### Table

<table>
<thead>
<tr>
<th>PI: Karplus, Martin</th>
<th>Title: Modeling atomic structure of the EmrE multidrug pump to design inhibitor peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td>Received: 06/14/2013</td>
<td>FOA: PA11-262 Council: 01/2014</td>
</tr>
<tr>
<td>Competition ID: ADOBE-FORMS-B2</td>
<td>FOA Title: NIH SMALL RESEARCH GRANT PROGRAM (PARENT R03)</td>
</tr>
<tr>
<td>1 R03 AI111416-01</td>
<td>Dual: Accession Number: 3599020</td>
</tr>
<tr>
<td>IPF: 3212901</td>
<td>Organization: HARVARD UNIVERSITY</td>
</tr>
<tr>
<td>Former Number:</td>
<td>Department: Chemistry and Chemical Biology</td>
</tr>
<tr>
<td>IRG/SRG: ZRG1 MSFD-N (08)F</td>
<td>AIDS: N Expedited: N</td>
</tr>
<tr>
<td>Subtotal Direct Costs (excludes consortium F&amp;A)</td>
<td>Animals: N Humans: N</td>
</tr>
<tr>
<td>Year 1:</td>
<td>Clinical Trial: N Current HS Code: 10</td>
</tr>
<tr>
<td>Year 2:</td>
<td>HESC: N</td>
</tr>
<tr>
<td>Senior/Key Personnel:</td>
<td>Organization: Role Category:</td>
</tr>
<tr>
<td>Martin Karplus Ph.D.</td>
<td>Harvard University PD/PI</td>
</tr>
<tr>
<td>Victor Ovchinnikov Ph.D</td>
<td>Harvard University Co-Investigator</td>
</tr>
</tbody>
</table>

*Always follow your funding opportunity's instructions for application format.* Although this application demonstrates good grantsmanship, time has passed since the grantee applied. The samples may not reflect the latest format or rules. NIAID posts new samples periodically: [https://www.niaid.nih.gov/grants-contracts/sample-applications](https://www.niaid.nih.gov/grants-contracts/sample-applications)

*The text of the application is copyrighted.* The awardee provided express permission for NIAID to post this grant application and summary statement for educational purposes. The awardee allows you to use the material (e.g., data, writing, graphics) they shared in the applications for nonprofit educational purposes only, provided the material remains unchanged and the principal investigators, awardee organizations, and NIH NIAID are credited.

*Freedom of Information Act (FOIA).* NIAID is strongly committed to protecting the integrity and confidentiality of the peer review process. When NIH responds to FOIA requests for grant applications and summary statements, the material will be subject to FOIA exemptions and include substantial redactions. NIH must protect all confidential commercial or financial information, reviewer comments and deliberations, and personal privacy information.

*Note on Section 508 Conformance and Accessibility.* We have reformatted this sample to improve accessibility for people with disabilities and users of assistive technology. If you have trouble accessing the content, contact the NIAID Office of Knowledge and Educational Resources at deaweb@niaid.nih.gov.
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 Enterobacteriaceae (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\(_\alpha\)-only structure is known. The C\(_\alpha\)-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 the drug 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 \(\alpha\)-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.

---


\(^2\) 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: Left panel: Three multidrug transporters from *E. coli* coordinate efflux of drugs from the cytoplasm into the medium (adapted from Ref. 1). Middle panel: Low-resolution X-ray 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. Right panel: A hydrocarbon-stapled \(\alpha\)-helical peptide (adapted from Ref. 6). The hydrocarbon staple is in red; in the notation \(X_n\), \(X\) refers to the stereochemistry at the \(\alpha\)-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, 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 quaternary 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., \(\beta\)-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 X-ray 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, data from cryo-electron microscopy (EM) and from NMR solution experiments, will be used together with recent developments in physics-based computer modeling of protein and membrane thermodynamics 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, it was found to be in conflict with subsequent EM, X-ray and NMR data. 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 and the optimal models will be tested experimentally. Stapled peptides have high α-helical propensities and have been used to disrupt protein-protein interactions important for the progression of cancer. 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 and neurotransmitter transporters, 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 X-ray and cryo-EM models. Although various structure prediction and refinement programs have been used before to predict optimal protein confirmations, ensembles of structures are needed to understand the function of membrane proteins as well as other large biomolecules. The use of extensive equilibration by Molecular Dynamics (MD) simulations of membrane proteins in a realistic lipid environment, with enhanced sampling methods that preserve the canonical ensemble (REMD or RLES) 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. 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, 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) as well as to inhibit HIV cell entry and to disrupt protein-protein interactions in tumor cells, 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 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]) and in protein structure prediction methods 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. 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 Verdi ne (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, 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 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. Trial sets of sidechain coordinates can be generated with several promising algorithms, and further refined by molecular dynamics simulations using restraints to the electron density map. 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.

