PI: **Karplus, Martin**	Title: Modeling atomic structure of the EmrE multidrug pump to design inhibitor peptides	
Received: 06/14/2013	FOA: PA11-262	Council: 01/2014
Competition ID: ADOBE-FORMS-B2	FOA Title: NIH SMALL RESEARCH GRANT PROGRAM (PARENT R03)	
1 R03 AI111416-01	Dual:	Accession Number: 3599020
IPF: 3212901	Organization: HARVARD UNIVERSITY	
Former Number:	Department: Chemistry and Chemical Biology	
IRG/SRG: ZRG1 MSFD-N (08)F	AIDS: N	Expedited: N
Subtotal Direct Costs (excludes consortium F&A) Year 1: ▉ Year 2: ▉	Animals: N Humans: N Clinical Trial: N Current HS Code: 10 HESC: N	New Investigator: N Early Stage Investigator: N

Senior/Key Personnel:	Organization:	Role Category:
Martin Karplus Ph.D.	Harvard University	PD/PI
Victor Ovchinnikov Ph.D	Harvard University	Co-Investigator

Specific Aims

Many aromatic compounds are used as antibiotic, antiseptic, and antineoplastic agents in the control of pathogens. In turn, bacteria have evolved complex mechanisms for active expulsion of such compounds from their cytoplasm, which leads to the phenomenon of drug resistance. The increasing emergence of organisms resistant to even the most powerful antibiotics is among the most serious problems for the treatment of infectious diseases in the developing world, such as tuberculosis and malaria. The rapid rise of carbapenem-resistant Enterobacteriacea (CRE) such as *Klebsiella* has recently prompted actions from the CDC and the HHS Department.[1] [2]

The *EmrE* transporter in *E. coli*, a member of the widespread small multidrug resistance (SMR) family of membrane-bound transporters, actively pumps toxic compounds out of the cytoplasm and thus contributes to the drug resistance. The aims of this proposal are (1) to develop an atomic-level model of the *EmrE* transporter and (2) to design peptide analogs that inhibit the formation of the active dimer in the bacterial membrane. The approach has the potential to enhance current antibacterial therapies for a wide range of pathogens, and also to test the applicability of synthetic peptides to target specific oligomerization motifs in membrane proteins.

Aim 1. **To construct and validate an atomic-level model of the *EmrE* membrane protein.** *EmrE* is a small antisymmetric homodimer (110 residues per monomer) with eight transmembrane helices whose low-resolution C_α-only structure is known. The C_α structures will be used with available modeling protocols will be used to construct a series of all-atom models. To improve the models, they will be subjected to molecular dynamics simulations with replica exchange in the presence of an explicit lipid membrane and water molecule environment. The free energy of the refined structural candidates will be computed with the confinement method, and the ones of the lowest free energy will be validated by comparison with experimentally determined ligand dissociation constants and residue pKa values.

Aim 2. **To design peptide analogues that inactivate the *EmrE* transporter by preventing dimerization.** Because three transmembrane helices from each monomer (TM1, TM2, and TM3) form the drug-binding site, and the remaining two helices (TM4) make contact in the low-resolution structure, peptide analogs will be designed to mimic TM4 as dimerization inhibitors resistant to proteolysis. The initial inhibitor model will be identical to TM4, and the starting structures of the *EmrE* monomer-inhibitor complex will be taken from the all-atom dimer determined in *Aim 1*. The peptide sequence will be optimized on the basis of the interaction energies between the *EmrE* monomer and the peptide. Proteolysis-resistant peptide analogs will be modeled to have similar shape and charge distribution as the highest-affinity peptides. In particular, all-hydrocarbon stapled α-helical peptides, whose synthesis has been worked out, will be used to replace the natural peptide. Optimum sidechain distributions that lead to strong binding interaction with the *EmrE* monomer model will be determined iteratively by MD simulations. The best candidate peptide analogs will be synthesized and tested experimentally.

[1] The Centers for Disease Control and Prevention. CDC: Action needed now to halt spread deadly bacteria[Press Release], 2013. URL http://www.cdc.gov/media/releases/2013/p0305_deadly_bacteria.html.
[2] The Department of Health and Human Services is awarding GlaxoSmithKline up to $200 million to develop new antibiotics. URL http://www.gsk.com/media/press-releases/2013/glaxosmithkline-awarded-up-to–200-million-by-u-s–government-to.html
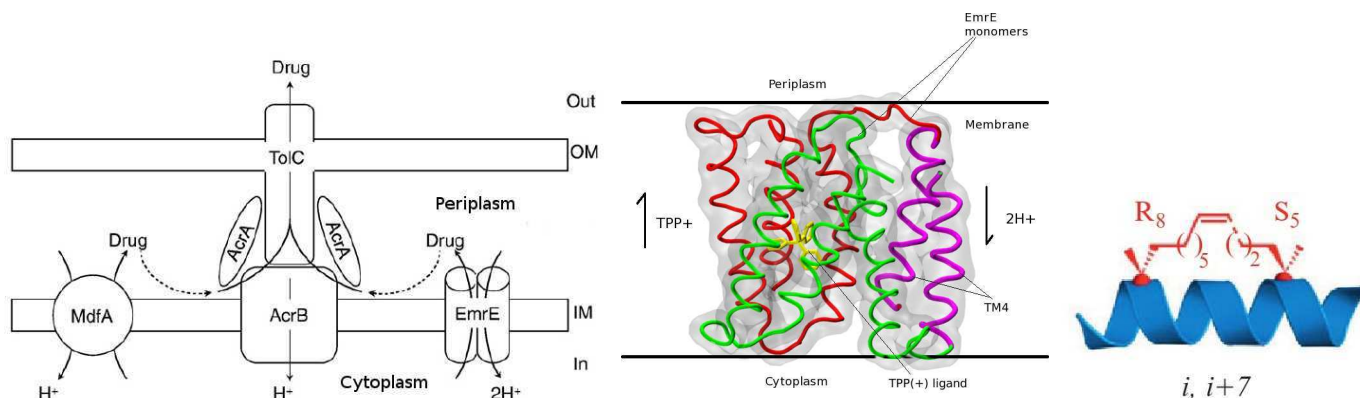
Figure 1: <u>Left panel</u>: Three multidrug transporters from *E. coli* coordinate efflux of drugs from the cytoplasm into the medium (adapted from Ref. 1). <u>Middle panel</u>: Low-resolution Xray crystal structure of ligand-bound dimeric *EmrE*[2]. The transport of aromatic drugs such as TPP$^+$ is coupled to the influx of two protons.[3] Transmembrane helices 4 (TM4), which appear to be important for dimerization,[4,5] are shown in magenta. <u>Right panel</u>: A hydrocarbon-stapled α-helical peptide (adapted from Ref. 6). The hydrocarbon staple is in red; in the notation X_n, X refers to the stereochemistry at the α-carbon and n gives the length of the alkenyl chain.

