Folding a viral peptide in different membrane environments: pathway and sampling analyses

Shivangi Nangia¹ · Jason G. Pattis¹ · Eric R. May¹

Received: 1 December 2017 / Accepted: 16 March 2018 / Published online: 11 April 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018, corrected publication April/2018

Abstract Flock House virus (FHV) is a well-characterized model system to study infection mechanisms in non-enveloped viruses. A key stage of the infection cycle is the disruption of the endosomal membrane by a component of the FHV capsid, the membrane active γ peptide. In this study, we perform all-atom molecular dynamics simulations of the 21 N-terminal residues of the γ peptide interacting with membranes of differing compositions. We carry out umbrella sampling calculations to study the folding of the peptide to a helical state in homogenous and heterogeneous membranes consisting of neutral and anionic lipids. From the trajectory data, we evaluate folding energetics and dissect the mechanism of folding in the different membrane environments. We conclude the study by analyzing the extent of configurational sampling by performing time-lagged independent component analysis.

Keywords Protein folding · Flock House virus · Molecular dynamics · Umbrella sampling · Non-enveloped virus · Membrane active peptides · TICA

1 Introduction

Understanding the thermodynamics of peptide association and folding in a membrane environment is critical to deciphering the underlying mechanisms of membrane disruption by membrane active peptides [1–7]. The transition from an unstructured solution state to an α -helical membrane bound state is a common trait of small amphipathic membrane proteins that have been researched extensively both experimentally [7–17] and computationally [18–29]. The thermodynamic driving forces for protein–membrane interactions and stabilization of the

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10867-018-9490-y>) contains supplementary material, which is available to authorized users.

✉ Eric R. May
eric.may@uconn.edu

¹ Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA

folded state are a delicate balance between enthalpic and entropic factors. Consequently, analyzing the peptide folding pathway and energetics in lipid bilayers can provide detailed insight into the biological activity of the peptide. A clear picture of the mode of membrane disruption employed by membrane active peptides has, in general, remained elusive [30, 31].

Over the last two decades computational studies have been instrumental in providing insight into lipid and protein dynamics by utilizing both equilibrium and biased sampling methodologies to study the energetics, thermodynamics, and structural dynamics of amphipathic membrane active proteins [18, 20, 23, 25–27, 32–36]. A major area of emphasis in understanding the mechanism of action of membrane active peptides is characterizing the initial stage of membrane association and peptide folding. Previous studies in this area include that of Brooks and coworkers who examined the folding dynamics of the designed WALP peptides in an implicit membrane model using temperature replica exchange molecular dynamics (T-REMD) [25]. Their studies showed that all variants of the WALP peptides penetrated the membrane with the N-terminal regions initiating the insertion and ultimately transitioning to an α -helix trans-membrane configuration consistent with experimental observations. More recently, molecular dynamics (MD) simulations pertaining to the folding and penetration of a single transmembrane WALP peptide were carried out in both all-atom and coarse-grained models [24]. The folding free-energy was determined in the coarse-grained representation as a function of helicity of the peptide using Hamiltonian REMD. Unbiased MD simulations have also been performed to study the folding dynamics of another widely studied membrane active peptide, the antimicrobial peptide melittin. These simulations have revealed that the peptide robustly associates with the membrane in a disordered state and attains helicity parallel to the surface of the membrane causing deformation of the bilayer as it folds [18]. Long time scale (17 μ s) unbiased MD simulations have shown that there is a narrow distribution of folded melittin conformers that partition into the membrane interface [23].

In our previous work [37], we investigated the thermodynamic aspects of binding and the structural dynamics of the FHV 21 N-terminal residues of the γ peptide (known as γ_1) using a multi-scale approach. We examined the binding and folding characteristics of γ_1 on pure phosphatidylcholine (PC), pure phosphatidylglycerol (PG) and a 50:50 mixed PC:PG membrane. Our findings from 1 μ s equilibrium all-atom simulations were in agreement with experimental measurements of the configuration of γ_1 on a PC bilayer, where we observed $\sim 70\%$ helicity of γ_1 [38]. On PG membranes, we observed low α -helical content ranging from 0 to 23%. The strong electrostatic interactions between cationic γ_1 and negatively charged PG may result in a higher entropy, less ordered bound state, which is consistent with ITC measurements [38]. The folding propensity of γ_1 on the 50:50 PC:PG bilayer could not be inferred from the behavior on the homogeneous membranes. γ_1 displayed low helical content (16%) on the mixed bilayer, leading us to conclude the correlation between the amount of charge present in the membrane and the folding propensity of the peptide is not linearly related. We also performed simulations starting from the folded state on PC, PG, and the mixed bilayer system, to probe the stability of the γ_1 helical conformation. We found that γ_1 does not unfold and remains embedded in the membranes throughout 1- μ s simulations. From these observations, we proposed that an energy barrier separates the folded state from the low helicity state, and the barrier height has a dependence on the amount of charge in the membrane [37]. Our equilibrium simulations were likely trapped in a metastable state, due to the rough energy landscape, which was insurmountable by conventional simulation approaches. To explore our proposition further, we appeal to enhanced sampling methods to overcome the limited sampling in equilibrium MD to produce barrier crossing events.

