Concerted Rolling and Penetration of Peptides during Membrane Binding

Jacob M. Remington, Jonathon B. Ferrell, Severin T. Schneebeli, Jianing Li*.

Department of Chemistry, the University of Vermont, Burlington, VT 05405

*Jianing Li (jianing.li@uvm.edu)

Abstract

Peptide binding to membranes is common and fundamental in biochemistry and biophysics, and critical for applications ranging from drug delivery to the treatment of bacterial infections. However, it is largely unclear, from a theoretical point of view, what peptides of different sequences and structures share in the membrane binding and insertion process. In this work, we analyze three prototypical membrane-binding peptides (alpha-helical magainin and PGLa, and beta-hairpin tachyplesin) during membrane binding, using molecular details provided by Markov state modeling and microsecond-long molecular dynamics simulations. By leveraging both geometric and data-driven collective variables that capture the essential physics of the amphiphilic and cationic peptide-membrane interactions, we reveal how the slowest kinetic process of membrane binding is the dynamic rolling of the peptide from an attached to fully bound state. These results not only add fundamental knowledge into the theory of how peptides bind biological membranes, but also open new avenues to study general peptides in more complex environments for further applications.

Introduction

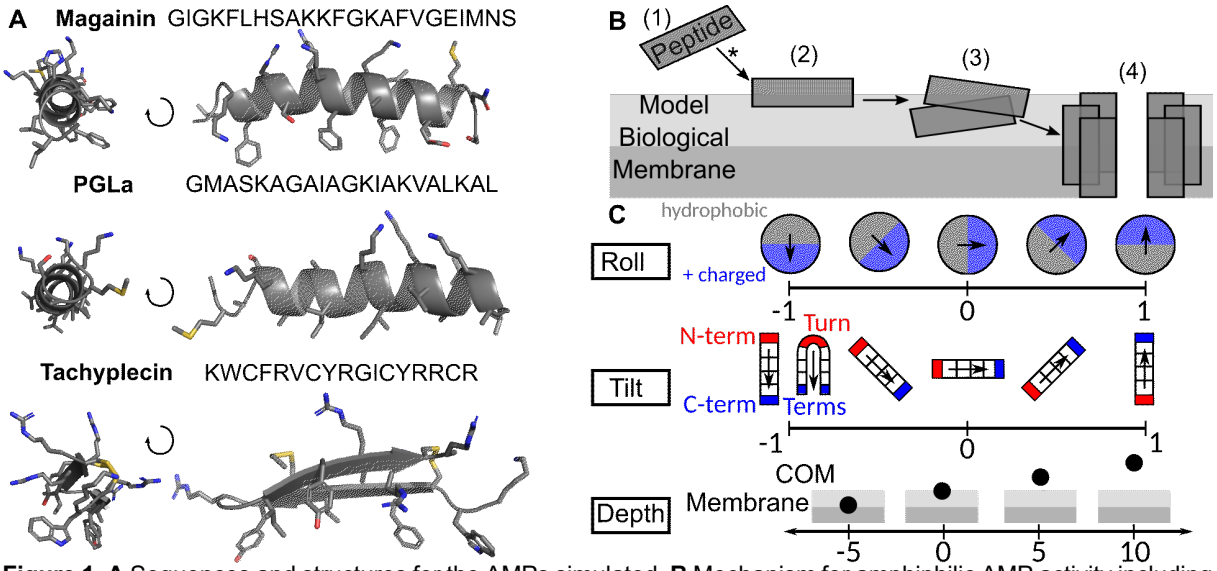

Figure 1. A Sequences and structures for the AMPs simulated. **B** Mechanism for amphiphilic AMP activity including the initial AMP binding (1) to (2) studied here, dimerization (2) to (3), and aggregation to form pores (3) to (4). **C** Proposed collective variables, (roll, tilt, and depth) for AMP binding. The roll is defined as the z-component of the unit normalized vector pointing from the center of mass of hydrophobic residues to positively charged residues. The tilt is the z-component of the unit normalized vector between the center of masses of the N-term and C-term of helical AMPs and the turn and terminal residues on hairpin AMPs. The depth is the z-component of the center of mass of the CA atoms on the peptide and the upper leaflet P atoms.

Membrane-binding peptides (MBP) have intriguing biological functions[1,2] and applications,[3,4] either on their own or as part of a multi-domain proteins. Most of these peptides are known as amphipathic helices[5], while a small but increasing number of them are now found to be non-helical.[6–11] Although a large body of molecular dynamics (MD) simulation studies have provided valuable detailed mechanisms of MBP[12,13] to complement experimental evidence, accurately modeling the peptide-membrane interactions remains a challenging task. It is mainly because the entire process occurs over large time- and length-scales with multiple intermediate steps.[14,15] As such, there lacks general insight — for both helical and non-helical MBPs — into the key steps of the membrane binding process. To delineate such process from a theoretical perspective, we report the use of the peptide roll, tilt, and depth coordinates (Figure 1) to resolve the initial membrane attachment and binding of three prototypical MBPs — helical magainin 2 (MAG) and PGLa, and hairpin Tachyplesin 1 (TAC). The simulations, totaling 9.6 μs per peptide, were analyzed with Markov state models (MSMs). Our results revealed that the roll of the amphiphilic peptides is dynamically coupled with the insertion depth of the peptide in the membrane, which together accurately resolve the slowest kinetic process of membrane binding. Further, the collective variables may be appropriate for enhanced sampling techniques of peptide binding to biological membranes, enabling future efforts to quantitatively assess the peptide-membrane binding propensity. Overall, the demonstration of this binding pathway will enable its manipulation and hopefully lead to the targeted enhancement of membrane activity.

Three peptides (MAG, PGLa, and TAC) were selected in this work, owing to a rich history of experimental data that was considered in the design of our computational approaches. Generally, at low peptide concentrations all three peptides form membrane bound states with hydrophobic residues interacting with the membrane core.[16–24] While increasing peptide concentration generally leads to membrane disruption often through pore formation[25–29] (Figure 1B). These experimentally resolved mechanistic steps emphasize the requirement of the MBP to attach and bind to the membrane, a step whose atomistic details are the subject of investigation herein. Building on prior knowledge, we aim to provide the missing dynamical details about membrane binding that has eluded experiments. Our atomistic MD simulations are designed to explore this key mechanistic step, offering valuable theoretical insight. We begin our efforts with a microsecond simulation of MAG