Research Strategy

Significance

Antimicrobial multidrug resistance poses serious challenges in the treatment of many infectious diseases. A ubiquitous mechanism by which multidrug resistance is conferred to bacteria involves the active efflux of a broad variety of cytotoxic compounds by transmembrane multidrug resistance proteins.[7] This efflux is driven either by ATP hydrolysis[8] (*e.g.* in the ABC-cassette superfamily of proteins), or by the electrochemical transmembrane potential.[9] The multidrug transport complex AcrAB/TolC in *E. coli* provides a major drug extrusion pathway from the periplasm (Fig. 1). There is strong evidence, however, that the AcrAB/TolC system cannot confer significant resistance to compounds such as ethidium and acriflavine without single-component 'suppliers' (*e.g.* MdfA or *EmrE*) that move aromatic compounds into the periplasm from the cytoplasm.[1] The various suppliers have overlapping specificities, and collectively form a multidrug resistance network of proteins that renders bacteria effectively immune to a broad range of compounds.[10–12] The small multidrug resistance (SMR) family of membrane-bound transporters is ubiquitous in bacteria, including the pathogens *M. Tuberculosis, P. Aeruginosa, B. Pertussis, N. Meningitis, B. Anthracis, S. aureus* among others,[9,13,14] but continues to receive less attention than other transporter families (*e.g.* ABC cassette proteins). Like many bacterial multidrug transporters (*e.g.* AcrB and MdfA), SMR proteins use the electrochemical potential of proton influx to actively pump toxic compounds from the cytoplasm. SMR pumps confer resistance to a wide variety of quarternary ammonium compounds (QACs),[15] some of which are used as disinfectants (*e.g.* benzalkonium chloride), and also promote resistance to antibiotics such as ampicillin, erythromycin and tetracycline.[16,17] SMR genes are often found on plasmids and transposons that contain other genes responsible for antibiotic resistance (to, *e.g.*, β-lactams, cephalosporins and dihydrofolate inhibitors,[13,18]) implying that the most dangerous antibiotic-resistant bacteria also tend to be resistant to common disinfectants. Although the *E. coli EmrE* multidrug transporter from the SMR family has been the subject of experimental studies (see Ref. 1 for a review) atomic-level structural information that is sufficient for a rational design of SMR inhibitors is not available. Such structures have been very difficult to obtain by experimental methods alone because of the generally high mobility of SMR proteins inside the lipid membrane. Recent studies of *EmrE* have produced low-resolution structures, which show that the 110 residue *EmrE* protein is a loose four-helix bundle with an active site formed by the active dimer.[2,19,20] On the basis of a low-resolution Xray structure,[2] it was proposed that certain synthetic peptides could disrupt *EmrE* dimerization, inhibiting its activity.[5] Although a transmembrane helical fragment of *EmrE* was indeed found to reduce the rate of drug efflux, the reduction was moderate (60%) and disappeared after $\sim$20 min due

to peptide degradation (an effective *D*-peptide resistant to proteolysis was not reported in the study). In the proposed research, molecular simulation tools will be used to carry out rational *in silico* design of SMR inhibitor peptides. In the first stage, low resolution X-ray crystal structures of the *EmrE* transporter,[2] data from cryo-electron microscopy (EM)[19,20] and from NMR solution experiments,[21] will be used together with recent developments in physics-based computer modeling of protein[22,23] and membrane thermodynamics[24–26] to construct and validate an atomic-resolution model of dimeric *EmrE* inside the lipid membrane and in complex with various drugs. Although a computer model of *EmrE* was published in 2004,[27] it was found to be in conflict with subsequent EM, Xray and NMR data.[2,21,28] In the second stage, the dimeric structures will be used to create starting models of *EmrE*-peptide inhibition complexes. The models will be optimized using directed *in silico* residue mutagenesis to generate peptides inhibitors with the highest affinity for *EmrE*. The optimized peptides will be stabilized to resist proteolysis using all-hydrocarbon *staples*,[6] and the optimal models will be tested experimentally. Stapled peptides have high α-helical propensities,[29] and have been used to disrupt protein-protein interactions important for the progression of cancer.[6,30–32] Because of the ubiquity of SMR proteins in bacteria, high-affinity peptide-analog inhibitors have the potential to be effective against a broad range of pathogens, especially in combination with antibiotic drugs. The structure of *EmrE* has similarities to the internal repeats in larger membrane proteins, such as aquaporins[33] and neurotransmitter transporters,[34] suggesting how the latter proteins may have arisen *via* gene duplication and mutation. The results of the proposed modeling study are therefore expected to be of general interest, and to be relevant for other membrane proteins as well.

