

Analysis of Charged Peptide Loop-flipping Across a Lipid Bilayer Using the String Method with Swarms-of-trajectories

Samarthaben J. Patel and Reid C. Van Lehn*

Department of Chemical and Biological Engineering

University of Wisconsin – Madison, Madison, WI, 53706, USA.

*send correspondence to: vanlehn@wisc.edu

Abstract

The hydrophobic core of the lipid bilayer is conventionally considered a barrier to the translocation of charged species such that the translocation of even single ions occurs on long timescales. In contrast, experiments have revealed that some materials, including peptides, proteins, and nanoparticles, can translocate multiple charged moieties across the bilayer on experimentally relevant timescales. Understanding the molecular mechanisms underlying this behavior is challenging, however, because resolving corresponding free energy landscapes with molecular simulation techniques is computationally expensive. To address this challenge, we use atomistic molecular dynamics simulations with the swarms-of-trajectories (SOT) string method to analyze charge translocation pathways across single-component lipid bilayers as a function of multiple collective variables. We first demonstrate that the SOT string method can reproduce the free energy barrier for the translocation of a charged lysine amino-acid analogue in good agreement with previous literature. We then obtain minimum free energy pathways for the translocation, or flipping, of charged peptide loops across the lipid bilayer by utilizing trajectories from prior temperature-accelerated molecular dynamics (TAMD) simulations as initial configurations. Corresponding potential of mean force calculations reveal that the protonation of a central membrane-exposed aspartate residue substantially reduces the free energy barrier for flipping charged loops by modulating the water content of the bilayer. These results provide new insight

into the thermodynamics underlying loop-flipping processes and highlight how the combination of TAMD and the SOT string method can be used for understanding complex charge translocation mechanisms.

Introduction

The cell membrane separates the internal cytosolic environment of a cell from the extracellular environment and plays a vital role in controlling intracellular transport. The primary structural component of the cell membrane is the lipid bilayer. Nonionic molecules can favorably partition into and directly translocate across the lipid bilayer, but the hydrophobic core of the bilayer inhibits the partitioning of polar and charged molecules.¹ Consequently, the translocation of individual ions across the lipid bilayer is thought to occur on a long timescale (~minutes to hours) whereas larger species bearing multiple charges are not expected to translocate at all. In contrast to this expectation, some charged materials have been found to cross the bilayer over experimentally relevant timescales. For example, certain charged peptides²⁻⁵ and small nanoparticles⁶⁻¹⁰ can passively translocate across the bilayer and into cells despite bearing multiple charged moieties. Recent experiments have also shown that α -helical peptides containing a hydrophobic transmembrane domain flanked by water-soluble loops with multiple charged residues can interconvert from a state in which the peptide is adsorbed to the surface of a single-component lipid bilayer to a state in which the peptide embeds within the bilayer on a timescale of a few seconds.¹¹ This process requires the charged flanking loops to cross the hydrophobic bilayer core, which we refer to as “loop-flipping.” Similar loop-flipping processes have even been identified for integral membrane proteins containing multiple helical domains and loops; notably, multiple loops in the proteins LacY and EmrE are thought to flip across the bilayer.¹²⁻¹⁴ While substantial work has identified mechanisms by which ions or charged small-molecules can cross the bilayer,¹⁵⁻¹⁷ little is known regarding the mechanisms underlying loop-flipping or related processes in larger, more complex molecules.

Molecular dynamics (MD) simulations have the potential to provide mechanistic understanding of loop-flipping or other charge translocation processes. However, a major challenge is the balance between the need for atomistic simulations to capture the interplay of lipids, charged groups, and water molecules during charge translocation and the limited timescale achievable by these simulations. Because unbiased atomistic MD simulations can only model timescales on the order of μs , loop-flipping is a rare event that cannot be observed without specialized sampling techniques. Consequently, a range of enhanced sampling techniques, including umbrella sampling¹⁵, metadynamics¹⁸, milestoning¹⁹, the adaptive biasing force method²⁰, and replica-exchange umbrella sampling¹⁸, have been applied to investigate the translocation of charged molecules across lipid bilayers.²¹⁻²² Most of these techniques rely on biasing a predefined set of collective variables (CVs), defined as variables that can be expressed as a function of the Cartesian coordinates of a set of atoms, in order to calculate a potential of mean force (PMF) in the CV space. The distance between the center of mass of the translocating molecule from the bilayer center projected onto the bilayer normal is widely used as a single CV in umbrella sampling and other free energy calculation methods and is sufficient to estimate PMFs for the translocation for small nonionic molecules.²²⁻²⁶ However, additional CVs are needed for computing PMFs for translocating small ions or charged amino acids across the lipid bilayer and for understanding the role of water defects/pores during charge translocation.^{15-17, 27-28}

Computing PMFs for more complex processes, such as loop-flipping for peptides or proteins, requires the selection of multiple CVs *a priori* that can adequately describe the given process. For example, these materials could have multiple charged residues that could separately translocate, requiring multiple CVs to describe.²⁹ Translocation processes involving multiple charged groups may also be impacted by molecular conformation, interactions with other charged

species, or the local reorganization of lipid and water molecules. Hence, capturing the correct free energy barrier for loop-flipping becomes increasingly difficult if using just one or two CVs, and choosing poor CVs could lead to an unphysical estimate of the corresponding free energy barrier.³⁰ Increasing the number of CVs may better describe complex loop-flipping or charge translocation processes, but the computational expense of most common enhanced sampling techniques (*e.g.*, umbrella sampling) is too large to permit the systematic exploration of more than two or three CVs. As a result, new methodologies are needed to explore charge translocation processes involving complex molecules with multiple charged groups.

Previously, we utilized temperature-accelerated molecular dynamics (TAMD) simulations to explore loop-flipping mechanisms for peptides with a single hydrophobic domain and charged flipping loops.³¹ In a TAMD simulation, the values of a set of CVs are dynamically varied during a simulation by representing the CVs as fictitious particles that slowly diffuse according to their own equations of motion at a high fictitious temperature.³² The fictitious particles will cross free energy barriers in the system and, by coupling the values of the CVs to the real system via harmonic potentials, bias the system to explore the free energy landscape. TAMD offers the ability to simultaneously bias a large number of CVs to overcome the limitations discussed above and moreover permits exploration of a free energy landscape without predefining specific values of CVs to be sampled. These features were exploited in our past work by biasing ten CVs simultaneously to identify multiple peptide loop-flipping mechanisms, including a mechanism in which flipping is promoted by water defects that form due to membrane-exposed charged residues.³¹ However, a limitation of TAMD is that a PMF cannot generally be obtained when sampling multiple CVs, although some variants of TAMD can obtain free energies for small numbers of CVs.³³ A simple way to estimate the free energy barrier associated with a process is

based on calculated attempt frequencies,³⁴ but this approach requires an assumption for the transition state in the system and does not provide details on the complete PMF.

In this work, we apply the string method³⁵, implemented using the swarms-of-trajectories (SOT) approach³⁶, to identify minimum free energy paths (MFEPs) for charge translocation and loop-flipping in a high-dimensional CV space and calculate corresponding free energy barriers. The MFEP is the path through a high-dimensional CV space that has the highest likelihood of a transition between two free energy minima. Once identified, the distance along the MFEP can then be used as reaction coordinate to obtain a PMF. We first demonstrate that the SOT string method predicts the PMF for the translocation of a charged lysine residue in good agreement with prior literature estimates using two CVs. We then apply this method to determine MFEPs using 11 CVs to describe the loop-flipping of a peptide with a hydrophobic transmembrane domain and two charged flanking loops. Using initial trajectories generated from TAMD, we find that the string method identifies distinct MFEPs for loop-flipping depending on the protonation of a membrane-exposed residue, highlighting the complexity of loop-flipping mechanisms even for single peptides. Moreover, corresponding free energy barriers qualitatively agree with estimates from TAMD, suggesting that TAMD estimates can be used as a first approximation for comparing the free energy barriers for various peptide sequence mutations. These results provide new insight into how TAMD and the SOT string method can be used in combination to explore complex processes for the translocation of charged groups across lipid bilayers and indicate that this combination of methods can be useful for studying loop-flipping processes in more complex systems (*e.g.*, membrane proteins).