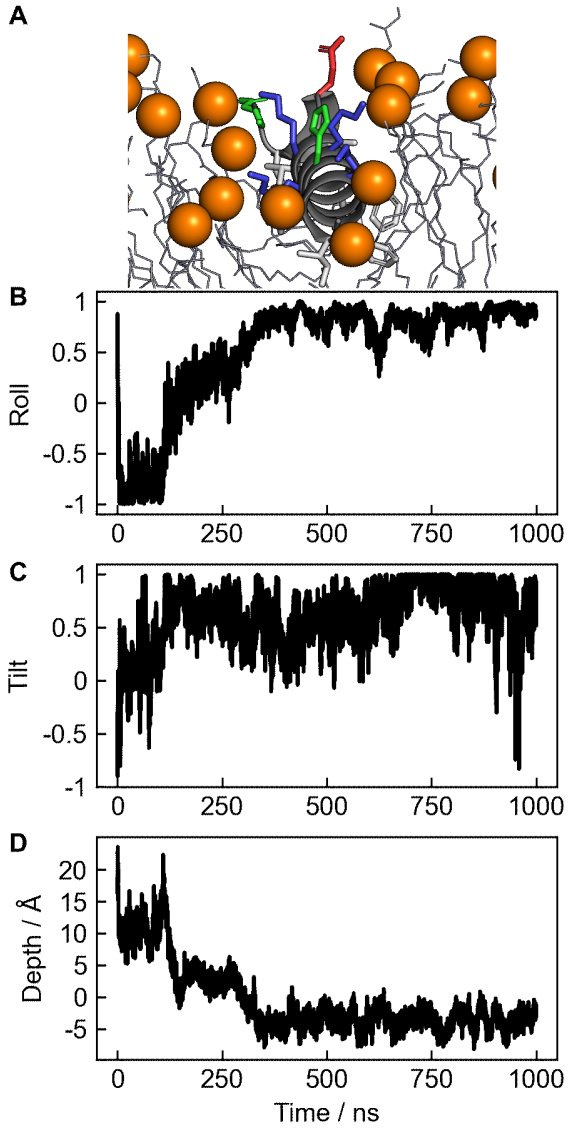

Figure 2. A Final snapshot of MAG from a 1-μs simulation initiated above a model bacterial membrane. The peptide sidechains are color-coded by residue type with the non-polar residues white, polar residues green, positively charged residues blue, and negatively charged residues red. The P atoms on the upper leaflet are shown as olive spheres. The positive z-axis is aligned vertically. **B, C,** and **D** Roll, tilt, and depth of MAG demonstrate it initially attaches to the membrane with a negative roll at times < 100 ns before fulling binding at 375 ns with a positive roll.

initialized above a lipid membrane wherein a single binding event is observed. Statistical sampling of such events for each of MAG, PGLa, and TAC is then demonstrated with Markov State Modeling and a guided search of the peptide's location and orientation relative to the membrane. Finally, the converged high dimensional free energy surfaces of the MBPs and minimum free energy pathways for MBP binding reveal a common underlying theme.

Results and Discussion

Initial Equilibrium Simulation of MAG Binding Reveals Slow Peptide Rolling

To explore the mechanistic details of peptide binding, we first performed a microsecond-long all-atom MD simulation of a single MAG peptide with a model bacterial membrane (Figure S1). During this simulation, MAG was observed to attach and embed into the bilayer with its helix nearly perpendicular to the bilayer normal (Figure 2A and C). This final structure is reminiscent of the MAG bound states observed experimentally[16], where the hydrophobic residues have penetrated the hydrophobic core of the bilayer and the positive sidechains form strong interactions with the upper-leaflet phosphate groups. This observation tracks well with the roll of MAG obtaining positive values at the end of the simulation (Figure 2B). Importantly, the MD simulations reveal key mechanistic details about how MAG may form this fully bound state. First, MAG rapidly attached to the membrane in the early 20 ns with the opposite value of roll near -1 and a tilt near 0. This prebound state resulted from the formation of the strong electrostatic interactions between the positive sidechains and the phosphate groups of the bilayer. However, the negative roll of this prebound state left the hydrophobic sidechains exposed to water. Over the next 200 ns of this simulation (from 100 to 375 ns) the peptide slowly reoriented by rolling into the helix to allow the more favorable hydrophobic-hydrophobic interactions and form the fully bound state with a corresponding decrease in the depth. Although just a single trajectory, this simulation suggests rolling plays a critical role in the membrane-binding mechanism of amphiphilic peptides. Therefore, these roll and depth coordinates well distinguish unbound and bound peptide orientations.

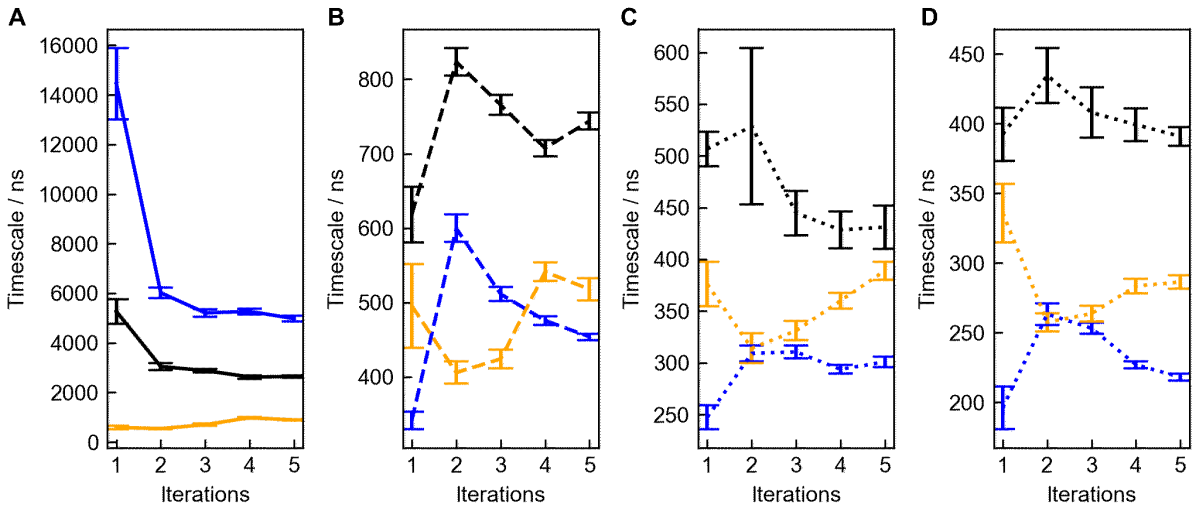

Figure 3. A The slowest estimated timescale for Bayesian MSMs of MAG (black), PGLa (blue), and TAC (orange) after successive iterations of the ensemble simulation protocol. **B, C,** and **D** The next three slowest timescales of the MSMs.

