser, F., Rosenberg, J., Paz, I., Pupko, T., Fariselli, P., Casadio, R., and Ben-Tal, N. (2004) ConSeq: the identification of functionally and structurally important residues in protein sequences. *Bioinformatics*. **20**, 1322–1324
3. Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533
4. Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T., and Ben-Tal, N. (2013) ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Isr. J. Chem.* **53**, 199–206
5. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350
6. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–402
7. Eddy, S. R. (2009) A new generation of homology search tools based on probabilistic inference. *Genome Inform.* **23**, 205–11
8. Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518
9. Kermani, A. A., Macdonald, C. B., Gundepudi, R., and Stockbridge, R. B. (2018) Guanidinium export is the primal function of SMR family transporters. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3060–3065
10. Mordoch, S. S., Granot, D., Lebendiker, M., and Schuldiner, S. (1999) Scanning cysteine accessibility of EmrE, an H+-coupled multidrug transporter from Escherichia coli, reveals a hydrophobic pathway for solutes. *J. Biol. Chem.* **274**, 19480–6
11. Rotem, D., Steiner-Mordoch, S., and Schuldiner, S. (2006) Identification of tyrosine residues critical for the function of an ion-coupled multidrug transporter. *J. Biol. Chem.* **281**, 18715–22
12. Wang, J., Rath, A., and Deber, C. M. (2014) Functional response of the small multidrug resistance protein EmrE to mutations in transmembrane helix 2. *FEBS Lett.* **588**, 3720–3725
13. Sharoni, M., Steiner-Mordoch, S., and Schuldiner, S. (2005) Exploring the Binding Domain of EmrE, the Smallest Multidrug Transporter. *J. Biol. Chem.* **280**, 32849–32855
14. Robinson, A. E., Thomas, N. E., Morrison, E. A., Balthazor, B. M., and Henzler-Wildman, K. A. (2017) New free-exchange model of EmrE transport. *Proc. Natl. Acad. Sci.* 10.1073/pnas.1708671114
15. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
16. Marks, D. S., Colwell, L. J., Sheridan, R., Hopf, T. A., Pagnani, A., Zecchina, R., and Sander, C. (2011) Protein 3D structure computed from evolutionary sequence variation. *PLoS One*. **6**, e28766
17. Ekeberg, M., Lövkvist, C., Lan, Y., Weigt, M., and Aurell, E. (2013) Improved contact prediction in proteins: Using pseudolikelihoods to infer Potts models. *Phys. Rev. E*. **87**, 12707
18. Goldberg, R. N., Kishore, N., and Lennen, R. M. (2002) Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data*. **31**, 231–370