Innovation

The novelty of the proposed work is that it brings to bear the state of the art protein and peptide modeling methodology, some of which was developed in the applicant's laboratory, to produce a high-resolution model of a small flexible membrane protein, the *EmrE* multidrug transporter, and to use it for the design of proteolytically stable peptides as inhibitors. For highly mobile small membrane proteins, experimental methods such as X-ray Crystallography and Nuclear Magnetic Resonance (NMR) have been unable to determine atomic-level structures. A rigorous computational methodology is proposed here to generate and analyze high-resolution structures of the *EmrE* transporter using as the starting point existing low-resolution Xray[2] structures and cryo-EM[19,20] models. Although various structure prediction and refinement programs have been used before to predict optimal protein confirmations,[35–39] ensembles of structures are needed to understand the function of membrane proteins[40,41] as well as other large biomolecules.[42] The use of extensive equilibration by Molecular Dynamics (MD) simulations of membrane proteins in a realistic lipid environment,[26] with *enhanced sampling* methods that preserve the canonical ensemble (REMD[43] or RLES[44]) is expected to produce high-quality sets of atomistic structures that exist in the membrane environment. Indeed, the increasing accuracy of MD simulations has recently been demonstrated with the *in silico* folding of small proteins.[45] Additionally, to the knowledge of this applicant, validation based on free energy simulations to rank the modeled structures, has not been performed previously.

Motivated by recent experiments,[4,5] we propose to design peptide inhibitors of *EmrE* dimerization. Although peptides and peptide analogs have been used previously to disrupt dimerization in membrane proteins (the *E. coli* aspartate receptor[46]) as well as to inhibit HIV cell entry[47] and to disrupt protein-protein interactions in tumor cells,[6,30–32] this is the first rigorous computational proposal to design peptide SMR inhibitors. Because SMR transporters are widely distributed in bacteria, and because they appear to have a conserved dimerization motif[5,13,48] such inhibitors promise to be effective against many pathogens, especially in combination with other therapies.

Although the computational approach outlined here is challenging and involves extensive computer simulations, recent advances in the efficiency of molecular dynamics (MD) simulation of proteins (*e.g.* using Graphical Processing Units [GPUs][49–51]) and in protein structure prediction methods[37–39] make computational modeling an important tool to supplement experimental studies. This applicant has contributed to the foundation of MD simulations in biology, has decades of experience with the computational methods to be used in the study, and continues to oversee the development of novel simulation methodologies.[22,23,52,53] Thus, the proposed study, in spite of its novelty, has an excellent chance of success in the two-year time period. Moreover, the inhibitor

peptide analogs designed *in silico* will be synthesized and tested by experimental collaborators. In particular, the laboratories of Dr. Gregory Verdine (Harvard University) and Dr. Charles Deber (University of Toronto) have agreed to synthesize the designed peptides and to carry-out efflux inhibition assays (letters of support attached).

Approach

This section is organized to address sequentially the specific objectives of the research proposal. As indicated in the timeline (provided at the end), the approach is based on the expectation that the design of several peptide inhibitor candidates will be completed in less than two years. At least some of these candidates will be synthesized and tested within the two-year period by our experimentalist collaborators. Computational modeling software programs will be used to create and validate the atomic resolution models of the *EmrE* transporter protein. Many of these programs were developed by the applicant and his co-workers and former students. In particular, the CHARMM program,[52,54] a workhorse for the development and application of many biomolecular simulation methods, is under continuous development in the applicant's laboratory.

Aim 1. Preparation, refinement and validation of high-resolution structures

The modeling strategy proposed for this *Aim* is based on the emerging realization[55–57] that traditional structure prediction methods for proteins can be combined with limited experimental information and with new molecular simulation methods to produce high-quality atomic structures. In the planned research, we will make use of low-resolution Xray structures that do not include coordinates of the residue sidechains.[2] Trial sets of sidechain coordinates can be generated with several promising algorithms,[37–39] and further refined by molecular dynamics simulations using restraints to the electron density map.[56,58,59] The best structures obtained from this procedure will be taken as the starting point for free-energy calculations of substrate binding, and for MD simulations of the conformational change associated with drug transport.

1*a*. **Computer side-chain modeling and refinement.**

I). Two low-resolution ∼4Å structure of dimeric *EmrE* complexed with Tetraphenylphosphonium (TPP+) are available from the Protein Data Bank (PDB accession codes 3B5D and 3B62).[2] Only C_α carbon atoms are provided at the low resolution. The remaining structure with PDB code 3B61 in the absence of ligand will not be used because the packing of transmembrane helices appears unrealistic. Starting from the C_α coordinates, coordinates of the protein backbone will be generated using the program CHARMM.[52,54] Since most of the structure is α-helical (as expected for membrane proteins), the backbone coordinates will be generated by specifying the ϕ, ψ, and ω dihedral angles, which can be done using the internal coordinate facility in CHARMM. Additional restraints on the ϕ and ψ dihedral angles will be applied as needed to keep the corresponding values within the allowed Ramachandran region. The resulting backbone-only models will be energy-minimized using a specially adapted CHARMM energy function.[60]

II). The backbone-only models will be the starting point for generating coordinates of the residue sidechains. Several algorithms that have been developed for sidechain modeling will be used to generate initial trial models of complete structures. These are SCWRL,[37] OPUS-Rota,[61] Self-Consistent Mean Field theory (SCMF),[35,62] and ROSETTA[39]. We note that modeling tools that rely on homology modeling cannot be used because there are no appropriate homologs of *EmrE* with known atomic structures. The modeled structures will be energy-minimized with the program CHARMM using an implicit membrane model to simulate the effects of a lipid environment.[24] Any chirality errors that arise in the modeling and minimization stages will be corrected at this stage.

1*b*. **Molecular dynamics simulations for refinement of structures.**

I). The modeled structures obtained in 1*a* will be inserted in pre-equilibrated patches of DPMC lipid bilayers using the CHARMM-gui interface.[63] Although other membrane compositions may be used, *EmrE* is

known to be functional in DPMC membranes.[1] Explicit water molecules and ions will be added to the protein-membrane complex to mimic the cytoplasm and the periplasm.