Guided Sampling of Peptide Binding with Markov State Modeling

To test if the observations from a single MBP binding event hold statistically, large-scale ensembles of simulations were performed and guided by iterative Bayesian estimations of the underlying MSM. To

start, 96 structures spanning the roll, tilt, and depth space were generated for each MBP using 96 distinct 50 ns Steered MD (SMD) simulations (Figure S1 and Table S1). During this stage, a simplifying assumption was made that the helical MBPs, which are known to be mostly disordered in solution[30], fold into helices on their encounter with a lipid bilayer. Accordingly, during SMD the peptides were restrained to their bound conformations which restricts our analysis to states where the peptides have attached to the membrane. While in principle the work done by the steering force during the SMD trajectories could provide free-energy estimates using the Jarzynski relationship,[31,32] this approach is notoriously slow to converge[33] and requires multiple simulations of each path. Further, many of the SMD trajectories penetrating deep in the bilayer resulted in large-scale deformations of the bilayer and involved significant work (>100 kcal/mol) to travel through, likely improbable, regions of the coordinate space. Thus, we relaxed each of the resultant structures and proceeded with 20-ns unbiased simulations. Bayseian MSMs were then estimated from the resultant 8 μs of unbiased production simulations using a lag time of 7.5 ns during the initial four stages of the protocol (Figure S2), and 15 ns for post-analysis and estimation of RTD MSMs described in the next section. The relative uncertainties of the eigenvectors corresponding to the four slowest kinetic processes were computed and used as estimates for the regions of the coordinate space where sampling was insufficient. Restarting new simulations from these regions would likely enhance the overall convergence of the MSM by focusing simulation efforts. Indeed, during four additional iterations of 96 distinct 20-ns simulations we observed a decrease in the error of the four slowest kinetic processes for each peptide (Figure S3). Further, in most cases changes between the 4th and 5th iteration are within the expected error from 4th iteration, signifying convergence of the protocol (Figure 3).

Coupling of Roll and Depth on Free Energy Surfaces for Peptide Binding

The MSMs with converged slowest timescale using the five iterations of each simulated MBP enabled the calculation of three-dimensional (3D) free energy surfaces (FES) on the roll, tilt, and depth (RTD) coordinates using the probabilities obtained from the equilibrium eigenvectors (details provided in the methods). Here, the 3D FES are visualized more intuitively using depth-wise slices. These slices reveal how the energy landscape along the roll and tilt coordinates change as the peptide binds the membrane (Figure 4). First, we focus on the similar features of these high dimensional FES shared among the three peptides. The most favorable roll value of the peptides changes from negative to positive values as the MBP inserts into the membrane. This supports the original observations from the long simulation, in that there are both attached and bound states of the MBP in the membrane. In the attached state, hydrophobic residues are exposed to the solvent while in the bound state, hydrophobic residues are embedded in the hydrophobic core of the bilayer. The later bound state of all three peptides agrees well with previous experimental observations.[16,17,19] The minimum free energy binding pathways computed using structures with roll, tilt, and depth values closest to the depth-wise slice minima are shown in Figure S4 and support this assignment of residue-membrane interactions.

Despite the overall similarities among the FES in Figure 4, we observed apparent differences between membrane binding in the three peptides. Focusing on MAG, two energy minima are observed; one above (depth, D = 7.5 Å) and one below (D = -5.0 Å) the membrane surface defined by the upper-leaflet P atoms (Figure 4A). For MAG the difference in free energy between the two states is less than RT (~0.6 kcal/mol) with the bound state slightly stabilized (~0.1 kcal/mol). However, for PGLa the bound state is stabilized much more (~1 kcal/mol). This difference between PGLa and MAG could be due to differences in the sequence including the presence of a negatively charged residue on MAG or the lack of aromatic hydrophobic residues in PGLa that each may affect binding (Figure 1).

In addition to the MSMs derived using the geometric RTD coordinates, we also used Time-lagged Independent Coordinate Analysis (TICA) to extract data-driven coordinates for peptide binding. A coarse grained and radial basis set was used as a featurization that smoothly counted how many interactions between hydrophobic sidechains, positively charged sidechains, lipid phosphate groups, and lipid tail groups were present at a range of distances in a simulation frame (More details in the Methods). Physical