Different enhanced sampling methods offer varied advantages and disadvantages and one should make a well-informed selection of a method most applicable to addressing the scientific scenario being explored. A popular choice of method for studying protein–membrane systems is T-REMD [39]. T-REMD works on the principle that molecules can sample through a rugged energy surface by making repeated swaps among its replicas that are simulated simultaneously at different temperatures. Although replica-exchange is a widely used method, there are some disadvantages to of this approach. The number of replicas required for efficient sampling scale as $f^{1/2}$ (where f is the degrees of freedom), which for large biological systems with explicit solvent can lead to very high computational costs [40, 41]. Moreover, T-REMD is not an effective method to overcome entropic barriers, which are present in folding transitions [42]. Path sampling techniques such as milestoning [43], forward flux sampling [44], and transition path sampling [45] also offer non-biased simulation approaches that can be employed to study activated processes by exploiting transition path theory and calculating the key transitions in the trajectory space rather than focusing on the stable states. The basis of these methods is to sample the fast-occurring infrequent rare events involving a transition. Other enhanced sampling methods involve application of bias potentials to accelerate the sampling in a desired region of configurational space such as computational flooding [46], metadynamics (MetaD) [47], and umbrella sampling (US) [48]. These methods require the user to select an order parameter (collective variable, CV) along which a biasing potential can be applied to surmount free-energy barriers in the landscape. Both computational flooding and MetaD methods rely on biasing potentials being added “on the fly” to the energy landscape of the system with the objective to sample all energy minima, but avoid excessive and re-sampling of local minima.

Umbrella sampling is a mature and heavily utilized method in different biophysics studies including protein folding [49, 50], peptide–peptide interactions [51, 52], protein–DNA interactions [53], binding energies and interactions with lipid membranes [20, 37, 54, 55], and conformational sampling of small molecules [56, 57], among others. US relies on a stratification strategy; intermediate states along an order parameter are simulated with a restraint potential that keeps the system localized to a chosen point along the order parameter. A series of restrained simulations spanning the entire range of interest along the order parameter are simulated, and provided there are overlapping distributions between the umbrella windows, the probabilities can be unbiased and the potential of mean force (PMF) can be determined. Convergence in US is non-trivial to achieve or to evaluate and there are also choices regarding the restraint spacing and restraint force constant, though these are relatively easy to evaluate and there is significant literature to inform these choices. One important requirement of US is that an initial path needs to be defined. In the limit of infinite sampling, the initial pathway would be irrelevant. Although in practice it can have a significant effect, especially if there are slow degrees of freedom or energy minima states separated by large energy barriers in orthogonal degrees of freedom to the US coordinate. Factors that make US a more robust and advantageous technique are that additional sampling can be carried out where sampling is sparse or in windows which display slow transitions and it is an appropriate method to maximally utilize parallel computing.

Many non-enveloped viruses [58, 59] contain a membrane active component of their capsid that in some systems is an amphipathic peptide, which is disconnected from the capsid. We have been investigating the membrane lytic peptide of the non-enveloped Flock House virus (FHV), which displays characteristics similar to antimicrobial peptides [37, 60, 61]. In our previous study, we employed microsecond equilibrium simulations to examine the folding characteristics of γ_1 , on membranes of different compositions [37]. For the present study, we