II). MD simulations to refine the structures will be carried out according to the protocols listed below.
(i) Equilibrium MD simulations will be carried out with the program AceMD[50] using Graphical Processing Units (GPUs). The anticipated size of the simulation systems is $\sim$50,000 atoms, and the expected speed of the MD simulation is 30 nanoseconds per day. Several hundred nanoseconds will probably be required to reach steady-state behaviour of the simulation system.
(ii) Temperature replica exchange MD simulations[43] will be carried out using the program NAMD.[64] (NAMD is significantly faster than CHARMM for straightforward MD but more limited in terms of simulation options.) Simulations will be conducted at several temperature values ranging from 300K (physiological) to about 600K. Only the physiological (300K) simulations will be used for structure evaluation; the higher-temperature simulation exchange simulation coordinates with the 300K systems, so as to accelerate the effective rate and search capabilities of the MD simulation.
(iii) Locally Enhanced Sampling (LES)[65,66] simulations will be performed with the program CHARMM using implicit solvent representation for the membrane and cytoplasm[24], and with the program NAMD with explicit solvent. The use of variable restraints allows the transformation of the artificial replicated LES ensemble into the correct ensemble at the physiological temperature.[44]

1c. **Evaluation of refined and equilibrated structures.** After the structures simulated by MD in the previous step attain steady state (*i.e.* when the total energy and coordinates are fluctuating randomly in the vicinity of the corresponding averages in response to the thermostat), the following procedure will be used to determine the optimal structures.
(i) The coordinates will be checked to determine whether they fit into the experimental electron density (EM) map.[28] In the case that the coordinates do not fit completely, additional MD simulations with restraints to the EM map will be performed.[56,58,59] Final structures whose coordinates are consistent with the EM map will be retained for further analysis.
(ii) The structures will be removed from the explicit lipid membrane environment and placed into an implicit membrane model[24,25] in CHARMM. The removal of explicit lipid and water molecules simplifies the separation of the free energy of protein solvation from the intramolecular free energy of the protein. The solvation free energy (FE) of each structure is provided by the implicit solvent model, and the configurational free energy of the protein structure will be estimated from a confinement analysis.[53]
(iii) In case that several distinct structures have very similar FE values, all of them will be retained for further validation and analysis (1*d* and 1*e* below).

1d **Free energy simulations of ligand binding to** *EmrE***.** The mechanism by which *EmrE* confers drug resistance involves active extrusion of aromatic drug compounds from the *E. coli* cytoplasm in exchange for two protons.[1] At the atomic level, the active site *EmrE* dimer alternately binds two protons and a drug molecule (*e.g.* TPP+, or Ethidium). To achieve drug efflux from the cytoplasm, the affinities of *EmrE* for the drug molecules (reported in experiments as dissociation [K_D] constants[2,67,68]) have to be tuned to the cross-membrane pH gradient. To establish the physiological relevance of the structures modeled in step 1, we will use free energy simulation to compute the FE of ligand binding to *EmrE*.

I). Coordinates of the ligands TPP+[2] and ethidium[69] will be submitted to CHARMM-CGENFF server[70] to generate force field parameters, which are necessary to perform energy calculations. The server produces trial parameters automatically which can be improved iteratively using alternating quantum and molecular mechanical calculations. This is a standard procedure in computational chemistry, and will be carried out according to the CHARMM parametrization protocol.[60,70]

II). Using the ligand force field parameters obtained in step *I* and the binding pose of the TPP+ ligand found in the low-resolution structure[2], the TPP+ and ethidium ligands will be reversibly "annihilated" using alchemical FE simulation.[71–74] The FE calculations will be performed with the MD program NAMD,[64] and the analysis and postprocessing of simulation data will be done using the Visual Molecular Dynamics

(VMD) software.[75] The alchemical annihilation simulations will be performed both in the active site of *EmrE*, and in the aqueous and lipid solvent environments to compute the FE *difference* of ligand binding.

1*e* **Determination of the pKa's of the amino acid residues involved in drug transport.** Mutagenesis and pH titration experiments have established that residue Glu14 is essential for proton-coupled transport of *EmrE* substrates.[76–78] Using Poisson-Boltzmann solvation theory,[79–81] we will estimate the free energy difference between protonated and deprotonated Glu14 of both monomers in the dimeric structure, and compare the value with that obtained from pKa values in the literature (7.3 – 8.5). In addition, the pKa of the Asp-14 mutant will be calculated and compared with the experimental value (5.6). For further validation, the more computationally expensive alchemical FE simulations will be performed to compute the pKa shifts inside the protein relative to the pKa in solution. Specifically, Glu-14 will be alchemically protonated both the protein and in the solvent; the FE difference between the two processes is proportional to the pKa shift. The pKa calculations are a particularly stringent test of the quality of the structure of the active site, because the large experimental deviation of the Glu-14 pKa inside the *EmrE* drug binding site (~ 8) from that in bulk solvent (~ 4.5) requires a correct and precise arrangement of the neighboring residues.

Aim 2. Design of peptides for the inhibition of dimerization

In the low-resolution C_α dimer structure of *EmrE*, transmembrane helices 4 (TM4) are in close proximity for direct interaction[2] (see Fig. 1). In this *Aim*, proteolysis-resistant peptide analogs will be designed to mimic TM4 (sequence [84]DLPAIIGMMLICAGVLIINLLS[105]) for use as dimerization inhibitors.

2*a* **Initial model of inhibition peptide.** *I).* Two *EmrE*-peptide models will be obtained from the optimal all-atom dimer model obtained in stage 1. In the first model, all parts of the first *EmrE* monomer except transmembrane helix 4 (TM4) will be deleted. In the second model, all parts of the second monomer except TM4 will be deleted. The two models are expected to be structurally close but not identical and will provide better simulation statistics than a single model. Partial deletion of one of the monomers will require re-equilibration of the lipid membrane around the monomer-peptide complex, which will be done by MD simulations, as described in 1*b* above.

II). An initial 100ns long MD simulation will performed using each of the starting structures to collect statistics of the interaction energy between the peptide and the monomer, and between the peptide and the lipid membrane. The MD calculations will be performed with the program AceMD[50] in parallel using a dedicated GPU for each calculation. An additional MD simulation of the single TM4 peptide embedded in the lipid bilayer will be performed for 100ns. This simulation will be used to normalize interaction energies, as described in the next step. The MD simulation trajectory with the lowest average interaction energy will be the starting point of the directed *in silico* mutagenesis optimization performed in step 2*b*.