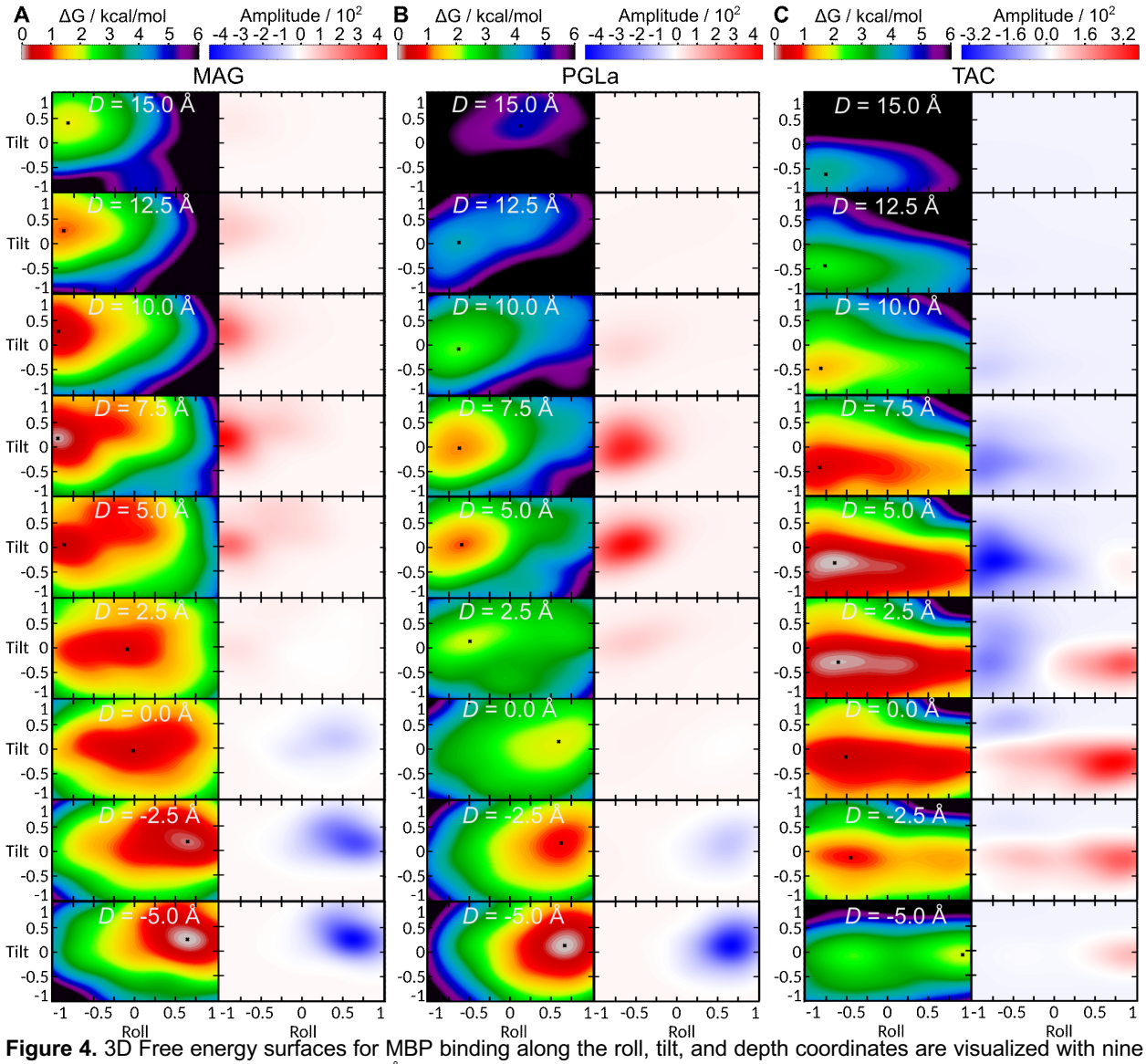

Figure 4. 3D Free energy surfaces for MBP binding along the roll, tilt, and depth coordinates are visualized with nine depth-wise slices ranging from 15 to -5 Å (left in each panel). The amplitude of the first non-equilibrium eigenvector of the MSMs in the same coordinate space reveal changes in MBP position and orientation associated with the slowest kinetic event in MBP binding (right in each panel). **A, B,** and **C** Surfaces for the three peptides MAG, PGL, and TAC are shown respectively. The surfaces are stacked so that the topmost correspond to MBP above the membrane P atoms, and the bottommost have the AMP below the membrane P atoms. The location of the minima and maxima of the non-equilibrium eigenvectors demonstrate that for all three peptides the slowest kinetic event is the rolling of the MBP across the hypothetical membrane surface defined by the P atoms.

interpretation of the first two TICA modes (TICA 1 and TICA 2) was achieved in two ways; (1) clustering the trajectories using TICA 1 and TICA 2 and averaging the RTD coordinates conditional on the cluster centers (Figure S5), and (2) plotting the correlation between the TICA modes with the input featurization components (Figures S6 to S11). First, the coupling between roll and depth can be clearly seen in TICA 1 for MAG and PGL, and TICA 2 for TAC (Figure S5). As these respective TICA components increase, the depth of the peptide decreases and the roll changes sign signifying a reorientation of the peptides as they enter the membrane akin to the physical events observed in the RTD derived MSM. Indeed, FES of the first two TICA components show clear energy basins corresponding to the attached and bound states of the peptides (Figure 5A1, 5B1, and 5C1). For MAG and PGL, TICA 2 was not correlated with RTD and TICA 1

alone for TAC only moderately distinguishes between the attached and bound states. The physical interpretation of these non-RTD related components can be seen in Figures S7, S9, and S10, that show how they, unexpectedly, correlate with changes in the lipid bilayer. For MAG and PGL the TICA 2 related membrane changes do not appear to couple to TICA 1 related peptide binding (Figure 5A1 and 5B1), however for TAC peptide binding and lipid changes are both mixed into TICA 1 in a complex pattern (Figure 5C1 and S10). The conformational preferences of the MBP and the interplay between lipid changes and peptide binding are clarified by exploring their kinetic properties in the next section.

Slowest Kinetic Event in Binding is Crossing the Bilayer's Phosphate Head Groups

Kinetic information available from the MSMs suggest the transition pathway between the attached and bound states involves a smooth coupling between the roll and depth coordinates among all three peptides (Figures 4 and S5). A key property of a MSM's transition probability matrix is that the eigenvalues can be sorted into a hierarchy of transition timescales whose associated right eigenvectors describe kinetic events that relax non-equilibrium state distributions[34]. For all three peptides and in both RTD and TICA derived MSMs the slowest of these kinetic processes is qualitatively the same and corresponds to transitions between the attached and bound states (Figure 4 and S13). This later finding follows from the location of the minima and maximum amplitude of the first non-equilibrium eigenvector closely matching the locations of these two states in the roll, tilt, and depth coordinates (Figure 4) and the TICA regions with associated positive roll and low depth (Figure S5 and S13). Interestingly, for MAG and PGL the second and third non-equilibrium eigenvectors are associated with TICA 2, a bilayer related mode described in the previous section. For these peptides, there is a clear separation of timescales between the rolling and membrane dynamics (Figure S12A and B) and the MSM analysis can distinguish these two processes. For TAC, the separation of timescales for the membrane binding mode from was poor (Figure S12C) and the resulting first two slowest processes, while being associated with similar changes in the roll and depth (Figure S5C), also were mixed with changes to the lipid bilayer seen in TICA 1 (Figure S10). However, given that roll and depth associated changes are present in the slowest kinetic event among these three MBPs it can be hypothesized that the phenomena may be more general among amphiphilic peptides.

Furthermore, the slowest timescales associated with the binding transition for the two helical MBPs, PGL and MAG, were 5 and 3 μs respectively (Figure 3A). The slower timescale for PGLa appears to correlate with the larger height of the barrier near D = 0.0 Å in comparison with MAG (Figure 4A and B). A similar increase in barrier is seen in the saddle near TICA 1 = -1 and -0.5 for MAG and PGL respectively (Figure 5). A coarse mean first passage time (MFPT) estimate between the two states using the TICA derived MSMs show the sets of attached and bound states can interconvert on faster timescales (Figure 5) although the microstates used for attached and bound state classification through Hidden MSM estimation suggest the MFPTs are underestimated (see Methods and Supporting Information for more details including Chapman-Kolmogorov (CK) tests). On the other hand, the bound state for TAC was disfavored by approximately 1 kcal/mol relative to the attached state, although the RTD coordinates poorly capture the barrier between the two states near D = 0.0 Å (Figure 4C). The TICA coordinates capture a similar difference in energy but were able to reveal a significant barrier at TICA 2 = 2.5 with a height of approximately 1 kcal/mol relative to the bound state. Thus, a more rapid exchange between these two states for TAC is expected and confirmed by the longest timescale for TAC, nearly 1 μs, 3-5 times faster than the helical peptides. This suggests that while the hairpin peptide can more rapidly transition between attached and bound states, the attached state predominates. The slow timescales for peptide binding from the MSMs highlights the requirements for significant computational efforts when modeling peptide-membrane interactions.