Methods

Simulation systems and parameters

We constructed lipid bilayers from 1,2-dioleoyl-*sn*-glycero-3-phosphatidylcholine (DOPC) as a model lipid because phosphatidylcholine (PC) lipids account for >50% of the phospholipids in most eukaryotic membranes and DOPC is in the fluid phase at room temperature.³⁷ We simulated systems containing one of two different charged molecules: either a charged analogue of a lysine side chain or an α -helical peptide with the sequence KKLALALLDWLLLLALALKK. The lysine analogue was studied to present baseline results for the translocation of a charged species. The peptide sequence was chosen because it has been shown to exhibit loop-flipping behavior in past computational³¹ and experimental studies.¹¹ The peptide was modeled in two states, with either the central aspartate residue charged (K4D-c) or neutral (K4D-n) to represent two extremes of pH explored in past studies^{11, 31}. The peptide has an α -helix formed by the central 17 residues, the N-terminus was acetylated, and the C-terminus was amidated. Peptides with neutral and charged aspartate residues were embedded within separate DOPC bilayers with a total of 177 and 175 lipids, respectively, obtained from our previous work.³¹ The CHARMM-GUI Membrane Builder³⁸⁻³⁹ was utilized to embed the lysine analogue within a DOPC lipid bilayer containing 128 lipids. The CHARMM36 force field was used for the lipids and peptide. The lysine, neutral aspartate peptide and charged aspartate peptide systems were solvated using 7644, 8769, and 8786 TIP3P water molecules and chloride ions were added to achieve charge neutrality.

The lysine-DOPC system was energy minimized and equilibrated in the *NPT* ensemble for 50 ns. The peptide-DOPC systems were previously equilibrated and obtained from previous work.³¹ All molecular dynamics simulations were performed using a leapfrog integrator with a 2-fs time step. Verlet lists were generated using a 1.2 nm neighbor list cutoff. van der Waals

interactions were modeled with a Lennard-Jones potential using a 1.2 nm cutoff that was smoothly shifted to zero between 1.0 and 1.2 nm. Electrostatic interactions were calculated using the smooth particle mesh Ewald method with a short-range cutoff of 1.2 nm, grid spacing of 0.12 nm, and fourth order interpolation. Bonds were constrained using the LINCS algorithm. For equilibration, the temperature was maintained at 298.15 K using a Berendsen thermostat with a time constant of 1.0 ps and the pressure was maintained at 1 bar using a semi-isotropic Berendsen barostat with a time constant of 5 ps and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. The energy minimization and the equilibration steps were performed using Gromacs 2016.⁴⁰

All SOT string method simulations were implemented using a combination of Gromacs 5.1 patched with PLUMED 2.3⁴¹ and python3. For all SOT string method simulations, the temperature was maintained at 298.15 K using a velocity-rescale thermostat with a time constant of 1.0 ps and the pressure was maintained at 1 bar using a semi-isotropic Parrinello-Rahman barostat with a time constant of 5 ps and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. Analysis of the trajectories were performed using python3 and mdtraj.⁴²

Collective variables

Separate sets of collective variables (CVs) were biased in the lysine-DOPC and peptide-DOPC systems. Two CVs (illustrated in Figure 1A) were used for the lysine-DOPC system: the distance between the center-of-mass of the lysine analogue and the bilayer midplane along the bilayer normal (referred to as the z -distance) and the water-lysine coordination number, $C_{\text{water-lys}}$. The z -distance was selected because it is a typical CV biased in studies of charge translocation.²²⁻²⁴ $C_{\text{water-lys}}$ quantifies the number of water molecules around the lysine analogue and was selected

to bias the formation of a water defect in the membrane as suggested by prior studies.⁴³⁻⁴⁴

$C_{water-lys}$, was computed using the PLUMED v2.3 package⁴¹, described by:

$$C_{water-lys} = \sum_{i \in water} \sum_{j \in lysine} s_{ij} \quad (1)$$

where s_{ij} is a switching function defined by:

$$s_{ij} = \frac{1 - \left(\frac{r_{ij}}{r_o}\right)^n}{1 - \left(\frac{r_{ij}}{r_o}\right)^m} \quad (2)$$

with $n = 12$, $m = 24$ and $r_o = 0.315$ nm.

11 CVs (illustrated in Figure 1B) were used for the peptide-DOPC systems: three CVs corresponding to the x -, y -, and z -positions of the center of mass of the helix, six CVs corresponding to the x -, y -, and z -positions of the centers of mass of the two flanking loops, the peptide tilt angle, and the peptide helicity. The peptide tilt angle was defined as the angle between the bilayer normal (*i.e.*, the z -axis of the simulation box) and a vector between the centers of mass of the two flanking loops. The helicity was calculated using the *alpharmsd* CV defined by PLUMED (details of the helicity CV are provided in the SI). These CVs were chosen based on the peptide geometry and can be extended to more complex proteins with multiple transmembrane helices and loops. All 11 CVs were biased in the SOT string method calculations (detailed below), which were initialized from prior TAMD simulations³¹ in which 10 of the CVs, excluding the peptide helicity, were biased. In these prior TAMD simulations, the helicity was instead restrained to prevent unfolding based on the large unfolding free energy penalty and prior simulation observations that peptide insertion occurs without unfolding.⁴⁵⁻⁴⁷

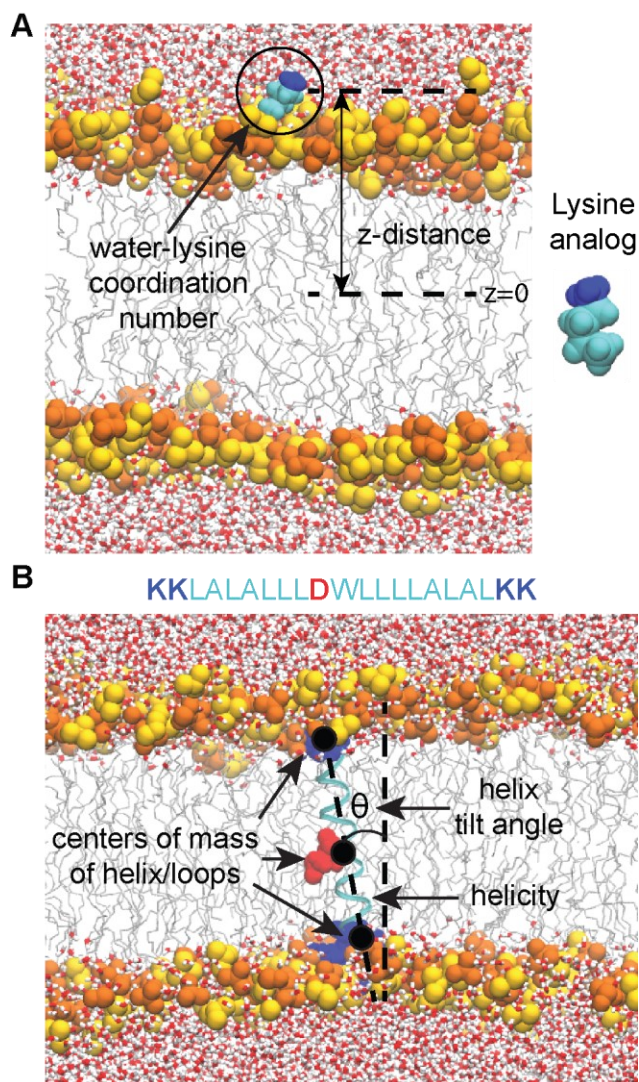


Figure 1. A) Snapshot of lysine side-chain analogue in a DOPC lipid bilayer and illustrations of the two CVs. **B)** Snapshot of an example peptide system in a DOPC lipid bilayer and illustration of the 11 CVs used. The cyan ribbon represents the α -helix of the peptide. The aspartate (red) and lysine (blue) residues are shown in a van der Waals representation. Lipid head groups are shown in yellow and orange, lipid tails in silver, and water in red/white.

Swarms-of-trajectories string method implementation

The SOT string method is briefly described here; a detailed description of the method is provided elsewhere.³⁶ The SOT string method estimates a MFEP through a high-dimensional CV space by determining a set of discrete images (*i.e.*, points in CV space) along that path. Large numbers of unbiased MD simulations initiated from these images are used to evolve images toward the MFEP.

Importantly, images are constrained to be equidistant along the string to ensure that images do not fall into the two free energy minima. The steps that we perform to implement this method are summarized here and illustrated schematically in Figure 2, then further detailed below:

- 1) A set of initial images for the discretized string is generated.
- 2) Replicas of the entire system are restrained to the CV values corresponding to each image.
- 3) 100 short unrestrained simulations (*i.e.*, a swarm of trajectories) are performed for each image. Simulations are initiated from different configurations and velocities with the same CV values obtained from the restrained.
- 4) For each image, the final values of the set of CVs are averaged across all the trajectories in the swarm to compute the drift in CV space.
- 5) The string is updated by adding the drift in each CV for each image.
- 6) The string is re-parameterized to ensure that images are equidistant.
- 7) Steps 2-6 are repeated until the string converges.