2*b* **Optimization of peptide sequence.** The optimization protocol will use Monte Carlo sampling as follows. A peptide residue with a side chain atom within a cutoff distance d from the *EmrE* monomer will be se-lected at random. The distance d will be initially set to 3Å, but will be tuned to reduce the computational complexity of the problem by selecting the residues that are expected to make the largest contributions to the binding affinity (measured by the interaction energy). The selected residue will be mutated ran-domly to a nonpolar residue (*i.e.* selected from AFGILMPV) and the MD simulations of the *EmrE*-peptide complex and of the peptide alone embedded in the lipid will be repeated for 100ns. The average interac-tion energy between the mutated peptide and the *EmrE* monomer and membrane will be computed and compared to the corresponding interaction energy from the previous simulation. Because the binding of inhibition peptides occurs in the membrane, the interaction energy between the peptide and the lipid molecules must be taken into account for the free and the bound peptide. The change in the interaction energy between the peptide and the membrane will be computed from the MD trajectories of the mutated and the original peptides, both simulated in the membrane. The essential interaction energy differences

will be calculated based on the equation

$$\Delta E = E^{complex}_{mutated} - E^{complex}_{original} - (E^{peptide}_{mutated} - E^{peptide}_{original}),$$ (1)

where $E^{complex}$ denotes the average interaction energy between the peptide and the *EmrE* monomer and membrane, $E^{peptide}$ denotes the average interaction energy between the peptide and membrane in the (normalization) simulation without *EmrE*, and the subscripts *original* and *mutated* refer to the structure of the peptide before and after the trial mutation. The quantity ΔE will be used in a Metropolis Monte-Carlo (MC) acceptance-rejection test, in which the probability of accepting the trial mutation is computed as

$$P_{accept} = \begin{array}{ll} 1 & \text{if } \Delta E \le 0 \\ \exp(-\beta\Delta E) & \text{if } \Delta E > 0, \end{array}$$ (2)

where the quantity β (inverse temperature) will be varied to explore different peptide compositions, increasing from the physiological value at 300K $\beta=1/(k_BT) \simeq 1.7(\text{kcal/mol})^{-1}$; higher temperature ($T$) values will be used to increase the range of sampled mutations. If the Monte Carlo move is accepted, the new structure becomes the 'original' structure; otherwise the new structure is discarded. The above procedure is repeated until the inhibitor peptide composition does not change after a specified number of trial moves.

The above procedure uses the interaction energy differences instead of the free energy differences. However, the difference between the enthalpy and the free energy (the entropic contribution) to the binding free energy difference is expected to be small because the mutations are not likely to change significantly the flexibility of the binding interface and the calculations involve the differences between the bound and free helix. To quantify the error of neglecting the entropy contribution, we will compute the binding free energy and enthalpy difference resulting from a trial mutation use confinement analysis[53,82] with an implicit membrane model.[24,25]

2c **Construction of peptide analogs.** The *EmrE*-peptide complexes with the lowest interaction energies from step 2b will be the starting point for constructing proteolysis-resistant peptide analogs by adding all-hydrocarbon peptides 'staples' (see Fig. 1). Stapled peptides appear to be ideal candidates for the present problem because of their generally high α-helical propensity,[29,83] resistance to proteolysis,[6] and hydrophobicity of the staple, which is compatible with a high membrane permeability.[83] Successful design of a stapled peptide requires that the staple does not interfere with the intended contact interface between the peptide and the protein. For this reason, only those trial staple positions will be considered that are away from the binding interface. On the basis of the low-resolution experimental structure of *EmrE*,[2] the following positions in TM4, indicated in bold, appear to be well suited for the placement of a staple: [84]**D**LPAIIGM**M**LICAGVLIINLLS[105] (conserved residue positions are underlined). However, optimal positions will be determined from the model in step 2b. All residue pairs separated by 3, 4 and 7 positions will be considered for stapling (subject to the above selection constraint). The displacements correspond to one or two turns of the α-helix (see *e.g.* Fig. 1 for a two-turn staple). The sidechains in the chosen residues will be replaced with the appropriate hydrocarbon staple,[6,29] and the stapled peptide will be simulated by MD in complex with *EmrE* and also in isolation from *EmrE*, as described in step 2b. The stapled peptide candidates will be ranked on the basis of the interaction energy differences between the peptide and the protein, and peptide and lipid, as described in step 2b, as well as the α-helical propensity of the peptide.[29] The highest scoring stapled peptides will be retained for experimental characterization.

Project Timeline

In the first year, most of the effort will focus on the development of the atomic-resolution *EmrE* structure. We anticipate that approximately six months will be required to produce several refined structures (Aims 1a–1c). The remaining Aims 1c-1d (validation of the refined structures) are expected to require an additional four to six months. The remaining time (up to 1 year) will be devoted to the optimization of the peptide sequence and of the peptide staple position (Aim 2), which will involve iterative interactions with the experimental groups.

References

[1] Shimon Schuldiner. EmrE, a model for studying evolution and mechanism of ion-coupled transporters. Biochimica et biophysica acta, 1794(5):748–62, May 2009. ISSN 0006-3002. doi: 10.1016/j.bbapap.2008.12.018.

[2] Yen-Ju Chen, Owen Pornillos, Samantha Lieu, Che Ma, Andy P Chen, and Geoffrey Chang. X-ray structure of EmrE supports dual topology model. Proc. Natl. Acad. Sci. U. S. A., 104(48):18999–9004, November 2007. ISSN 1091-6490. doi: 10.1073/pnas.0709387104.

[3] D. Rotem and 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.

[4] Bradley E. Poulsen, Fiona Cunningham, Kate K. Y. Lee, and Charles M. Deber. Modulation of substrate efflux in bacterial small multidrug resistance proteins by mutations at the dimer interface. J. Bacteriol., 193: 5929–5935, 2011.

[5] Bradley E. Poulsen and Charles M. Deber. Drug efflux by a small multidrug resistance protein is inhibited by a transmembrane peptide. Antimicrob. Agents Chemother., 56:3911–3916, 2012.