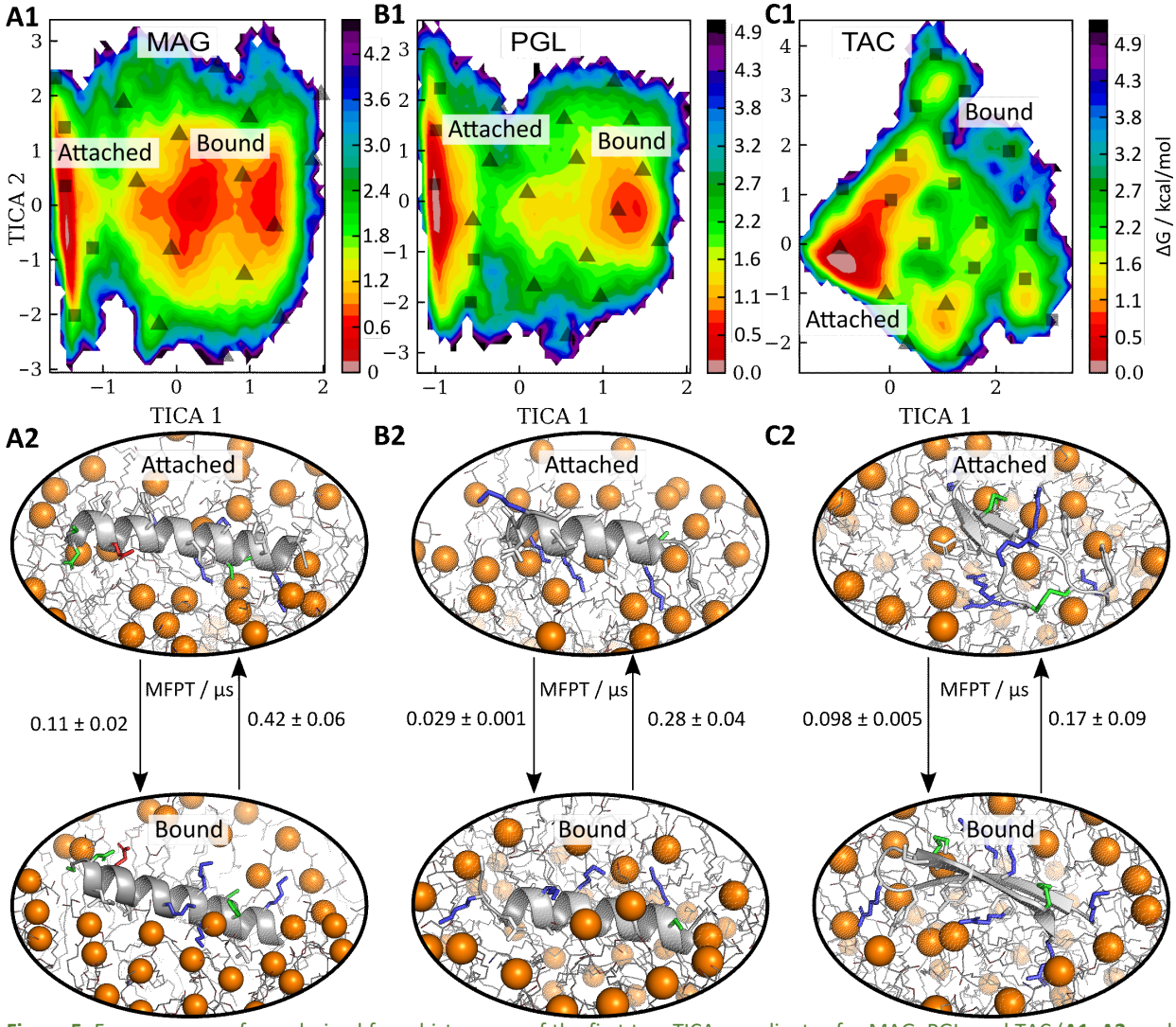

Figure 5. Free energy surfaces derived from histograms of the first two TICA coordinates for MAG, PGL, and TAC (**A1**, **A2**, and **A3** respectively). Markers depict the classification of regions in the TICA space using the memberships of two-state hidden Markov state models, where squares are attached peptides and triangles are bound peptides. Clear transition regions are observed for MAG and PGL at TICA 1 = -1 and -0.4 respectively, while for TAC the bound states are seen with large TICA 1 and TICA 2 values and a more generally diverse free energy surface is observed. Example simulation frames from the most populated microstate with each classification are shown in **B**. Also shown in **B** are the mean first passage times between the attached and bound states. Errors are standard deviations from the Bayesian MSMs estimated using 500 samples.

Conclusion

MBP are ubiquitous on their own (such as many antimicrobial peptides) and as part of complex proteins (such as human P450s,[35] Bax,[36] and the envelop proteins of viruses[37]). While extensive simulation studies have gained important molecular insight into how individual peptides are bound to biological membranes, we in this work revealed the common membrane binding mechanism of helical and non-helical MBPs, from a new theoretical viewpoint. Our ensembles of equilibrium MD simulations guided by iterative MSM estimation revealed the intricate details of peptide binding to model bacterial membranes. With this approach, ~8 μs of aggregate simulation time was required to converge the slowest four timescales in the MSMs of MAG, TAC, and PGL. The MSMs demonstrated that the slowest kinetic event in peptide binding

is the rolling of the peptide amphiphilic moment (defined by the orientation of the hydrophobic and positively charged residues relative to the membrane normal). This rolling is dynamically coupled to the binding process, suggesting multiple paths for future simulation efforts. For one, collective variables designed to help sample peptide binding events (e.g. for screening the membrane binding propensity of antimicrobial sequences generated by machine learning[38–40]) should accommodate changes in both the roll and penetration depth. While in principle data-driven collective variables[41–47] could capture these events, the clear mechanism provided by the geometric roll, tilt, and depth coordinates demonstrates the value in physics-based CVs in elucidating the chemical/biological process and dynamics. Additionally, differences in primary sequence between MAG and PGLa appear to alter the energy barrier height for membrane penetration, which suggest the rational design of MBPs could be approached through in silico mutagenesis and the measurement of such energetic profiles. Furthermore, the finding that all three of the studied peptides had qualitatively similar slowest kinetic processes suggests a more general mechanism for binding of MBPs into lipid bilayers. Overall, our findings from this work may open the avenues to study general MBPs with similar methodology and guide future investigations of more complex systems (e.g., peptide folding and aggregation on the membrane surface).

Methods

All-atom models of MAG, TAC, and PGL were prepared using CHARMM-GUI[48] from helical models for MAG and PGL, while TAC was built using PDBID: 2RTV.[49] The exploratory simulation used a 3:1 mixture of POPG and POPE lipids. The MSM simulations used a slightly more accurate to *E. Coli*. Inner membrane model including an additional 5% by mass cardiolipin, POCL[50]. NaCl was used as the salt and counter ions to mimic a salt concentration of 100 mM. A 15-Å buffer was added to the AMP's aligned with a tilt of 0 to define the x and y dimensions of the initial box size. The z-dimension was set to ~90 Å, which included space to allow peptides to freely diffuse at least 30 Å above the upper leaflet without interacting (coming within 10 Å) with the image of the lower leaflet. The CHARMM-36m forcefield was used to model all interactions. MD simulations were carried out using the AMBER simulation package[51] and the SMD simulations used PLUMED[52]. For SMD the spring constants for roll, tilt, and depth (see below) were held at 1000 kcal/mol, 1000 kcal/mol, and 50 kcal/mol Å respectively. During SMD simulations harmonic restraints were added on root mean squared displacement of peptide CA atoms to hold the peptides in their initial structures using PLUMED with a spring constant of 10 kcal/mol Å. In this way energetic contributions from the folding/unfolding of the peptides in solution were neglected. While these MBP are known to be more unstructured in solution than when in a membrane, it is unclear if folding occurs during attachment or binding. It is our hypothesis herein that folding occurs during attachment and can thus be neglected while studying peptide binding. Production simulations were conducted at 303.15 K in the NPγT ensemble using the GPU accelerated version of pmemd[53].