This workflow was applied in two stages for both the lysine-DOPC and peptide-DOPC systems. In the first stage, a string was computed with images separated by large distances in CV space. In the second stage, a subset of these images spanning a region of CV space expected to contain a free energy barrier was selected, new images were generated by interpolating between these existing images, and the string method was performed again. This two-stage approach thus enabled refinement of the string in a region of interest. Once the refined string was obtained, the potential of mean force (PMF) along the string was estimated by sampling from restrained simulations and numerically integrating the product of the mean force and the tangent to the string at each image following previous work.³⁵

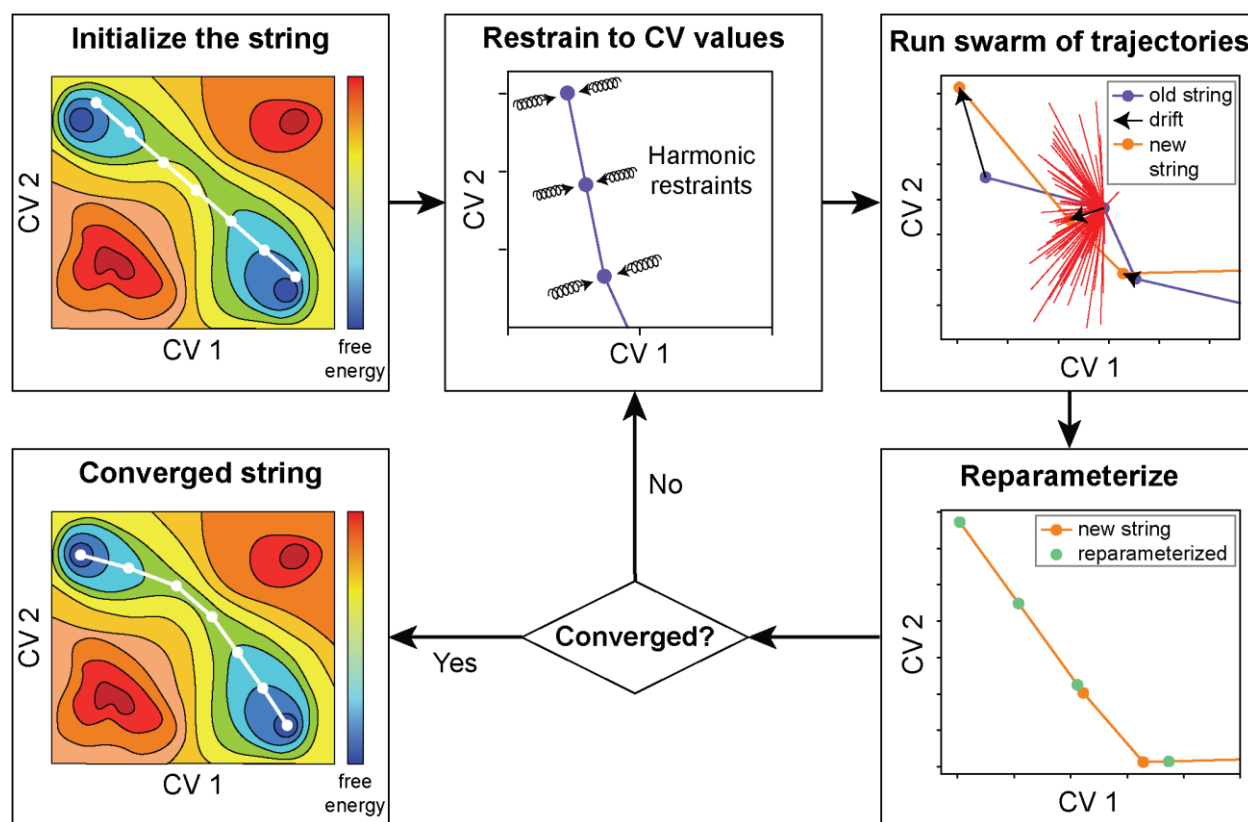


Figure 2. Schematic illustrating the workflow for the swarm-of-trajectories string method. Starting from an initial string connecting two states in a representative free energy landscape (top left), the string is iteratively evolved toward a minimum free energy path (bottom left) by initializing swarms of trajectories and updating the string according to calculated drifts in CV space.

Swarms-of-trajectories string method implementation for lysine-DOPC system

The initial string for the lysine-DOPC system was generated by pulling the center of mass of the lysine analogue along the bilayer normal using a harmonic restraint with a force constant of $717 \text{ kcal mol}^{-1} \text{ nm}^{-2}$. 21 initial string images were obtained, each separated by 0.25 nm along the z -distance CV from -2.5 to 2.5 nm with 0 nm corresponding to the bilayer center. In each iteration of the SOT string method, CVs were restrained to their desired values by applying a weak harmonic potential for 10 ps then a strong harmonic potential for 100 ps. For the z -distance CV, the spring constant was $40 \text{ kcal mol}^{-1} \text{ nm}^{-2}$ for the weak restraint and $1000 \text{ kcal mol}^{-1} \text{ nm}^{-2}$ for the strong restraint. For the $C_{\text{water-lys}}$ CV, the spring constant was $1 \text{ kcal mol}^{-1} \text{ rad}^{-2}$ for the weak

restraint and $8 \text{ kcal mol}^{-1} \text{ rad}^{-2}$ for the strong restraint. A swarm of 100 20-step (0.04 ps) unrestrained simulations were next performed for each image, the average drift in each CV was computed and added to the prior CV values to update the string, then the string was reparameterized. 50 iterations of this procedure were performed to obtain a converged string (see SI Figure S1). The initial string for the second stage was generated by selecting a smaller subset of images from the converged string obtained during the first stage that corresponded to configurations symmetric around the center of the bilayer with $-2 \text{ nm} < z < 2 \text{ nm}$. We linearly interpolated CV values between these images to obtain a denser population of 23 images as a new initial string. The workflow described above was repeated with the images corresponding to the ends of the string fixed in CV space. For every iteration, a swarm of 100 10-step (0.02 ps) unrestrained simulations was performed for each image and 40 iterations were sufficient for convergence of the refined string. The PMF along the refined string was then calculated using at least 20 ns restrained simulations for each image.

Swarms-of-trajectories string method implementation for peptide-DOPC system

The initial string for the peptide-DOPC system was generated by selecting 24 configurations from a TAMD loop-flipping trajectory from our previous work.³¹ In each iteration of the SOT string method, CVs were restrained to their desired values by applying a weak harmonic potential for 10 ps then a strong harmonic potential for 100 ps. For CVs corresponding to Cartesian coordinates, the spring constant was $350 \text{ kcal mol}^{-1} \text{ nm}^{-2}$ for the weak restraint and $8000 \text{ kcal mol}^{-1} \text{ nm}^{-2}$ for the strong restraint. For the peptide tilt angle, the spring constant was $21,875 \text{ kcal mol}^{-1} \text{ rad}^{-2}$ for the weak restraint and $50\,000 \text{ kcal mol}^{-1} \text{ rad}^{-2}$ for the strong restraint. For the helicity, the spring constant was $87.5 \text{ kcal mol}^{-1}$ for the weak restraint and $2000 \text{ kcal mol}^{-1}$ for the strong restraint. Upper and lower walls were also used to avoid peptide desorption from

the lipid bilayer surface to specifically study the loop-flipping process following our past work.³¹ A swarm of 100 500-step (1 ps) unrestrained simulations was performed for each image, the average drift in each CV was computed, and the string was updated and reparameterized as described above. This process was repeated for 40 iterations to obtain a converged string. Like the lysine-DOPC system, we selected a subset of configurations from this final converged string and interpolated CV values to generate an initial string with 31 and 27 images for the second stage for K4D-n and K4D-c respectively. The workflow was then repeated with the ends of the string fixed in CV space. For every iteration, a swarm of 100 250-step (0.5 ps) unrestrained simulations was performed for each image and 10 iterations was sufficient for convergence of the refined string. The PMF along the refined string was then calculated using 25 and 20 ns restrained simulations per image for K4D-n and K4D-c respectively.

Results and Discussion

String method analysis of the translocation of a lysine analogue across a DOPC bilayer

We first sought to compute the PMF for the translocation of a lysine analogue across a DOPC bilayer (referred to as lysine-DOPC translocation; see Figure 1A) to assess the accuracy of the SOT string method approach compared to more conventional umbrella sampling calculations (SI Section S2) and prior literature calculations⁴⁸. This PMF further serves as a benchmark for comparison to PMFs for the flipping of charged peptide loops containing lysine residues. While the translocation of charged molecules is often described by a single CV – typically the z -distance of the molecule center of mass from the bilayer center – this approach can lead to hysteresis associated with slow sampling of orthogonal degrees of freedom and produce unrealistically asymmetric PMFs.^{18, 30, 49} SI Figure S2 shows an example of an asymmetric PMF for lysine-DOPC translocation obtained from umbrella sampling calculations using only the z -distance CV (see

Methods). This asymmetry arises from the formation of a transient water defect around the lysine analogue as it crosses the bilayer that produces a favorable local solvation environment¹⁹. Because formation of the water defect is not biased by the z -distance CV, the asymmetry of the PMF reflects higher-energy defect structures that persist as a charge is translocating across the bilayer (depending on the method of system preparation). One approach to address this issue is to include an additional CV to bias water defect formation and capture the correct free energy barrier for charge translocation. Biasing two CVs using techniques like umbrella sampling or metadynamics can lead to a large computational cost due to the need to sample a larger CV space for convergence; conversely, the string method requires minimal additional computational cost to bias multiple CVs.