[6] Gregory L Verdine and Gerard J Hilinski. Stapled peptides for intracellular drug targets., volume 503. Elsevier Inc., 1 edition, January 2012. ISBN 9780123969620. doi: 10.1016/B978-0-12-396962-0.00001-X.

[7] H. Nikaido. Multiple antibiotic resistance and efflux. Curr. Opin. Microbiol., 1:516–523, 1998.

[8] M.M. Gottesman, T. Fojo, and S.E. Bates. Multidrug resistance in cancer: role of ATP- dependent transporters. Nat. Rev. Cancer, 2:48–58, 2002.

[9] I.T. Paulsen, M.H. Brown, and R.A. Skurray. Proton-dependent multidrug efflux systems. Microbiol. Rev., 60:575–608, 1996.

[10] A. Lee, W. Mao, M.S. Warren, A. Mistry, K. Hoshino, R. Okumura, H. Ishida, and O. Lomovskaya. Interplay between efflux pumps may provide either additive or multiplicative effects on drug resistance. J. Bacteriol., 182:3142–3150, 2000.

[11] S. Yang, S.R. Clayton, and E.L. Zechiedrich. Relative contributions of the AcrAB, MdfA and NorE efflux pumps to quinolone resistance in Escherichia coli. J. Antimicrob. Chemother., 51:545–556, 2003.

[12] N. Tal and S. Schuldiner. A coordinated network of transporters with overlapping specificities provides a robust survival strategy. Proc. Natl. Acad. Sci. USA, 9051-9056:106, 2009.

[13] D.C. Bay, K.L. Rommens, and R.J. Turner. Small multidrug resistance proteins: a multidrug transporter family that continues to grow. Biochim. Biophys. Acta, 1778:1814–1838, 2008.

[14] Denice C. Bay and Raymond J. Turner. Diversity and evolution of the small multidrug resistance protein family. BMC Evol. Biol., 9:140, 2009. doi: 10.1186/1471-2148-9-140.

[15] 17. Identification and characterization of quaternary ammonium compound resistant staphylococci from the food industry. Int. J. Food Microbiol., 48:211–219, 1999.

[16] 73. The product of the qacC gene of Staphylococcus epidermidis CH mediates resistance to beta-lactam antibiotics in Gram-positive and Gram-negative bacteria. Res. Microbiol., 156:472–477, 2005.

[17] K. Nishino and A. Yamaguchi. Analysis of a complete library of putative drug transporter genes in Escherichia coli. J. Bacteriol., 183:5803–5812, 2001.

[18] 28. Genetic linkage between resistance to quaternary ammonium compounds and beta-lactam antibiotics in food-related Staphylococcus spp. Microb. Drug Resist., 7:363–371, 2001.

[19] Iban Ubarretxena-Belandia, Joyce M Baldwin, Shimon Schuldiner, and Christopher G Tate. Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. EMBO J., 22(23):6175–81, December 2003. ISSN 0261-4189. doi: 10.1093/emboj/cdg611.

[20] Sarel J Fleishman, Susan E Harrington, Angela Enosh, Dan Halperin, Christopher G Tate, and Nir Ben-Tal. Quasi-symmetry in the cryo-EM structure of EmrE provides the key to modeling its transmembrane domain. J. Mol. Biol., 364(1):54–67, November 2006. ISSN 0022-2836. doi: 10.1016/j.jmb.2006.08.072.

[21] E. A. Morrison, G. T. DeKoster, S. Dutta, R. Vafabakhsh, M.W. Clarkson, A. Bahl, D. Kern T. Ha, and K.A. Henzler-Wildman. Antiparallel EmrE exports drugs by exchanging between asymmetric structures. Nature, 481:45–52, 2012.

[22] V. Ovchinnikov, M. Karplus, and E. Vanden-Eijnden. Free energy of conformational transition paths in biomolecules: The string method and its application to myosin VI. J. Chem. Phys., 134:085103, 2011. PMC3060930.

[23] V. Ovchinnikov and M. Karplus. Analysis and elimination of a bias in targeted molecular dynamics simulations of conformational transitions: Application to Calmodulin. J. Phys. Chem. B, 116:8584–8603, 2012. doi: doi=10.1021/jp212634z. PMC3406239.

[24] S. Tanizaki and M. Feig. A generalized Born formalism for heterogeneous electric environments: Application to the implicit modeling of biological membranes. J. Chem. Phys., 122:124706, 2005.

[25] S. Tanizaki and M. Feig. Molecular dynamics simulations of large integral membrane proteins with an implicit membrane model. J. Phys. Chem., 110:548–556, 2006.

[26] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyev, A. D. MacKerell Jr., and R. W. Pastor. Update of the charmm all-atom additive force field for lipids: Validation on six lipid types. J. Phys. Chem. B, 114:7830–7843, 2010.

[27] K.E. Gottschalk, M. Soskine, S. Schuldiner, and H. Kessler. A structural model of EmrE, a multi-drug transporter from Escherichia coli. Biophys. J., 86:3335–3348, 2004.

[28] I. Ubarretxena-Belandia, J.M. Baldwin, S. Schuldiner, and C.G. Tate. Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. EMBO J., 22:6175–6181, 2003.

[29] Zuojun Guo, Udayan Mohanty, Justin Noehre, Tomi K Sawyer, Woody Sherman, and Goran Krilov. Probing the alpha-helical structural stability of stapled p53 peptides: molecular dynamics simulations and analysis. Chemical biology & drug design, 75(4):348–59, April 2010. ISSN 1747-0285. doi: 10.1111/j.1747-0285.2010.00951.x.

[30] Sohee Baek, Peter S Kutchukian, Gregory L Verdine, Robert Huber, Tad a Holak, Ki Won Lee, and Grzegorz M Popowicz. Structure of the stapled p53 peptide bound to Mdm2. Journal of the American Chemical Society, 134(1):103–6, January 2012. ISSN 1520-5126. doi: 10.1021/ja2090367.