The RTD coordinates were computed using pytraj as shown in Figure 1[54]. The roll was computed using the z-component of the unit normalized vector pointing from the center of mass of the hydrophobic residues to positively charged residues. The tilt was computed similarly using the z-component of the of the unit normalized vector pointing from either (1) the center of mass of the last four residues to the first two for helical MBP, or (2) the center of masses of the four residues in the turn the first and last two residues for hairpin AMP. The depth was computed as the z-component of the displacement vector between the COM of the peptide CA atoms and the COM of the P atoms in the upper leaflet. One drawback of the roll and tilt coordinates being defined with respect to the positive z axis is that crossing the z-boundary of the periodic boundary instantaneously flips the sign of roll and tilt. Accordingly, before each 20 ns simulation, we recentered the system on the upper leaflet phosphate atoms to remove small drifting of membrane in the z-dimension. Further we removed a small number of simulations (< 10 %) where the depth value increased greater than 30 Å or penetrated too far in the membrane less than -10 Å (28 for MAG, 38 for PGL, and 2 for TAC). The roll and tilt have natural bounds at +/- 1, while the periodic boundary conditions employed in our simulations meant the depth is bounded by the z-dimension of the box. These two assumptions

restricted our simulations and analysis to interrogate interactions between the MBPs and the membrane that occur after attachment and during binding.

After our sampling protocol the roll, tilt, and depth from the aggregate of ensemble simulations were clustered using a uniform space clustering algorithm with a minimum distance of 0.5 in the roll, tilt, and depth space. We note that during the five iterations of short trajectory generation k-means clustering with 250 microstates was used to avoid the reparameterization of the minima distance parameter in uniform space clustering. Reversible Bayesian MSMs using 500 samples of the posterior were estimated using the Pyemma package[55]. At each iteration MSMs were re-estimated, and simulation frames randomly selected from 24 states with the highest relative uncertainty in the first four slowest eigenvectors of the MSM. As the amplitudes of eigenvectors are only unique up to a multiplicative scalar care needs to be taken when estimating statistical uncertainties directly from the eigenvectors from distinct transition probability matrices as is done with Bayesian MSMs. First, the eigenvectors were normalized using the 2-norm. In the following, an eigenvector of the i^{th} sampled MSM is denoted as v_i, and an arbitrary first sample is chosen as v_0. For every other sampled eigenvector, the scalar x in {-1,1} that maximized $x^* v_i \cdot v_0$ was used to pick the appropriate $x^* v_i$ before computing averages and variances. In this way, erroneous variance from sampled eigenvectors arbitrarily changing signs is eliminated. The 3D free energy surfaces and amplitudes for the 3D non-equilibrium eigenvectors were computed from a kernel density estimation (using gaussian kernels and a width of 0.05 in the roll and tilt coordinates and a width of 1 Å in the depth coordinate) using the probabilities (and amplitudes for non-equilibrium eigenvectors) of the microstates as weights. Normalization of the 3D surface were performed before slicing for proper comparison of slices.

A coarse grained (CG) and atom centered symmetry function[56] inspired featurization was used for input to TICA to describe the detailed peptide and membrane interactions. In this description four CG bead types were defined; (1) the hydrophobic residues (ALA, VAL, ILE, LEU, MET, PHE, TYR, TRP, and PRO), (2) the positively charged residues (ARG, HIS, and LYS), (3) the P atoms on the lipid head groups, and (4) the carbon tails defined by groups of 3 atoms (C23, C25, and C29), (C33, C36, and C39), (CA3, CA5 and CA9), (CB3, CB5, and CB9), (CC3, CC5, and CC9), or (CC3, CC5, and CC9). Using the center of mass of each CG bead type the atomistic structures were mapped to CG structures and the pair wise distance array between all beads were computed. The pairwise distance arrays were then symmetrized by counting the number of the 12 CG interaction types (whose CG distance is $r_{n,m}$ between CG types c_i and c_j) within 20 slices linearly spaced with radii, r_s, 3 to 15 Å and a width σ = 2 Å^{-1} using Equation 1.

$$Featurizaiton(r_s, c_i, c_j) = \sum_{n=1}^{N_i} \sum_{m=1}^{N_j} \frac{1}{(N_i N_j)} e^{-\sigma(r_{n,m} - r_s)^2} \quad \text{(Equation 1)}$$

In Equation 1 the summations are over the pairs of N_i beads of type c_i and N_j beads of type c_j. This resulted in an input featurization dimension of 240 for TICA computation. TICA[57] can be understood qualitatively by first considering principal component analysis (PCA) wherein a hierarchy of linear combinations of input variables are derived that capture decreasing amounts of variance in the input data. In PCA, this is achieved through diagonalization of the covariance matrix for the input variables. The key difference with TICA is that the covariance matrix is replaced with a time-lagged covariance matrix that measures the variations in input data given a fixed period has elapsed (defined by the TICA lag time, τ_{TICA}, here 3 ns for all three peptides). This has the effect of lowering the contribution to the time-lagged covariance matrix of more rapid fluctuations of input variables within metastable states and increasing the contributions of slow transitions between metastable states. As such, regular space clustering was used with a minimum distance of 0.9 for the MSM estimation using the TICA coordinates. Like the RTD MSMs, for the TICA trajectories Bayesian MSMs with 500 samples were estimated with lag times of 5.5 ns for MAG and PGL and 7.5 ns for TAC. CK tests were performed (Figures S14-S16) using 2 coarse state models as the separation in timescales was

only significant for the 1st and 2nd slowest processes (Figure S12). Memberships associated with the two states that are the probability of a microstate belonging to a given microstate were computed from Hidden MSMs[58] and a cut-off of 0.5 was used for microstate assignment in Figure 5. The same assignments were used for estimating MFPTs for MBP binding.

Acknowledgments

Computations were performed on the Vermont Advanced Computing Core supported in part by NSF award No. OAC-1827314. JMR and JL were partially supported by an NIH R01 award (R01GM129431 to J.L.) and J.B.F. was supported by an NSF award (CHE-1945394 to J.L.). S.T.S. was supported by the Army Research Office (Grant 71015-CH-YIP). Part of the computational facilities were also supported by an NSF CAREER award (Grant CHE-1848444 awarded to S.T.S).

References

(1) Liao, C.; Zhang, Z.; Kale, J.; Andrews, D. W.; Lin, J.; Li, J. Conformational Heterogeneity of Bax Helix 9 Dimer for Apoptotic Pore Formation. *Scientific reports* **2016**, *6*, 29502.

(2) Lambright, D. G. 5.11 Membrane Recruitment of Signaling Domains. In *Comprehensive Biophysics*; Egelman, E. H., Ed.; Elsevier: Amsterdam, 2012; pp 223–244.