We calculated the PMF for lysine-DOPC translocation using 2 CVs, including both the z -distance and $C_{water-lys}$ (see Methods), with the latter CV used to bias the formation of a water defect. Figure 3A shows the convergence of the string during the 40 iterations performed for the second stage of the SOT string method calculations. The final string is nearly symmetric with respect to the center of the bilayer, as expected based on the symmetry of the system. $C_{water-lys}$ is higher at the ends of the string, corresponding to positions near the bilayer-water interface, and decreases in the middle of the bilayer but remains finite due to the formation of a water defect as illustrated by the simulation snapshots in Figure 3B. Figure 3B further shows the PMF computed along the string. One additional image was added between images 11 and 12 by linearly interpolating between the CV values for these two images; this image permitted further resolution of the free energy maximum. We observe that the PMF is nearly symmetric around a maximum near $z = 0$, as expected (SI Figure S2). The calculated free energy barrier for lysine-DOPC translocation from the SOT string method simulations is 12.27 kcal/mol, which is comparable to the 12.67 kcal/mol free energy barrier estimated from 1D umbrella sampling (SI Figure S2). This

value is also comparable to previous literature values for lysine analogue translocation in DPPC lipid bilayer calculated using 1D umbrella sampling and the same force field.⁴⁸

Together, these results demonstrate that the SOT string method can identify the PMF for charge translocation when leveraging previous knowledge of appropriate CVs without requiring computationally expensive free energy landscape calculation techniques (*e.g.*, 2D umbrella sampling). Specifically, we obtained this symmetric profile by using 2 simple CVs with a total of $\sim 0.7 \mu\text{s}$ of simulation time (including 223 ns for convergence of the string and 480 ns for convergence of the PMF). For comparison, we were unable to obtain symmetric PMF profile using 1D umbrella sampling as shown in SI Figure S2 at a similar computational cost, highlighting the value of biasing multiple CVs with limited additional computational effort.

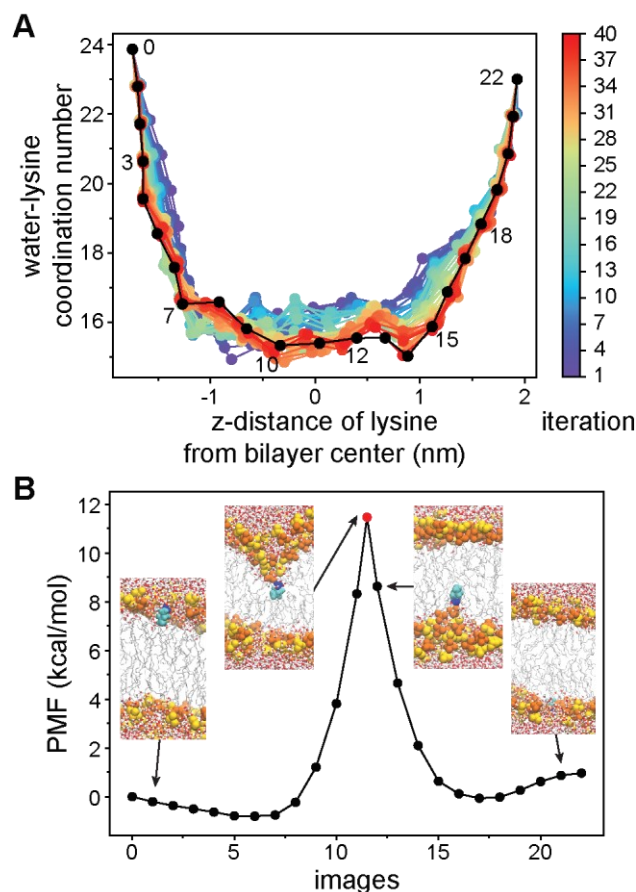


Figure 3. A) String convergence for lysine-DOPC translocation using 2 collective variables. The black line and points indicate the final converged string. Images are numbered in the plot. **B)** Potential of mean force (PMF) for the final string (black line in A). An additional image is included between images 11 and 12 to resolve the PMF maximum. Simulation snapshots show the relative orientation of the lysine analogue with respect to the lipid bilayer for representative images.

String method analysis of charged peptide loop-flipping

We next applied the SOT string method for determine free energy barriers for the loop-flipping of charged peptides, focusing on the K4D-n and K4D-c peptides studied in our previous work.³¹ In prior TAMD simulations, K4D-n and K4D-c primarily exhibited two distinct loop-flipping mechanisms that we refer to as the ‘direct’ and ‘cooperative’ mechanisms, respectively. In the ‘direct’ mechanism, the charged loop flipped across the bilayer by forming a water defect in the hydrophobic core of the bilayer in a similar manner as the lysine analogue (Figure 3). In the

‘cooperative’ mechanism, the charged loop flipped by interacting with a pre-formed water defect. The cooperative mechanism was identified for the K4D-c peptide because its charged central aspartate residue was exposed to the membrane and stabilized a pre-formed water defect prior to flipping. Conversely, the K4D-n peptide has a neutral central aspartate residue that does not stabilize a water defect, thus leading to the preferred direct mechanism. While the prior TAMD simulations revealed the preference for these mechanisms, they did not permit analysis of corresponding free energy barriers, motivating the application of the SOT string method.

For both peptides, initial configurations and corresponding CV values for the SOT string method were selected based on loop-flipping trajectories obtained from the prior TAMD simulations. Each trajectory was chosen to be representative of the primary loop-flipping mechanism (direct or cooperative) of the corresponding peptide. Configurations were selected to ensure that the CV corresponding to the z -distance between the flipping loop and the bilayer center spanned a range of values consistent with flipping, and moreover included key behaviors such as the interaction of the flipping loop with the pre-formed water defect that is key to the cooperative loop-flipping mechanism. Because 11 CVs were considered for the SOT string method, it is challenging to visualize the string directly. Figure 4 instead shows initial strings and simulation snapshots for the K4D-c and K4D-n peptides projected onto only two CVs (the z -distance and the peptide tilt angle) because these CVs clearly indicate the progression of the charged loop from one side of the bilayer to the other. The simulation snapshots show the transmembrane (image 0) and surface-adsorbed (image 23) states for both peptides, and moreover highlight the disruption of lipid head groups near the charged aspartate residue for the K4D-c peptide which we associate with a pre-formed water defect. Using the string method, we now explicitly calculate free energy

barriers for these loop-flipping mechanisms to determine the extent to which a pre-formed water defect (present in the cooperative mechanism) influences flipping free energy barriers.

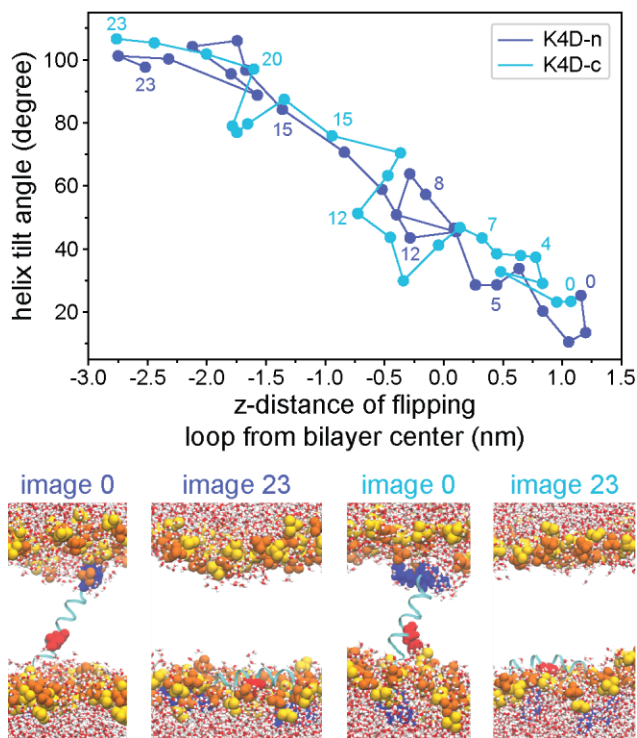


Figure 4. Example initial string (spanning a 11 CV space but drawn as a function of only 2 CVs) and simulation snapshots of the initial configurations of the start/end of the string for K4D-n (purple) and K4D-c (cyan). Images are numbered in the plot. The color scheme follows that of Figure 1 and lipid tails are not shown for clarity.