aim to calculate the free-energy profile of γ_1 folding in the presence of a membrane. We chose to employ US and have chosen helicity of γ_1 as the order parameter for these calculations. To study the folding process, one needs to consider two main aspects, the starting conformational state of the peptide and also its orientational features, i.e., the depth and angle of the peptide with respect to the bilayer. Our approach to addressing these initiation concerns was to initiate our simulations from the last snapshot of our previous work containing the bound helical state of γ_1 on different membrane compositions [37]. The bound conformations are derived from 1- μ s equilibrium simulations, which have sampled the insertion depth of γ_1 at a depth consistent with experimental measurements, based upon Trp fluorescence [38]. Initial unfolding pathways were constructed by applying steered molecular dynamics (SMD) to generate γ_1 conformations of varying helicity. In addition to evaluating energetic and structural aspects of the folding pathways, we performed additional analyses to evaluate the convergence and uncertainty in the US data. We have utilized time-lagged independent components analysis (TICA) to identify the slowest decorrelating degrees of freedom. By performing projections of the US data into TICA subspaces, we can evaluate the connectivity of the US data and observe if regions in the configurational space are undersampled leading to poor estimates of the free energy surfaces (FES).

2 Methods

2.1 System setup

The γ_1 peptide is a 21-residue peptide with sequence ASMWERVKSIKSSLAAASNI. Initial configurations used to generate unfolding pathways were obtained from the last snapshot of previous 1- μ s equilibrium simulations [37]. The three systems are helical γ_1 bound to PC, PG, and 50:50 mix of PC:PG bilayers (Table 1). The bilayers consisted of 270 lipids for the PC bilayer, 288 lipids for the PG bilayer, and a total of 200 lipid molecules for the mixed bilayer. The peptide carries a net +2 charge and both termini are charged, which was done to be consistent with deletion construct experiments of FHV virus-like particles (VLP), which contained γ_1 [62, 63]. The simulations were performed with the GROMACS package (version 5.0.1) [64] using the CHARMM36 force field [65, 66]. The protein and membrane coordinates were extracted from the equilibrium simulations, but the systems were resolvated and reionized to reduce the system sizes. All systems were solvated with the TIP3P water model and were neutralized with 0.15 M NaCl. The systems were minimized using the steepest descent algorithm followed by an NVT equilibration for 5 ns, followed by 10 ns in the NPT ensemble. Pressure coupling was maintained at 1.0 bar using semi-isotropic coupling using the Parrinello–Rahman barostat. The temperature was maintained at 303.15 K using the velocity (v)-rescale coupling method. The temperature and pressure coupling constants were 1.0 ps and the compressibility value was $4.5 \times 10^{-5} \text{ bar}^{-1}$. The long-range electrostatics were calculated by the particle mesh Ewald (PME) [67] method and non-bonded interactions were cut off at 1.2 nm. The bonds lengths were constrained with the between LINCS constraint algorithm and particle periodic boundary conditions were applied in X, Y, and Z-directions. A time step of 2 fs. was used for the equations of motion integration. All system configurations and snapshots were visualized using VMD software [68].

Table 1 Simulation system details

System description	No. of PC and/or PG lipids	No. of US windows	Time of simulation	
			Per window (ns)	Total (μs)
Pure PC	270	125	80	10.0
Pure PG	288	148	60	8.9
Mixed	100:100	148	80	11.8

2.2 Umbrella sampling

The helical content of γ_1 was chosen as the collective variable (CV) to perform US calculations. We have performed all-atom US simulations using GROMACS 5.0 with the PLUMED 2.2 plugin [69, 70] to quantify the free energy associated with the transition of γ_1 from a helical (H) to a random coil (C) conformation on different membrane compositions. The CV ALPHARMSD [71] from the PLUMED 2.2 plugin was applied to generate intermediate configurations from H to C. This CV reports a sum, S_α , computed from the backbone root-mean-square deviation (RMSD) of six residue segments with respect to an ideal α -helix. The S_α value is calculated based on the following switching function

$$S_\alpha = \sum_i \frac{1 - \left(\frac{\text{RMSD}_i}{0.1}\right)^8}{1 - \left(\frac{\text{RMSD}_i}{0.1}\right)^{12}}$$

where RMSD is in units of nm and the sum runs over all possible consecutive six residue segments, yielding a theoretical maximum of $S_\alpha = 16$ for the 21 residue γ_1 peptide. Conversions between S_α and percent helicity are calculated by simply dividing the S_α value by the theoretical maximum S_α for the peptide (16) or segment (5).

In each of the systems, the peptide was unfolded using SMD with a restraint of 500 kJ/mol and a velocity of 0.001 S_α units/ps. Intermediate configurations spaced in 0.1 S_α increments were selected for US. An umbrella potential of 500 kJ/mol was applied in each umbrella window. The number of intermediate configurations (windows) for each system and the simulation time associated with each window are reported in Table 1. A total of 30.7 μ s of data was collected.