1a. 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). Only Cα 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 Cα coordinates, coordinates of the protein backbone will be generated using the program CHARMM. 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.

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, OPUS-Rota, Self-Consistent Mean Field theory (SCMF), and ROSETTA. 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. Any chirality errors that arise in the modeling and minimization stages will be corrected at this stage.

1b. Molecular dynamics simulations for refinement of structures.

I. The modeled structures obtained in 1a will be inserted in pre-equilibrated patches of DPMC lipid bilayers using the CHARMM-gui interface. Although other membrane compositions may be used, EmrE is...
known to be functional in DPMC membranes. 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 using Graphical Processing Units (GPUs). The anticipated size of the simulation systems is ~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 will be carried out using the program NAMD, (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) simulations will be performed with the program CHARMM using implicit solvent representation for the membrane and cytoplasm, 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.

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. In the case that the coordinates do not fit completely, additional MD simulations with restraints to the EM map will be performed. 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 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.
(iii) In case that several distinct structures have very similar FE values, all of them will be retained for further validation and analysis (1d and 1e 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. 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 constants) 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+ and ethidium will be submitted to CHARMM-CGENFF server 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.

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, the TPP+ and ethidium ligands will be reversibly “annihilated” using alchemical FE simulation. The FE calculations will be performed with the MD program NAMD, and the analysis and postprocessing of simulation data will be done using the Visual Molecular Dynamics...
(VMD) software. 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.

1e 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. Using Poisson-Boltzmann solvation theory, 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 Cα dimer structure of EmrE, transmembrane helices 4 (TM4) are in close proximity for direct interaction (see Fig. 1). In this Aim, proteolysis-resistant peptide analogs will be designed to mimic TM4 (sequence ^DLPAIIGMMLICAGVLIINLL^105 ) for use as dimerization inhibitors.

2a 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 1b 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 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 2b.

2b 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 selected 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 randomly 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 interaction 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

\[ \Delta E = E_{\text{complex}}^{\text{mutated}} - E_{\text{complex}}^{\text{original}} - (E_{\text{peptide}}^{\text{mutated}} - E_{\text{peptide}}^{\text{original}}), \]  

(1)

where \( E_{\text{complex}} \) denotes the average interaction energy between the peptide and the EmrE monomer and membrane, \( E_{\text{peptide}} \) 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 \( \Delta E \) will be used in a Metropolis Monte-Carlo (MC) acceptance-rejection test, in which the probability of accepting the trial mutation is computed as

\[ P_{\text{accept}} = \frac{1}{\exp(-\beta \Delta E)} \text{ if } \Delta E \leq 0 \]

\[ = 0 \text{ if } \Delta E > 0, \]  

(2)

where the quantity \( \beta \) (inverse temperature) will be varied to explore different peptide compositions, increasing from the physiological value at 300K \( \beta=1/(k_B T) \approx 1.7 \text{kcal/mol}^{-1} \); 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 \( \alpha \)-helical propensity,\(^{29,83}\) resistance to proteolysis,\(^{6}\) and hydrophobicity of the staple, which is compatible with a high membrane permeability.\(^{53}\) 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: \( ^{84} \text{DLPAIIGMLICAGVLIINLLS}^{105} \) (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 \( \alpha \)-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 \( \alpha \)-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