String method analysis of loop-flipping for K4D-n

We first present SOT string method results for the K4D-n peptide, which has a neutral central aspartate residue and flips across the bilayer via the direct mechanism described above. Figure 5A shows the final string for K4D-n loop-flipping (see SI Figure S3 for string convergence). The string was calculated using 11 CVs but is projected onto two CVs – the helix tilt angle and the z-distance of the flipping loop from the bilayer center – to facilitate visualization; consequently,

the images along the string do not appear equidistant due to variations in the values of the other 9 CVs. The string correctly converges with respect to all 11 CVs as shown in Section S4. Image 0 corresponds to a transmembrane state for which the z -distance of the flipping loop is positive and the helix tilt angle is near 20° (*i.e.*, the helix is approximately aligned with the bilayer normal). With increasing image number (*i.e.*, progression along the string) the z -distance of the flipping loop decreases and eventually switches sign, indicating a flip across the bilayer midplane, while the helix tilt angle increases to approximately 90° at image 30. These values correspond to a surface-adsorbed state in which the charged loop is positioned at the opposite bilayer leaflet and the helix lies flat along the bilayer surface. The snapshots in Figure 5B illustrate configurations characteristic of similar states.

The corresponding PMF for the final string is shown in Figure 5B and permits analysis of the relationship between the string and corresponding free energy variations. Starting from a local free energy minimum associated with the transmembrane state (image 0), the PMF increases monotonically until image 8, which corresponds to the point at which the helix center of mass crosses the bilayer center (SI Figure S4I) for this particular loop-flipping pathway. After image 8, all charged residues are close to the aqueous medium at the bilayer interface leading to a plateau in the PMF that persists until image 15. At this point, the flipping loop enters the core of the bilayer, bringing in a water defect and leading to a further increase in the PMF. After image 23, the water defect relaxes because the loop has flipped to the other side of the bilayer and the slope of the PMF decreases. After image 26, the peptide has completely flipped across the bilayer to reach the surface-adsorbed state and an associated local free energy minimum. Trends in the PMF thus reflect variations in the positions in the helix and charged loop as expected.

In total, the free energy difference between the transmembrane state (image 0) and surface-adsorbed state (image 30) is 45.53 kcal/mol, indicating that the transmembrane state is substantially more favorable. For comparison, the total water-to-bilayer transfer free energy for the 17 amino-acid residues in the peptide transmembrane domain, estimated using the side-chain hydrophobicity scale of Moon and Fleming,⁵⁰ is -43.37 kcal/mol, in reasonable agreement with the string method PMF. We note that this hydrophobicity scale estimate assumes a charged aspartate side chain and hence is a slight underestimate. Based on the PMF, the free energy barrier for K4D-n loop-flipping is 49.17 kcal/mol. This barrier is substantially greater than the barrier to the translocation of a single lysine analogue (~12 kcal/mol as shown in Figure 3). This result suggests that there is a much larger perturbation to the lipid bilayer compared to the perturbation associated with lysine translocation, as observed in the snapshots shown in Figure 5B, which leads to the large free energy barrier. The magnitude of this barrier reinforces that the flipping of the charged loops via the direct mechanism is extremely rare.

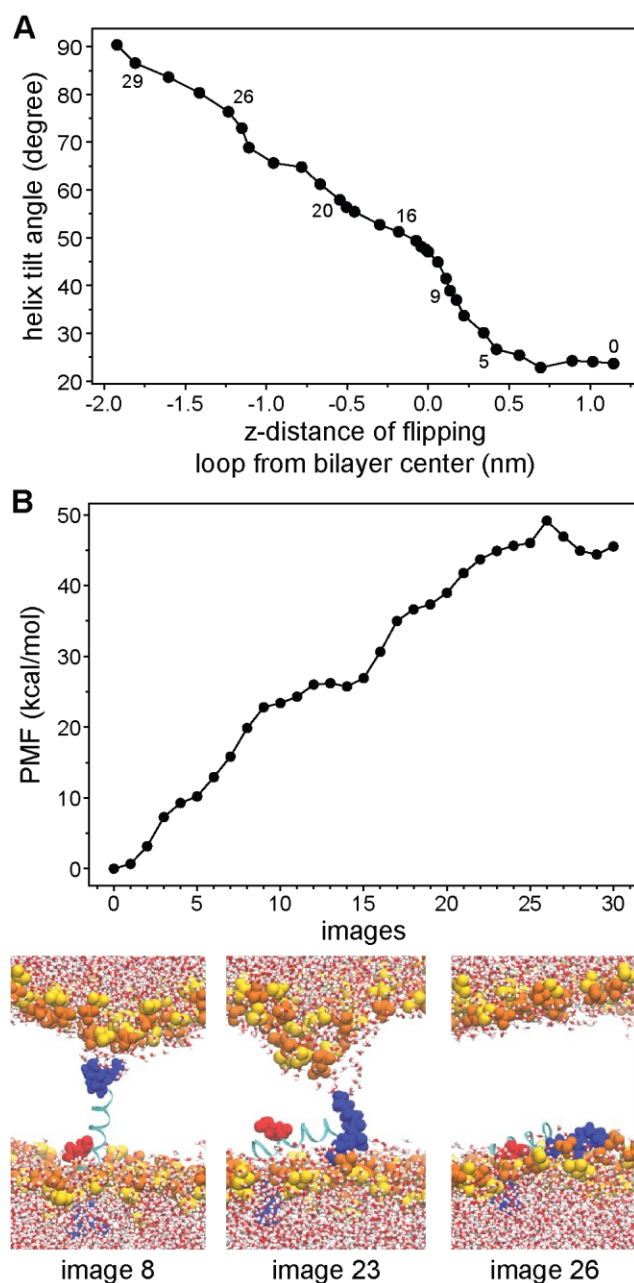


Figure 5. A) Final string (spanning a 11 CV space but drawn as a function of only 2 CVs) for K4D-n loop-flipping across DOPC bilayer via the direct mechanism. Images are numbered in the plot and do not appear equidistant because only two CVs are shown. **B)** Potential of mean force for the final string and representative simulation snapshots. The color scheme follows that of Figure 1 and lipid tails are not shown.

String method analysis of charged peptide loop-flipping for K4D-c

We next present SOT string method results for the K4D-c peptide, which has a charged central aspartate residue and flips across the bilayer via the cooperative mechanism described above. Figure 6A shows the final string after convergence (see SI Figure S5 for string convergence) for K4D-c loop-flipping. The transmembrane state (image 0) has comparable values of the z -distance and helix tilt angle CVs as the transmembrane state of the K4D-n peptide (Figure 5). The string passes through similar values of the tilt angle and z -distance as the string for the K4D-n peptide, but in general the helix tilt angle samples slightly lower values than the string in Figure 5A. The difference in tilt angle is most pronounced for the end of the string; for example, the tilt angle is approximately 70° for image 26 as is also observable in the snapshots in Figure 6B.

Figure 6B shows the PMF calculated using the final converged string. Compared to the results for K4D-n in Figure 5, the magnitude of the free energy difference between the transmembrane and surface-adsorbed states of 10.12 kcal/mol is substantially reduced, which can be attributed to two factors. First, the charged aspartate residue destabilizes the transmembrane state. Second, the reduced tilt angle associated with the surface-adsorbed state causes the hydrophobic residues of the peptide to deeply embed in the hydrophobic core of the bilayer while the charged aspartate residue is favorably coordinated by the zwitterionic lipid head groups, stabilizing the surface-adsorbed state. Supporting this point, the total interface-to-bilayer transfer free energy for the 17 amino-acid residues in the peptide transmembrane domain, estimated based on recent updates to the Fleming scale to capture the effect of the bilayer interface⁵¹, is -9.97 kcal/mol, in better agreement with the K4D-c result. This comparison suggests that the surface-adsorbed state for the K4D-c peptide corresponds to the water-bilayer interface whereas the surface-adsorbed state for the K4D-n peptide is more similar to bulk water, with the difference

possibly reflecting the stronger association of the charged aspartate residue of the K4D-c peptide with lipid head groups. We also note that the helicity of the K4D-n is slightly higher for the surface-adsorbed state compared to that of K4D-c (see SI Figure S4), which might also influence the stability of the surface-adsorbed state.