2.3 Trajectory analysis

The trajectory data excluding the first 20 ns from each umbrella window was used for both 1D and 2D weighted histogram analysis method (WHAM) [72] to calculate the PMFs. Both 1D and 2D WHAM analysis was performed using Alan Grossfield's code and the PMFs were converged to 10^{-3} kcal/mol [73]. The GROMACS tool g_mindist was used to calculate the number of contacts between γ_1 and lipid molecules within 5 Å radius. To compute the segmental S_α in a given conformation, three overlapping regions of γ_1 were defined: residues 1–10, residues 6–15, and residues 12–21, described as the N-term, middle, and C-term segments, respectively. The PLUMED plugin was also used for calculating the insertion depth of γ_1 . For this parameter, the three non-overlapping regions of γ_1 are defined as residues 1–7 as N-term, 8–14 as middle, and 15–21 as C-term. An evaluation of the flexibility of the S_α

restraint to allow the peptide to sample helicity in different regions of the peptide was performed using the `compute_dssp` utility of MDTraj [74]. For each system, the helicity was calculated in a low ($S_\alpha = 2$), medium ($S_\alpha = 7$), and high ($S_\alpha = 12$) window. The helicity was averaged in 10-ns segments and is presented in Fig. S1, which shows that the restraint does not rigidly fix the helical segments, but allows for some shifting of the conformation within an umbrella sampling window.

2.4 TICA and dTRAM analyses

The US trajectories were analyzed using pyEMMA 2.4 [74]. As was done for WHAM, the first 20 ns were excluded from each US window and the frames in the input trajectory were separated by 10 ps. The system was featurized by using the distances between all the C_α atoms of the peptide in order to reduce the degrees of freedom for analysis. The choice of C_α distance pairs has been shown to be a good metric to describe folding transitions and overcomes limitations inherent in position-based metrics [75]. The time-lagged covariance matrix, $C(\tau)$, with components

$$c_{ij}(\tau) = \langle r_i(t)r_j(t + \tau) \rangle$$

(where $r_i(t)$ represents different C_α atom pair distances at time t) was constructed. Time-lagged independent component analysis (TICA) [76, 77] was performed with several different lag times, all showing similar landscapes. A TICA lag time of 8 ns was chosen for further analysis because it gave the highest cumulative variance captured by the first few time-lagged independent components (tICs) and the slowest TICA timescales. TICA was performed on an individual system basis (system tICs) and also by combining all the systems into a master trajectory (global tICs). 2D projections onto tIC subspaces were performed and representative structures were chosen to visualize the transition pathways. To select the structures, the trajectory data were clustered into 60 clusters based on the global tIC 1, 2, and 3 values using mini batch k-means clustering (30 clusters were used when using system tICs). A representative structure was pulled from each of the clusters we chose to highlight. Folding PMFs were also estimated using a transition-based approach, discrete transition-based reweighting analysis method (dTRAM) [78]. Regular space clustering [79] was used to generate microstates using a minimal distance separation of $0.015 S_\alpha$ units, from which the transition matrix was computed. A 2-ns lag time was used in the dTRAM analysis, and all PMFs were converged to a tolerance of at least 10^{-7} kcal/mol.

3 Results and discussion

3.1 Folding energetics

We assessed the folding of γ_1 as function of helicity (S_α value) on membranes of different lipid compositions. The three systems and the simulation details are described in Table 1. The three membrane compositions are PC (zwitterionic), PG (anionic), and a mixed bilayer composed of equal number of PC and PG lipid molecules. The energetics of γ_1 folding on the membrane were calculated from the US data, using the WHAM algorithm. Figure 1 presents the 1D PMF for the three systems. We evaluated the total free energy change of folding of γ_1 to be similar

Fig. 1 PMF of folding FHV γ_1 peptide in different membrane environments

for PC (− 15.0 kcal/mol) and PG (− 14.4 kcal/mol), while the mixed system had a smaller free energy change (− 12.3 kcal/mol). The PC and mixed system both have a global minimum in a highly folded state (~ 92% helicity, state *H*), while for PG the highly folded state is metastable and the global minimum is at an intermediately helical configuration (~ 68% helicity, state *I*). The PC system also shows (meta)stability for *I* state and there are barriers for both PC and PG separating the *H* and *I* basins. The barrier in the folding direction for the PG system is 3.6 kcal/mol and 3.0 kcal/mol for the PC system. Interestingly, the mix system PMF is qualitatively different than the homogenous systems as it does not display stability at the *I* state. The barrier in the folding direction for the mixed system is small (1.8 kcal/mol), while the unfolding barrier is 4.2 kcal/mol.