(3) Gori, A.; Romanato, A.; Bergamaschi, G.; Strada, A.; Gagni, P.; Frigerio, R.; Brambilla, D.; Vago, R.; Galbiati, S.; Picciolini, S.; Bedoni, M.; Daaboul, G. G.; Chiari, M.; Cretich, M. Membrane-Binding Peptides for Extracellular Vesicles on-Chip Analysis **2020**, *9* (1), 1751428.

(4) Taylor, R. E.; Zahid, M. Cell Penetrating Peptides, Novel Vectors for Gene Therapy. *Pharmaceutics* **2020**, *12* (3), 225.

(5) Liao, C.; Esai Selvan, M.; Zhao, J.; Slimovitch, J. L.; Schneebeli, S. T.; Shelley, M.; Shelley, J. C.; Li, J. Melittin Aggregation in Aqueous Solutions: Insight from Molecular Dynamics Simulations. *The journal of physical chemistry. B* **2015**, *119* (33), 10390–10398.

(6) Last, N. B.; Schlamadinger, D. E.; Miranker, A. D. A Common Landscape for Membrane-Active Peptides. *Protein Sci* **2013**, *22* (7), 870–882.

(7) Yang, J.; Weliky, D. P. Solid-State Nuclear Magnetic Resonance Evidence for Parallel and Antiparallel Strand Arrangements in the Membrane-Associated HIV-1 Fusion Peptide. *Biochemistry* **2003**, *42* (40), 11879–11890.

(8) Wang, W.; Tan, J.; Ye, S. Unsaturated Lipid Accelerates Formation of Oligomeric β-Sheet Structure of GP41 Fusion Peptide in Model Cell Membrane. *J. Phys. Chem. B* **2020**, *124* (25), 5169–5176.

(9) Khan, M. M.; Filipczak, N.; Torchilin, V. P. Cell Penetrating Peptides: A Versatile Vector for Co-Delivery of Drug and Genes in Cancer. *Journal of Controlled Release* **2021**, *330*, 1220–1228.

(10) Bolintineanu, D.; Hazrati, E.; Davis, H. T.; Lehrer, R. I.; Kaznessis, Y. N. Antimicrobial Mechanism of Pore-Forming Protegrin Peptides: 100 Pores to Kill E. Coli. *Peptides* **2010**, *31* (1), 1–8.

(11) Langham, A. A.; Ahmad, A. S.; Kaznessis, Y. N. On the Nature of Antimicrobial Activity: A Model for Protegrin-1 Pores. *J. Am. Chem. Soc.* **2008**, *130* (13), 4338–4346.

(12) Reid, L. M.; Verma, C. S.; Essex, J. W. The Role of Molecular Simulations in Understanding the Mechanisms of Cell-Penetrating Peptides. *Drug Discovery Today* **2019**, *24* (9), 1821–1835.

(13) Bolintineanu, D. S.; Kaznessis, Y. N. Computational Studies of Protegrin Antimicrobial Peptides: A Review. *Peptides* **2011**, *32* (1), 188–201.

(14) Palmer Nicholas; Maasch Jacqueline R. M. A.; Torres Marcelo D. T.; de la Fuente-Nunez César; Richardson Anthony R. Molecular Dynamics for Antimicrobial Peptide Discovery. *Infection and Immunity 89* (4), e00703-20.

(15) Remington, J. M.; Liao, C.; Sharafi, M.; Ste.Marie, E. J.; Ferrell, J. B.; Hondal, R. J.; Wargo, M. J.; Schneebeli, S. T.; Li, J. Aggregation State of Synergistic Antimicrobial Peptides. *J. Phys. Chem. Lett.* **2020**, *11* (21), 9501–9506.

(16) Bechinger, B.; Zasloff, M.; Opella, S. J. Structure and Orientation of the Antibiotic Peptide Magainin in Membranes by Solid-State Nuclear Magnetic Resonance Spectroscopy. *Protein Sci* **1993**, *2* (12), 2077–2084.

(17) Bechinger, B.; Zasloff, M.; Opella, S. J. Structure and Dynamics of the Antibiotic Peptide PGLa in Membranes by Solution and Solid-State Nuclear Magnetic Resonance Spectroscopy. *Biophys J* **1998**, *74* (2 Pt 1), 981–987.
(18) Glaser, R. W.; Sachse, C.; Dürr, U. H. N.; Wadhwani, P.; Ulrich, A. S. Orientation of the Antimicrobial Peptide PGLa in Lipid Membranes Determined from 19F-NMR Dipolar Couplings of 4-CF3-Phenylglycine Labels. *Journal of Magnetic Resonance* **2004**, *168* (1), 153–163.
(19) Kawano, K.; Yoneya, T.; Miyata, T.; Yoshikawa, K.; Tokunaga, F.; Terada, Y.; Iwanaga, S. Antimicrobial Peptide, Tachyplesin I, Isolated from Hemocytes of the Horseshoe Crab (Tachypleus Tridentatus). NMR Determination of the Beta-Sheet Structure. *J Biol Chem* **1990**, *265* (26), 15365–15367.
(20) Tamamura, H.; Kuroda, M.; Masuda, M.; Otaka, A.; Funakoshi, S.; Nakashima, H.; Yamamoto, N.; Waki, M.; Matsumoto, A.; Lancelin, J. M.; Kohda, D.; Tate, S.; Inagaki, F.; Fujii, N. A Comparative Study of the Solution Structures of Tachyplesin I and a Novel Anti-HIV Synthetic Peptide, T22 ([Tyr5,12, Lys7]-Polyphemusin II), Determined by Nuclear Magnetic Resonance. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1993**, *1163* (2), 209–216.
(21) Oishi, O.; Yamashita, S.; Nishimoto, E.; Lee, S.; Sugihara, G.; Ohno, M. Conformations and Orientations of Aromatic Amino Acid Residues of Tachyplesin I in Phospholipid Membranes. *Biochemistry* **1997**, *36* (14), 4352–4359.
(22) Park, N. G.; Lee, S.; Oishi, O.; Aoyagi, H.; Iwanaga, S.; Yamashita, S.; Ohno, M. Conformation of Tachyplesin I from Tachypleus Tridentatus When Interacting with Lipid Matrixes. *Biochemistry* **1992**, *31* (48), 12241–12247.
(23) Doherty, T.; Waring, A. J.; Hong, M. Peptide–Lipid Interactions of the β-Hairpin Antimicrobial Peptide Tachyplesin and Its Linear Derivatives from Solid-State NMR. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2006**, *1758* (9), 1285–1291.
(24) Haris, P. I.; Chapman, D. The Conformational Analysis of Peptides Using Fourier Transform IR Spectroscopy. *Biopolymers* **1995**, *37* (4), 251–263.
(25) Kim, C.; Spano, J.; Park, E.-K.; Wi, S. Evidence of Pores and Thinned Lipid Bilayers Induced in Oriented Lipid Membranes Interacting with the Antimicrobial Peptides, Magainin-2 and Aurein-3.3. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2009**, *1788* (7), 1482–1496.
(26) Ludtke, S. J.; He, K.; Heller, W. T.; Harroun, T. A.; Yang, L.; Huang, H. W. Membrane Pores Induced by Magainin. *Biochemistry* **1996**, *35* (43), 13723–13728.
(27) Glaser, R. W.; Sachse, C.; Dürr, U. H. N.; Wadhwani, P.; Afonin, S.; Strandberg, E.; Ulrich, A. S. Concentration-Dependent Realignment of the Antimicrobial Peptide PGLa in Lipid Membranes Observed by Solid-State 19F-NMR. *Biophysical Journal* **2005**, *88* (5), 3392–3397.
(28) Tremouilhac, P.; Strandberg, E.; Wadhwani, P.; Ulrich, A. S. Conditions Affecting the Re-Alignment of the Antimicrobial Peptide PGLa in Membranes as Monitored by Solid State 2H-NMR. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2006**, *1758* (9), 1330–1342.
(29) Strandberg, E.; Esteban-Martín, S.; Salgado, J.; Ulrich, A. S. Orientation and Dynamics of Peptides in Membranes Calculated from 2H-NMR Data. *Biophysical Journal* **2009**, *96* (8), 3223–3232.
(30) Matsuzaki, K. Magainins as Paradigm for the Mode of Action of Pore Forming Polypeptides. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1998**, *1376* (3), 391–400.
(31) Jarzynski, C. Equilibrium Free-Energy Differences from Nonequilibrium Measurements: A Master-Equation Approach. *Phys. Rev. E* **1997**, *56* (5), 5018–5035.
(32) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78* (14), 2690–2693.
(33) Goette, M.; Grubmüller, H. Accuracy and Convergence of Free Energy Differences Calculated from Nonequilibrium Switching Processes. *Journal of Computational Chemistry* **2009**, *30* (3), 447–456.
(34) Husic, B. E.; Pande, V. S. Markov State Models: From an Art to a Science. *J. Am. Chem. Soc.* **2018**, *140*, 2386–2396.
(35) Baylon, J. L.; Lenov, I. L.; Sligar, S. G.; Tajkhorshid, E. Characterizing the Membrane-Bound State of Cytochrome P450 3A4: Structure, Depth of Insertion, and Orientation. *J. Am. Chem. Soc.* **2013**, *135* (23), 8542–8551.
(36) Guo, B.; Zhai, D.; Cabezas, E.; Welsh, K.; Nouraini, S.; Satterthwait, A. C.; Reed, J. C. Humanin Peptide Suppresses Apoptosis by Interfering with Bax Activation. *Nature* **2003**, *423* (6938), 456–461.