The free energy barrier for loop-flipping from the transmembrane state at image 0 to the surface-adsorbed state is 14.64 kcal/mol, which is comparable to the barrier for the translocation of the lysine analogue. The lower free energy barrier compared to that of K4D-n can be explained by the cooperative interaction between the flipping lysine loop and the central charged aspartate in the transmembrane helix, as hypothesized in our previous work.³¹ The PMF maximum corresponds to image 15, at which point the flipping loop is near the bilayer midplane. The snapshot in Figure 6B indicates that in this configuration the bilayer is already locally perturbed by the charged aspartate residue (*i.e.*, due to the pre-formed water defect) and the flipping loop interacts with this local perturbation. The presence of this local perturbation even in the initial image (image 0) thus indicates that less of a free energy penalty is induced during loop-flipping, agreeing with our previous hypothesis. Similar cooperative interactions with water defects have also been reported to reduce the free energy barrier for the translocation of related charged molecules.^{19, 45, 52-53}

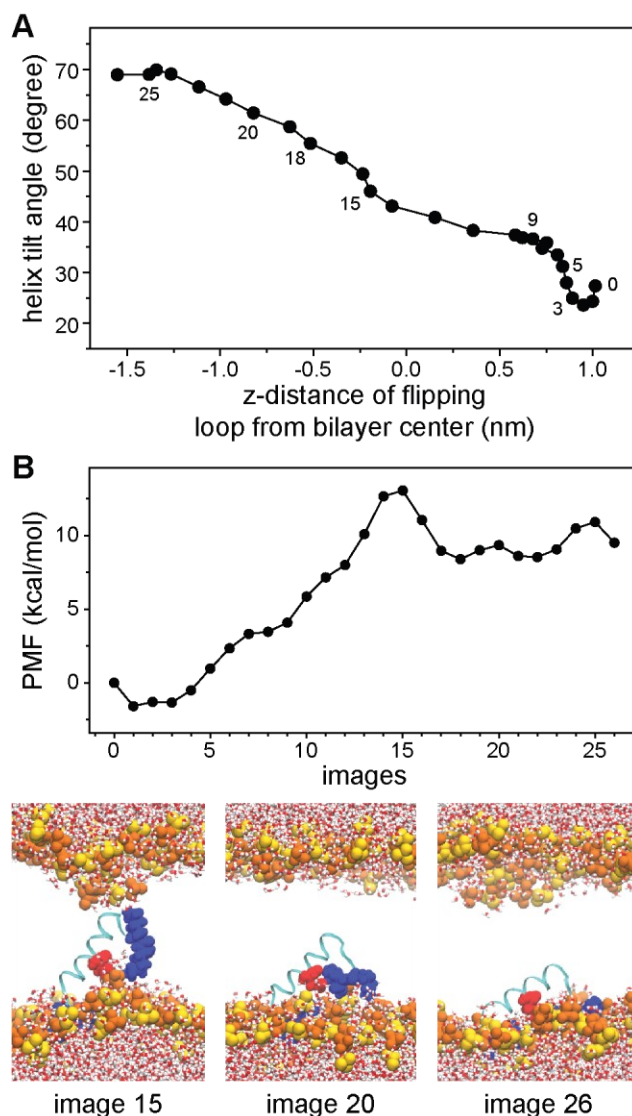


Figure 6. A) Final string (spanning a 11 CV space but drawn as a function of only 2 CVs) for K4D-c loop-flipping across DOPC bilayer via the cooperative mechanism. Images are numbered in the plot and do not appear equidistant because only two CVs are shown. **B)** Potential of mean force for the final string and representative simulation snapshots. The color scheme follows that of Figure 1 and lipid tails are not shown.

Finally, we repeated the SOT string method and PMF calculations of K4D-c loop-flipping for a separate set of initial images obtained from a different TAMD loop-flipping trajectory. The calculation was repeated to determine how the selection of initial CV values would impact the final strings. Here, it is worth highlighting that the string method identifies a MFEP that depends on local features of the free energy landscape and corresponds to one possible high-transition

probability of path out of multiple such paths.⁵⁴⁻⁵⁵ Consequently, repetition of the calculation using different initial conditions permits comparison of distinct paths through the high-dimensional CV space associated with loop-flipping. SI Figure S6 shows the second string and corresponding PMF, which exhibit many of the same features found in Figure 6. While the string itself passes through distinct CV values, key events (*e.g.*, the interaction of the flipping loop with the water defect) are preserved, as is the reduced loop-flipping free energy barrier relative to the neutral K4D-n. The primary difference is variations in the PMF associated with loop-lipid interactions prior to flipping, pointing to the complexity of the interactions that underlie peptide loop-flipping events. Further discussion is included in Section S6.2.

PMF differences reflect differences in average water content of the bilayer

To further investigate how the local variation in water density influences the thermodynamics of loop-flipping, we plotted the average number of water molecules near the peptide (within 4 nm of the aspartate residue center of mass as measured in the *xy* plane) and in the hydrophobic core of the lipid bilayer (defined by $-0.8 \text{ nm} < z < 0.8 \text{ nm}$). Figure 7 shows these values, calculated from the restrained simulations used to determine the PMFs, for both peptides. The K4D-n peptide leads to a consistent increase in the amount of water within the bilayer compared to K4D-c, which reflects a larger perturbation to the bilayer structure and overall increase in the free energy. For the K4D-c peptide, the images near the PMF maximum have a peak in the average water content of the bilayer, again highlighting that the free energy barrier for loop-flipping largely depends on the water molecules in the hydrophobic core of the lipid bilayer irrespective of the charge state of the aspartate residue. A similar result was obtained for the second SOT string method calculate for K4D-c peptide (SI Section S7).

As discussed in the introduction, selecting appropriate CVs to describe a charge translocation or loop-flipping process remains difficult *a priori*. In this work, we primarily selected CVs for peptide loop-flipping based on the peptide and loop geometry, which are intuitive because loop-flipping involves the movement of the loops and helix along the lipid bilayer normal and leads to a change in the helix tilt angle. From Figure 7, we observe that although we do not directly bias water molecules in this work, having a CV based on the interaction of water molecules with the peptide charged residues might be beneficial. This result is consistent with the analysis of the translocation of the lysine analogue, which showed that including the $C_{water-lys}$ CV reduces the hysteresis observed otherwise with a 1D PMF. Previous literature results have also shown benefits of incorporating water coordination-based CVs for obtaining PMFs for the translocation of charged or polar molecules across lipid bilayer.⁴³⁻⁴⁴ These results indicate that future work should consider further refinement of current CVs⁵⁶⁻⁵⁷ and alternative CVs that capture variations in local water density to augment geometric CVs when considering the translocation or flipping of complex molecules like the peptides studied here.

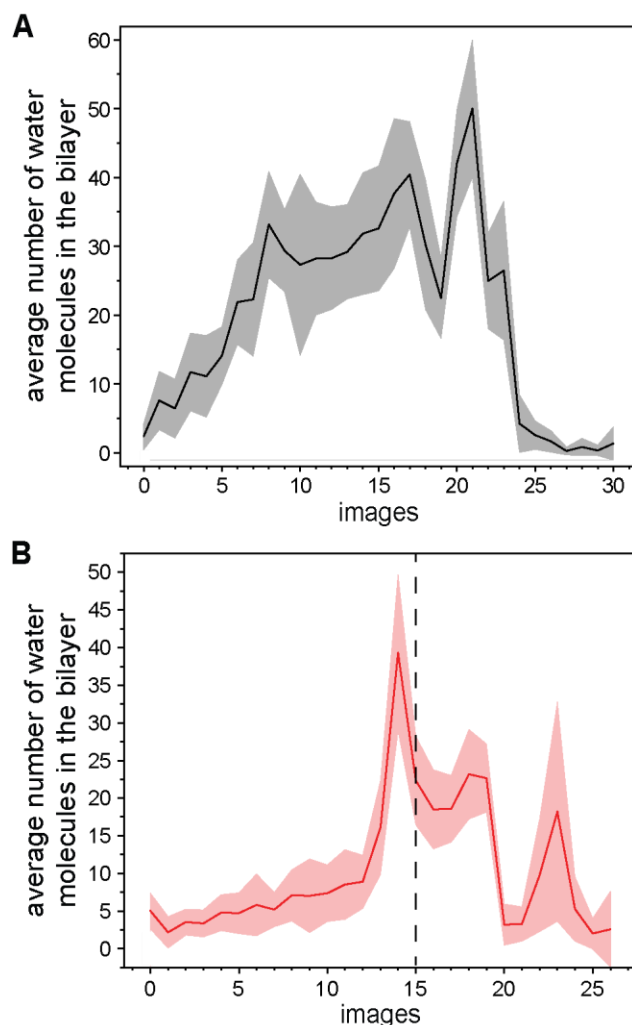


Figure 7. Average number of water molecules in the hydrophobic core of the lipid bilayer for the **A)** K4D-n and **B)** K4D-c peptides. Values were calculated from restrained simulations used to calculate the PMFs in Figures 5 and 6. Image numbers likewise correspond to the final strings in those figures. Shaded error bars represent the standard deviation. The vertical dashed line indicates the image corresponding to the PMF maximum for the K4D-c peptide.