Some features of the free energy profiles are not consistent with our expectations based from our previous work [37]. In equilibrium simulations, we observed rapid folding on the microsecond timescale for γ_1 in a PC systems, but did not observe significant folding in PG or mixed bilayer systems. Therefore, our prediction was that PG and mixed bilayers would have higher energy barriers in the folding direction which is in contrast with the PMFs. However, the PMFs are consistent with our observation that the folded state of γ_1 is stable in all membrane compositions and must overcome an energy barrier to significantly unfold. In our equilibrium simulations, we observed differing degrees of membrane penetration in the different bilayers, so to examine this feature in the current study, we calculated 2D PMFs using peptide insertion depth as the second coordinate. Figure 2 presents the 2D PMFs and it can be seen that PC (Fig. 2a) system has several local minima along the folding pathway and

Fig. 2 Two-dimensional free energy surface of γ_1 folding on PC (a), PG (b), and mixed (c) bilayers. The coordinates of the FES are the helicity and the peptide membrane insertion depth, measured relative to the upper leaflet mean phosphate positions

samples the most deeply inserted conformations, which correlates with our equilibrium simulation observations. The PG FES (Fig. 2b) shows the least insertion and the mixed system (Fig. 2c) displays a smooth landscape without significant local minima.

3.2 Folding mechanisms

Given the different thermodynamic stability of states along the folding pathways in different membranes environments, we wanted to understand if this reflected different folding pathways. To characterize the folding mechanisms, we performed a segmental analysis of the peptide structure and insertion depth. We calculated the S_α value of γ_1 over three different segments of the peptide (based on the definitions described in the Methods section). Figure 3 shows the segmental S_α value (averaged over the last 10 ns of the US simulations in each US window) with respect to the overall S_α value in each US window. We observe that for all three systems, the folding is initiated in the N-term, followed by the middle region, and the C-term folds last. While the PC and mixed system both show the N-term folds to near completion before the middle region begins to fold, the PG system shows there is concomitant (cooperative) folding between the N-term and middle regions. There are three basic and one acidic residue in γ_1 , and all of these charged residues are within the first 12 N-terminal residues of the peptide. The concentration of charges in the N-term and middle regions may be a driving force in the cooperative folding to align the charges in an energetically favorable manner when folding on the anionic PG membrane.

The mixed system also has an interesting feature not present in the homogeneous systems, which is that the C-term segment undergoes backtracking. The mixed system has a slightly metastable state around 40% helicity (Fig. 1, Fig. 2c), which is when the C-term segment begins to fold. Although, the C-term segment initiates folding, it returns to an unfolded state at around 60% total helicity. This backtracking feature may be indicative of a more frustrated folding landscape and barriers in orthogonal directions to the folding coordinate [80]. We examined two orthogonal coordinates along the folding pathway, the segmental insertion depth for all systems (Fig. S2) and the ratio of PG to PC contacts in the mixed system (Fig. 4). The segmental insertion depths qualitatively showed similar features for all systems, namely the N-term and middle regions remained inserted through the pathway and the C-term segment transitions from solvent exposed to embedded during the folding. The deeper penetration of the N-term and middle regions of γ_1 as opposed to the C-term region is in agreement with experimental [38] and our previous computational

Fig. 3 Segmental folding behavior of γ_1 in PC (a), PG (b), and mixed (c) bilayers. The total peptide folding was dissected into three overlapping ten-residue segments: residues 1–10 (N-term), residues 6–15 (middle), and residues 12–21 (C-term). The helicity values are obtained by averaging the S_α values of the segments over the final 10 ns in each US window. The absolute S_α values are divided by the theoretical limit which is 5 for a ten-residue segment. Curves are smoothed by performing a five-point moving average

results [37]. The contacts formed by γ_1 with the lipids in the mixed system did not show a significant dependence on the folding coordinate, and did not shift dramatically from the initial pathway (Fig. 4). In general, the peptide maintains a higher ratio of PG to PC contacts, forming ~ 1.5 PG contacts for every PC contact. This PG:PC contact ratio was consistent with the previous equilibrium simulations [37].