[31] Avinash Muppidi, Kenichiro Doi, Selvakumar Edwardraja, Eric J Drake, Andrew M Gulick, Hong-Gang Wang, and Qing Lin. Rational design of proteolytically stable, cell-permeable peptide-based selective Mcl-1 inhibitors. Journal of the American Chemical Society, 134(36):14734–7, September 2012. ISSN 1520-5126. doi: 10.1021/ja306864v.

[32] Christopher J Brown, Soo T Quah, Janice Jong, Amanda M Goh, Poh C Chiam, Kian H Khoo, Meng L Choong, May a Lee, Larisa Yurlova, Kourosh Zolghadr, Thomas L Joseph, Chandra S Verma, and David P Lane. Stapled peptides with improved potency and specificity that activate p53. ACS chemical biology, 8(3):506–12, March 2013. ISSN 1554-8937. doi: 10.1021/cb3005148.

[33] K. Murata, K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J.B. Heymann, A. Engel, and Y. Fujiyoshi. Structural determinants of water permeation through aquaporin-1. Nature, 407:599–605, 2000.

[34] A. Yamashita, S.K. Singh, T. Kawate, Y. Jin, and E. Gouaux. Crystal structure of a bacterial homologue of Na+/Cl-dependent neurotransmitter transporters. Nature, 437:215–223, 2005.

[35] P. Koehl and M. Delarue. Application of a self-consistent mean field theory to predict protein side-chains conformation and estimate their conformational entropy. J. Mol. Biol., 239:249–275, 1994.

[36] A.T. Brünger, Paul D. Adams, G. Marius Clore, Warren L. DeLano, Piet Gros, Ralf, W. Grosse-Kunstleve, Jian-Sheng Jiang, John Kuszewski, Michael Nilges, Navraj S. Pannu, Randy J. Read, Luke M. Rice, Thomas Simonson, and Gregory L. Warren. Crystallography & NMR system: A new software suite for macro-molecular structure determination. Acta Cryst., D54:905–921, 1998.

[37] Georgii G. Krivov, Maxim V. Shapovalov, and Roland L. Dunbrack Jr. Improved prediction of protein side-chain conformations with scwrl4. Proteins, 77:778–795, 2009.

[38] Mingyang Lu, Athanasios D Dousis, and Jianpeng Ma. OPUS-Rota : A fast and accurate method for side-chain modeling. Protein Sci., 17:1576–1585, 2008. doi: 10.1110/ps.035022.108.likely.

[39] Kira M S Misura, Dylan Chivian, Carol a Rohl, David E Kim, and David Baker. Physically realistic homol-ogy models built with ROSETTA can be more accurate than their templates. Proc. Natl. Acad. Sci. USA, 103(14):5361–6, April 2006. ISSN 0027-8424. doi: 10.1073/pnas.0509355103.

[40] José D Faraldo-Gómez and Lucy R Forrest. Modeling and simulation of ion-coupled and ATP-driven membrane proteins. Curr. Opin. Struct. Biol., 21(2):173–9, April 2011. ISSN 1879-033X. doi: 10.1016/j.sbi.2011.01.013.

[41] Davide Provasi, Marta Camacho Artacho, Ana Negri, Juan Carlos Mobarec, and Marta Filizola. Ligand-induced modulation of the free-energy landscape of G protein-coupled receptors explored by adaptive biasing techniques. PLoS Comput Biol, 7(10):e1002193, 10 2011.

[42] K.K. Frederick. Conformational entropy in molecular recognition by proteins. Nature, 448:325–329, 2007.

[43] Y. Sugita and Y. Okamoto. Replica exchange molecular dynamics method for protein folding. Chem. Phys. Lett., 314:141–151, 1999.

[44] V. Ovchinnikov and M. Karplus. A restrained locally enhanced sampling method (RLES) for finding free energy minima in complex systems. in preparation, 2013.

[45] Kresten Lindorff-Larsen, Stefano Piana, Ron O. Dror, and David E. Shaw. How fast-folding proteins fold. Science, 334:517, 2011. doi: 10.1126/science.1208351.

[46] Neta Sal-Man, Doron Gerber, and Yechiel Shai. Hetero-assembly between all-L- and all-D-amino acid transmembrane domains: forces involved and implication for inactivation of membrane proteins. Journal of molecular biology, 344(3):855–64, November 2004. ISSN 0022-2836. doi: 10.1016/j.jmb.2004.09.066.

[47] D M Eckert, V N Malashkevich, L H Hong, P a Carr, and P S Kim. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. Cell, 99(1):103–15, October 1999. ISSN 0092-8674.

[48] Edda De Rossi, Manuela Branzoni, Rita Cantoni, Anna Milano, Giovanna Riccardi, and Orio Ciferri. mmr, a mycobacterium tuberculosis gene conferring resistance to small cationic dyes and inhibitors. J. Bacteriol., 188:6068–6071, 1998.

[49] J.A. Anderson, C.D. Lorenz, and A. Travesset. General purpose molecular dynamics simulations fully implemented on graphics processing units. J. Comput. Phys., 227:5342–5359, 2008.

[50] M. Harvey, G. Giupponi, and G. De Fabritis. Acemd: Accelerated molecular dynamics simulations in the microseconds timescale. J. Chem. Theory and Comput., 5:1632, 2009.

[51] M.S. Friedrichs, P. Eastman, V. Vaidyanathan, M. Houston, S. Legrand, A.L. Beberg, D.L. Ensign, C.M. Bruns, and V.S. Pande. Accelerating molecular dynamic simulation on graphics processing units. J. Comput. Chem., 30:864–872, 2009.

[52] B.R. Brooks, C.L. Brooks III, A.D. Mackerell Jr., L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, and et. al. CHARMM: The biomolecular simulation program. J. Comput. Chem., 30: 1545–1614, 2009. PMC2810661.

[53] V. Ovchinnikov, M. Cecchini, and M. Karplus. A Simplified Confinement Method (SCM) for Calculating Absolute Free Energies and Free Energy and Entropy Differences. J. Phys. Chem. B, 117:750–62, 2013. doi: doi=10.1021/jp3080578. PMC3569517.