(37) Alsaadi, E. A. J.; Neuman, B. W.; Jones, I. M. Identification of a Membrane Binding Peptide in the Envelope Protein of MHV Coronavirus. *Viruses* **2020**, *12* (9).
(38) Van Oort, C. M.; Ferrell, J. B.; Remington, J. M.; Wshah, S.; Li, J. AMPGAN v2: Machine Learning-Guided Design of Antimicrobial Peptides. *J. Chem. Inf. Model.* **2021**, *61* (5), 2198–2207.
(39) Müller, A. T.; Gabernet, G.; Hiss, J. A.; Schneider, G. ModlAMP: Python for Antimicrobial Peptides. *Bioinformatics* **2017**, *33* (17), 2753–2755.
(40) Nagarajan, D.; Nagarajan, T.; Roy, N.; Kulkarni, O.; Ravichandran, S.; Mishra, M.; Chakravortty, D.; Chandra, N. Computational Antimicrobial Peptide Design and Evaluation against Multidrug-Resistant Clinical Isolates of Bacteria. *Journal of Biological Chemistry* **2018**, *293* (10), 3492–3509.
(41) Chen, W.; Ferguson, A. L. Molecular Enhanced Sampling with Autoencoders: On-the-Fly Collective Variable Discovery and Accelerated Free Energy Landscape Exploration. *Journal of Computational Chemistry* **2018**, *39* (25), 2079–2102.
(42) Sultan, M. M.; Pande, V. S. Automated Design of Collective Variables Using Supervised Machine Learning. *J. Chem. Phys.* **2018**, *149* (9), 094106.
(43) Chen, W.; Tan, A. R.; Ferguson, A. L. Collective Variable Discovery and Enhanced Sampling Using Autoencoders: Innovations in Network Architecture and Error Function Design. *J. Chem. Phys.* **2018**, *149* (7), 072312.
(44) Hashemian, B.; Millán, D.; Arroyo, M. Modeling and Enhanced Sampling of Molecular Systems with Smooth and Nonlinear Data-Driven Collective Variables. *J. Chem. Phys.* **2013**, *139* (21), 214101.
(45) Mardt, A.; Pasquali, L.; Wu, H.; Noe, F. VAMPnets for Deep Learning of Molecular Kinetics. *Nat. Commun.* **2018**, *9*, 11.
(46) Chen, W.; Sidky, H.; Ferguson, A. L. Nonlinear Discovery of Slow Molecular Modes Using State-Free Reversible VAMPnets. *J. Chem. Phys.* **2019**, *150* (21), 214114.
(47) Bonati, L.; Rizzi, V.; Parrinello, M. Data-Driven Collective Variables for Enhanced Sampling. *J. Phys. Chem. Lett.* **2020**, *11* (8), 2998–3004.
(48) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *Journal of computational chemistry* **2008**, *29* (11), 1859–1865.
(49) Kushibiki, T.; Kamiya, M.; Aizawa, T.; Kumaki, Y.; Kikukawa, T.; Mizuguchi, M.; Demura, M.; Kawabata, S.; Kawano, K. Interaction between Tachyplesin I, an Antimicrobial Peptide Derived from Horseshoe Crab, and Lipopolysaccharide. *Biochim Biophys Acta* **2014**, *1844* (3), 527–534.
(50) Sohlenkamp, C.; Geiger, O. Bacterial Membrane Lipids: Diversity in Structures and Pathways. *FEMS Microbiology Reviews* **2016**, *40* (1), 133–159.
(51) Amber 2021, University of California, San Francisco. 2021.
(52) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Computer Physics Communications* **2014**, *185* (2), 604–613.
(53) Lee, T.-S.; Cerutti, D. S.; Mermelstein, D.; Lin, C.; LeGrand, S.; Giese, T. J.; Roitberg, A.; Case, D. A.; Walker, R. C.; York, D. M. GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J. Chem. Inf. Model.* **2018**, *58*, 2043–2050.
(54) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
(55) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernàndez, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J. H.; Noé, F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015**, *11*, 5525–5542.
(56) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 074106.
(57) Pérez-Hernández, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noé, F. Identification of Slow Molecular Order Parameters for Markov Model Construction. *J. Chem. Phys.* **2013**, *139*, 015102.
(58) Noé, F.; Wu, H.; Prinz, J. H.; Plattner, N. Projected and Hidden Markov Models for Calculating Kinetics and Metastable States of Complex Molecules. *J. Chem. Phys.* **2013**, *139*, 184114.