Combination of TAMD and SOT string method for free energy barrier calculations

In our previous work, we used TAMD simulations to investigate loop-flipping mechanisms because TAMD permits the exploration of the free energy landscape without predefining a specific pathway. The prior use of TAMD is advantageous in this work because it permits the identification of initial states for the SOT string method that correspond to a high-probability transition between the transmembrane and surface-adsorbed states, facilitating string convergence as highlighted

previously⁵⁸. A limitation of TAMD is that it does not directly measure a free energy landscape. In our previous work, we roughly estimated free energy barriers based on calculated attempt frequencies³⁴. For those calculations, we estimated the bounds on the free energy barrier for K4D-n loop-flipping to be 31.5-33.9 kcal/mol and for K4D-c to be 21.6-23.7 kcal/mol. Conversely, the PMFs obtained in this work lead to free energy barriers of 49.17 kcal/mol for K4D-n loop-flipping and 14.64 kcal/mol for K4D-c loop-flipping. While the barriers estimated by the prior TAMD approach are not quantitatively equal to the barriers obtained from the SOT string method calculations, they follow correct trends, suggesting that even the simple TAMD estimates can resolve free energy differences associated with different peptide sequences. TAMD calculations can thus facilitate comparisons between barriers expected for different peptides prior to more detailed SOT string method calculations. We also note that the TAMD estimates are based on multiple TAMD loop-flipping events, whereas the free energy barrier calculated using the SOT string method could reflect local features of the CV space for that path. Another possible limitation of the SOT string method is that it might not resolve the maximum free energy barrier without denser sampling of CV space and without computing multiple paths, thereby increasing the computational expense of the method. Initializing multiple initial strings from different TAMD trajectories (as done here for the K4D-c peptide) could thus enable more robust free energy barrier estimates via repeated SOT string method calculations. In sum, the SOT string method in combination with TAMD is valuable for gaining insight on loop-flipping behavior and could be used to screen variations in loop-flipping free energy barriers for various peptide sequences in future work.

Conclusions

In this work, we apply the SOT string method to identify minimum free energy paths and corresponding free energy barriers associated with the translocation of charged species across a single-component DOPC lipid bilayer. We first show that this approach reproduces the free energy barrier for the translocation of a charged lysine analogue in agreement with past calculations. The SOT string method enables lysine translocation to be resolved as a function of two CVs without extensive additional sampling compared to conventional 1D umbrella sampling calculations. We then apply the SOT string method to study the flipping of charged peptide loops across the bilayer. In past work, we utilized TAMD simulations to identify loop-flipping pathways in a high-dimensional CV space without calculating free energy barriers.³¹ Using trajectories from the past TAMD simulations as initial configurations, we use the SOT string method to identify paths as a function of 11 CVs for both K4D-n (a peptide with a neutral membrane-exposed aspartate residue) and K4D-c (a peptide with a charged membrane-exposed aspartate residue). Corresponding free energy barriers qualitatively agree with past TAMD estimates. The loop-flipping free energy barrier for K4D-n is higher than the barrier for K4D-c, suggesting that a charged membrane-exposed group substantially reduces the free energy penalty of translocation. These findings highlight the general applicability of the string method to the study of charge translocation and provide molecular insight into factors influencing the translocation of charged peptides. We further envision applying the combination of TAMD and the SOT string method to study the translocation of species bearing multiple charged molecules, including charged nanoparticles⁶⁻¹⁰ or membrane proteins.¹²⁻¹⁴

Supporting Information

Convergence of all strings and PMFs, convergence of all CVs for peptide loop-flipping SOT string method calculations, details on the implementation of the helicity CV, data for a separate SOT string method simulation for K4D-c loop-flipping, and data on the average water density in bilayer core for K4D-c. Raw data and simulation codes are available at [10.5281/zenodo.4741259](https://zenodo.org/record/4741259).

Acknowledgements

This work was supported by the National Science Foundation under Grant No. MCB-1817292.

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1549562.

References

1. Alberts, B.; Bray, D.; Hopkin, K.; Johnson, A. D.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P., *Essential Cell Biology*. 4 ed.; Garland Publishing: Philadelphia, PA, 2013.
2. Cruz, J.; Mihailescu, M.; Wiedman, G.; Herman, K.; Searson, P. C.; Wimley, W. C.; Hristova, K., A Membrane-Translocating Peptide Penetrates into Bilayers without Significant Bilayer Perturbations. *Biophysical Journal* **2013**, *104* (11), 2419-2428.
3. Fuselier, T.; Wimley, W. C., Spontaneous Membrane Translocating Peptides: The Role of Leucine-Arginine Consensus Motifs. *Biophysical Journal* **2017**, *113* (4), 835-846.
4. Ablan, F. D. O.; Spaller, B. L.; Abdo, K. I.; Almeida, P. F., Charge Distribution Fine-Tunes the Translocation of Alpha-Helical Amphipathic Peptides across Membranes. *Biophysical Journal* **2016**, *111* (8), 1738-1749.
5. Wheaton, S. A.; Ablan, F. D. O.; Spaller, B. L.; Trieu, J. M.; Almeida, P. F., Translocation of Cationic Amphipathic Peptides across the Membranes of Pure Phospholipid Giant Vesicles. *Journal of the American Chemical Society* **2013**, *135* (44), 16517-16525.
6. Van Lehn, R. C.; Alexander-Katz, A., Pathway for Insertion of Amphiphilic Nanoparticles into Defect-Free Lipid Bilayers from Atomistic Molecular Dynamics Simulations. *Soft Matter* **2015**, *11* (16), 3165-3175.
7. Van Lehn, R. C.; Atukorale, P. U.; Carney, R. P.; Yang, Y. S.; Stellacci, F.; Irvine, D. J.; Alexander-Katz, A., Effect of Particle Diameter and Surface Composition on the Spontaneous Fusion of Monolayer-Protected Gold Nanoparticles with Lipid Bilayers. *Nano Lett* **2013**, *13* (9), 4060-4067.
8. Panja, P.; Jana, N. R., Arginine-Terminated Nanoparticles of < 10 Nm Size for Direct Membrane Penetration and Protein Delivery for Straight Access to Cytosol and Nucleus. *J Phys Chem Lett* **2020**, *11* (6), 2363-2368.

9. Jiang, Y.; Huo, S. D.; Mizuhara, T.; Das, R.; Lee, Y. W.; Hou, S.; Moyano, D. F.; Duncan, B.; Liang, X. J.; Rotello, V. M., The Interplay of Size and Surface Functionality on the Cellular Uptake of Sub-10 Nm Gold Nanoparticles. *Acs Nano* **2015**, *9* (10), 9986-9993.
10. Dubavik, A.; Sezgin, E.; Lesnyak, V.; Gaponik, N.; Schwille, P.; Eychmuller, A., Penetration of Amphiphilic Quantum Dots through Model and Cellular Plasma Membranes. *Acs Nano* **2012**, *6* (3), 2150-2156.
11. LeBarron, J.; London, E., Highly Hydrophilic Segments Attached to Hydrophobic Peptides Translocate Rapidly across Membranes. *Langmuir* **2016**, *32* (41), 10752-10760.
12. Woodall, N. B.; Hadley, S.; Yin, Y.; Bowie, J. U., Complete Topology Inversion Can Be Part of Normal Membrane Protein Biogenesis. *Protein Science* **2017**, *26* (4), 824-833.
13. Vitrac, H.; MacLean, D. M.; Jayaraman, V.; Bogdanov, M.; Dowhan, W., Dynamic Membrane Protein Topological Switching Upon Changes in Phospholipid Environment. *P Natl Acad Sci USA* **2015**, *112* (45), 13874-13879.
14. Van Lehn, R. C.; Zhang, B.; Miller, T. F., Regulation of Multispanning Membrane Protein Topology Via Post-Translational Annealing. *Elife* **2015**, *4*.
15. Guan, X. Q.; Wei, D. Q.; Hu, D., Free Energy Calculation of Transmembrane Ion Permeation: Sample with a Single Reaction Coordinate and Analysis Along Transition Path. *Journal of Chemical Theory and Computation* **2019**, *15* (2), 1216-1225.
16. Zhang, H. Y.; Xu, Q.; Wang, Y. K.; Zhao, T. Z.; Hu, D.; Wei, D. Q., Passive Transmembrane Permeation Mechanisms of Monovalent Ions Explored by Molecular Dynamics Simulations. *J Chem Theory Comput* **2016**, *12* (10), 4959-4969.
17. Cao, Z. X.; Bian, Y. Q.; Hu, G. D.; Zhao, L. L.; Kong, Z. Z.; Yang, Y. D.; Wang, J. H.; Zhou, Y. Q., Bias-Exchange Metadynamics Simulation of Membrane Permeation of 20 Amino Acids. *Int J Mol Sci* **2018**, *19* (3).
18. Pokhrel, N.; Maibaum, L., Free Energy Calculations of Membrane Permeation: Challenges Due to Strong Headgroup-Solute Interactions. *Journal of Chemical Theory and Computation* **2018**, *14* (3), 1762-1771.
19. Fathizadeh, A.; Kogan, M.; Anderson, C. M.; Webb, L. J.; Elber, R., Defect-Assisted Permeation through a Phospholipid Membrane: Experimental and Computational Study of the Peptide Wkw. *J Phys Chem B* **2019**, *123* (31), 6792-6798.
20. Hu, Y.; Ou, S. C.; Patel, S., Free Energetics of Arginine Permeation into Model Dmpc Lipid Bilayers: Coupling of Effective Counterion Concentration and Lateral Bilayer Dimensions. *Journal of Physical Chemistry B* **2013**, *117* (39), 11641-11653.
21. Martinotti, C.; Ruiz-Perez, L.; Deplazes, E.; Mancera, R. L., Molecular Dynamics Simulation of Small Molecules Interacting with Biological Membranes. *Chemphyschem* **2020**, *21* (14), 1486-1514.
22. Lee, C. T.; Comer, J.; Herndon, C.; Leung, N.; Pavlova, A.; Swift, R. V.; Tung, C.; Rowley, C. N.; Amaro, R. E.; Chipot, C., *et al.*, Simulation-Based Approaches for Determining Membrane Permeability of Small Compounds. *Journal of Chemical Information and Modeling* **2016**, *56* (4), 721-733.
23. MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P., Distribution of Amino Acids in a Lipid Bilayer from Computer Simulations. *Biophysical Journal* **2008**, *94* (9), 3393-3404.
24. Bennion, B. J.; Be, N. A.; McNerney, M. W.; Lao, V.; Carlson, E. M.; Valdez, C. A.; Malfatti, M. A.; Enright, H. A.; Nguyen, T. H.; Lightstone, F. C., *et al.*, Predicting a Drug's Membrane Permeability: A Computational Model Validated with in Vitro Permeability Assay Data. *Journal of Physical Chemistry B* **2017**, *121* (20), 5228-5237.
25. Loverde, S. M., Molecular Simulation of the Transport of Drugs across Model Membranes. *J Phys Chem Lett* **2014**, *5* (10), 1659-1665.
26. Jin, T. Y.; Patel, S. J.; Van Lehn, R. C., Molecular Simulations of Lipid Membrane Partitioning and Translocation by Bacterial Quorum Sensing Modulators. *Plos One* **2021**, *16* (2).