[54] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem., 4:187–217, 1983.

[55] Oliver F Lange, Nils-Alexander Lakomek, Christophe Farès, Gunnar F Schröder, Korvin F a Walter, Stefan Becker, Jens Meiler, Helmut Grubmüller, Christian Griesinger, and Bert L de Groot. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. Science (New York, N.Y.), 320(5882):1471–5, June 2008. ISSN 1095-9203. doi: 10.1126/science.1157092.

[56] Leonardo G Trabuco, Elizabeth Villa, Eduard Schreiner, Christopher B Harrison, and Klaus Schulten. Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. Methods (San Diego, Calif.), 49(2):174–80, October 2009. ISSN 1095-9130. doi: 10.1016/j.ymeth.2009.04.005.

[57] Michael Habeck. Statistical mechanics analysis of sparse data. J. Struct. Biol., 173(3):541–8, March 2011. ISSN 1095-8657. doi: 10.1016/j.jsb.2010.09.016.

[58] M. Lorenz and K. C. Holmes. The actin-myosin interface. Proc. Natl. Acad. Sci. USA, 107:12529–12534, 2010.

[59] H. Vashisth, G. Skiniotis, and C.L. Brooks III. Using enhanced sampling and structural restraints to refine atomic structures into low-resolution electron microscopy maps. Structure, 20:1453–1462, 2012.

[60] A.D. MacKerell Jr, M. Feig, and C.L. Brooks III. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J. Comput. Chem., 25:1400–1415, 2004.

[61] B.Z. Lu, Y.C. Zhou, M.J. Holst, and J.A. McCammon. Recent progress in numerical methods for the Poisson-Boltzmann equation in biophysical applications. Commun. Comput. Physics, 3:973–1009, 2008.

[62] P. Koehl, H. Orland, and M. Delarue. Adapting Poisson-Boltzmann to the self-consistent mean fielf theory: Application to protein side-chain modeling. J. Chem. Phys., 135:055104, 2011.

[63] S. Jo, J.B. Lim, J.B. Klauda, and W. Im. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. Biophys. J., 97:50–58, 2009.

[64] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with NAMD. J. Comput. Chem., 26:1781–1802, 2005.

[65] R. Elber and M. Karplus. Enhanced sampling in molecular dynamics: Use of a time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. J. Am. Chem. Soc., 112: 9161–9175, 1990.

[66] A. Roitberg and R. Elber. Modeling side chains in peptides and proteins: Application of the locally enhanced sampling methods to find minimum energy conformations. J. Chem. Phys., 95:9277–9287, 1991.

[67] H. Yerushalmi, M. Lebendiker, and S. Schuldiner. EmrE, an Escherichia coli 12-kDa multidrug transporter, exchanges toxic cations and H+ and is soluble in organic solvents. J. Biol. Chem., 270:6856–6863, 1995.

[68] Y. Elbaz, S. Steiner-Mordoch, T. Danieli, and S. Schuldiner. In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. Proc. Natl. Acad. Sci. U. S. A., 101:1519–1524, 2004.

[69] K.M. Peters, B.E. Brooks, M.A. Schumacher, R.A. Skurray, R.G. Brennan, and M.H. Brown. A single acidic residue can guide binding site selection but does not govern qacr cationic-drug affinity. PLoS ONE, 6: E1597, 2011.

[70] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darwin, O. Guvench, P. Lopes, I. Vorobyev, and A.D. MacKerell Jr. CHARMM general force field: A force field and drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J. Comput. Chem., 31:671–690, 2009.

[71] S. Boresch, F. Tettinger, M. Leitgeb, and M. Karplus. Absolute binding free energies: A quantitative approach for their application. J. Phys. Chem. B, 107:9535–9551, 2003.

[72] J. Wang, Y. Deng, and B. Roux. Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. Biophys. J., 91:2798–2814, 2006.

[73] A. Pohorille, C. Jarzynski, and C. Chipot. Good practices in free energy calculations. J. Phys. Chem. B, 114: 10235–10253, 2010.

[74] M. Shirts. Best practices in free energy calculations for drug design. Methods Mol Biol., 819:425–467, 2012. doi: 10.1007/978-1-61779-465-0_26.

[75] W. Humphrey, A. Dalke, and K. Schulten. VMD - visual molecular dynamics. J. Mol. Graphics, 14:33–38, 1996.

[76] H. Yerushalmi and S. Schuldiner. An essential glutamyl residue in EmrE, a multidrug antiporter from Escherichia coli. J. Biol. Chem., 275:5264–5269, 2000.

[77] N. Gutman, S. Steiner-Mordoch, and S. Schuldiner. An amino acid cluster around the essential Glu-14 is part of the substrate and proton binding domain of EmrE, a multidrug transporter from Escherichia coli. J. Biol. Chem., 278:16082–16087, 2003.

[78] I. Lehner, D. Basting, B. Meyer, W. Haase, T. Manolikas, C. Kaiser, M. Karas, and C. Glaubitz. The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric. J. Biol. Chem., 283:3281–3288, 2008.

[79] D. Bashford and M. Karplus. pK_a's of Ionizable Groups in Proteins: Atomic Detail from a Continuum Electrostatic Model. Biochemistry, 29:10219–10225, 1990.

[80] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, and J.A. McCammon. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA, 98:10037–10041, 2001.

[81] S. Jo, M. Vargyas, J. Vasko-Szedlar, B. Roux, and W. Im. PBEQ-solver for online visualization of electrostatic potential of biomolecules. Nucleic Acids Res., 36:W270–W275, 2008.

[82] M. Tyka, A. Clarke, and R. Sessions. Absolute free-energy calculations of liquids using a harmonic reference state. J. Phys. Chem. B, 111:9571–9580, 2007.

[83] Federico Bernal, Andrew F. Tyler, Stanley J. Korsmeyer, Loren D. Walensky, and Gregory L. Verdine. Reactivation of the p53 tumor suppressor pathway by a stapled p53 peptide. Journal of the American Chemical Society, 129(9):2456–2457, 2007. doi: 10.1021/ja0693587.