27. Hu, Y.; Sinha, S. K.; Patel, S., Investigating Hydrophilic Pores in Model Lipid Bilayers Using Molecular Simulations: Correlating Bilayer Properties with Pore-Formation Thermodynamics. *Langmuir* **2015**, *31* (24), 6615-6631.
28. Vorobyov, I.; Olson, T. E.; Kim, J. H.; Koeppe, R. E.; Andersen, O. S.; Allen, T. W., Ion-Induced Defect Permeation of Lipid Membranes. *Biophys J* **2014**, *106* (3), 586-597.
29. Kabelka, I.; Brozek, R.; Vacha, R., Selecting Collective Variables and Free-Energy Methods for Peptide Translocation across Membranes. *J Chem Inf Model* **2021**, *61* (2), 819-830.
30. Neale, C.; Pomes, R., Sampling Errors in Free Energy Simulations of Small Molecules in Lipid Bilayers. *Bba-Biomembranes* **2016**, *1858* (10), 2539-2548.
31. Patel, S. J.; Van Lehn, R. C., Characterizing the Molecular Mechanisms for Flipping Charged Peptide Flanking Loops across a Lipid Bilayer. *Journal of Physical Chemistry B* **2018**, *122* (45), 10337-10348.
32. Maragliano, L.; Vanden-Eijnden, E., A Temperature Accelerated Method for Sampling Free Energy and Determining Reaction Pathways in Rare Events Simulations. *Chem Phys Lett* **2006**, *426* (1-3), 168-175.
33. Abrams, C. F.; Vanden-Eijnden, E., On-the-Fly Free Energy Parameterization Via Temperature Accelerated Molecular Dynamics. *Chem Phys Lett* **2012**, *547*, 114-119.
34. Abrams, C. F.; Vanden-Eijnden, E., Large-Scale Conformational Sampling of Proteins Using Temperature-Accelerated Molecular Dynamics. *P Natl Acad Sci USA* **2010**, *107* (11), 4961-4966.
35. Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G., String Method in Collective Variables: Minimum Free Energy Paths and Isocommittor Surfaces. *Journal of Chemical Physics* **2006**, *125* (2).
36. Pan, A. C.; Sezer, D.; Roux, B., Finding Transition Pathways Using the String Method with Swarms of Trajectories. *Journal of Physical Chemistry B* **2008**, *112* (11), 3432-3440.
37. van Meer, G., Membrane Lipids, Where They Are and How They Behave: Sphingolipids on the Move. *Faseb J* **2010**, *24*.
38. Jo, S.; Kim, T.; Iyer, V. G.; Im, W., Charmm-Gui: A Web-Based Graphical User Interface for Charmm. *J Comput Chem* **2008**, *29* (11), 1859-65.
39. Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Davila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M., *et al.*, Charmm-Gui Membrane Builder toward Realistic Biological Membrane Simulations. *J Comput Chem* **2014**, *35* (27), 1997-2004.
40. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D., *et al.*, Gromacs 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29* (7), 845-854.
41. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., Plumed 2: New Feathers for an Old Bird. *Comput Phys Commun* **2014**, *185* (2), 604-613.
42. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernandez, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S., Mdtraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal* **2015**, *109* (8), 1528-1532.
43. Ghaemi, Z.; Minozzi, M.; Carloni, P.; Laio, A., A Novel Approach to the Investigation of Passive Molecular Permeation through Lipid Bilayers from Atomistic Simulations. *Journal of Physical Chemistry B* **2012**, *116* (29), 8714-8721.
44. Guan, X. Q.; Wei, D. Q.; Hu, D., Free Energy Calculations on the Water-Chain-Assisted and the Dehydration Mechanisms of Transmembrane Ion Permeation. *Journal of Chemical Theory and Computation* **2020**, *16* (1), 700-710.
45. Ulmschneider, J. P., Charged Antimicrobial Peptides Can Translocate across Membranes without Forming Channel-Like Pores. *Biophysical Journal* **2017**, *113* (1), 73-81.
46. Bereau, T.; Bennett, W. F. D.; Pfaendtner, J.; Deserno, M.; Karttunen, M., Folding and Insertion Thermodynamics of the Transmembrane Walp Peptide. *Journal of Chemical Physics* **2015**, *143* (24), 243127.

47. Ulmschneider, M. B.; Ulmschneider, J. P.; Schiller, N.; Wallace, B. A.; von Heijne, G.; White, S. H., Spontaneous Transmembrane Helix Insertion Thermodynamically Mimics Translocon-Guided Insertion. *Nature Communications* **2014**, *5*, 4863.
48. Sun, D. L.; Forsman, J.; Woodward, C. E., Evaluating Force Fields for the Computational Prediction of Ionized Arginine and Lysine Side-Chains Partitioning into Lipid Bilayers and Octanol. *Journal of Chemical Theory and Computation* **2015**, *11* (4), 1775-1791.
49. Neale, C.; Madill, C.; Rauscher, S.; Pomes, R., Accelerating Convergence in Molecular Dynamics Simulations of Solutes in Lipid Membranes by Conducting a Random Walk Along the Bilayer Normal. *J Chem Theory Comput* **2013**, *9* (8), 3686-703.
50. Moon, C. P.; Fleming, K. G., Side-Chain Hydrophobicity Scale Derived from Transmembrane Protein Folding into Lipid Bilayers. *P Natl Acad Sci USA* **2011**, *108* (25), 10174-10177.
51. Marx, D. C.; Fleming, K. G., Local Bilayer Hydrophobicity Modulates Membrane Protein Stability. *J Am Chem Soc* **2021**, *143* (2), 764-772.
52. MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P., Transfer of Arginine into Lipid Bilayers Is Nonadditive. *Biophysical Journal* **2011**, *101* (1), 110-117.
53. Van Lehn, R. C.; Alexander-Katz, A., Grafting Charged Species to Membrane-Embedded Scaffolds Dramatically Increases the Rate of Bilayer Flipping. *Acs Central Sci* **2017**, *3* (3), 186-195.
54. E, W. N.; Vanden-Eijnden, E., Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events. *Annu Rev Phys Chem* **2010**, *61*, 391-420.
55. Vanden-Eijnden, E.; Venturoli, M., Revisiting the Finite Temperature String Method for the Calculation of Reaction Tubes and Free Energies. *Journal of Chemical Physics* **2009**, *130* (19).
56. Peters, B.; Trout, B. L., Obtaining Reaction Coordinates by Likelihood Maximization. *J Chem Phys* **2006**, *125* (5), 054108.
57. Levintov, L.; Paul, S.; Vashisth, H., Reaction Coordinate and Thermodynamics of Base Flipping in Rna. *J Chem Theory Comput* **2021**, *17* (3), 1914-1921.
58. Vashisth, H.; Maragliano, L.; Abrams, C. F., "Dfg-Flip" in the Insulin Receptor Kinase Is Facilitated by a Helical Intermediate State of the Activation Loop. *Biophysical Journal* **2012**, *102* (8), 1979-1987.

TOC graphic