Overall, our ability to compare the US results with our previous equilibrium simulations is likely hampered by differences in the pathways being sampled. The equilibrium study was initiated from a surface-bound low helicity state in which the middle segment was folded while the N-term and C-term segments were unfolded. In the current study, the pathways were generated from SMD unfolding from a well-equilibrated high-helicity state. The US pathway does not pass through a conformation where the middle segment is folded in the absence of the N-term being folded and therefore the peptide is sampling different regions of the helicity-insertion phase space in comparing the equilibrium and US pathways.

3.3 Sampling analyses

Our abilities to interpret the US results are dependent on several factors, including the pathway which is being sampled, but also the convergence of sampling along the generated path. We have performed several analyses to evaluate the quality of the sampling in this study. We have reconstructed the 1D free-energy profile, using different segments of the US data with WHAM (Fig. 5a) and using a method based on transition probabilities, dTRAM [78] (Fig. 5b). In Fig. 5a, the block-averaged profiles for WHAM indicate the PG profile has low uncertainty, the PC profile has moderate uncertainty, and the mix profile has high uncertainty. Not surprisingly the higher uncertainties are observed for the low helicity region of the PMF, which has a high degeneracy of conformational states. TRAM methods have been shown to have a lower error than WHAM, especially for shorter sampling times [78, 81]. It is seen that the results of WHAM (Fig. 1, Fig. 5a) and dTRAM (Fig. 5b) are in qualitative agreement, and show good quantitative agreement in the higher helicity region ($> 70\%$). However, the WHAM profiles underestimate the total free energy change of folding, as the dTRAM profiles have higher free energy values in the unfolded states.

In addition to assessing the reliability of the reconstructed free-energy profile, we utilized TICA analysis to gain information on the mechanism of folding and to evaluate the landscape in the slowest decorrelating degrees of freedom (slowest tICs). TICA has been used on US data previously to analyze protein dynamics and as the basis for adaptive sampling [81, 82]. Our

Fig. 4 Ratio of PG:PC lipid contacts to the γ_1 peptide in the mixed bilayer system. Data are shown for the initial pathway (blue) and averaged over the final 10 ns in each US window (red)

Fig. 5 Alternative PMF reconstructions. **a** WHAM was performed using 20-ns blocks and averaged over the blocks. For the PC and mixed system, there were three blocks (20–40 ns, 40–60 ns, and 60–80 ns) and for the PG system there were two blocks (20–40 ns and 40–60 ns). The *shaded regions* are the standard error () of the block-averaged PMFs. **b** dTRAM analysis was performed using all data after excluding the first 20 ns for equilibration

systems were featurized using the distances between all C_{α} atoms in the peptide, which provides a reduced data set for performing TICA. Therefore, the TICA space is related to but not identical to the US coordinate, which is the S_{α} parameter. We performed a global TICA analysis by combining the systems into a single, master system, which allows us to obtain global tICs, and compare the systems on a consistent basis. The slowest global tIC (tIC1) describes the US coordinate very well as the two are strongly correlated in all three systems. By projecting each US window onto the tIC1-helicity space, it can be observed that the windows are well connected and evenly spaced in the tIC1 coordinate (Fig. S3). Of the three systems, only the mixed system (Fig. S3C), has a substantial gap which occurs around 50 % helicity and may be an indication of inadequate sampling and/or energy barriers in this region.

We also performed two-dimensional projections of the US data onto the global tIC1-tIC2 and global tIC1-tIC3 subspaces to observe the sampling in the slowest motions orthogonal to tIC1 (and by proxy orthogonal to the US coordinate). Figure 6 presents the 2D projections, where the surface is colored based on the biased probabilities, though relative probabilities in the 2nd and 3rd tICs, at fixed tIC1, should not be affected by the US restraints. We also performed the projections against the system specific tICs in both the tIC1-tIC2 and tIC1-tIC3 subspaces (Fig. S4). The same projections onto the global tICs are presented in Fig. S5 and S6, but colored by S_{α} and US time, respectively. Figure S5 shows that the folded state is occurring at high (positive) tIC1 values, while Fig. S6, indicates there is no detectable drift occurring in these slow degrees of freedom during US within a window. In general, the 2D projections show a narrowing of sampling at high tIC1 values, which is consistent with the folded state having a low structural degeneracy. Four peptide configurations are shown for each system (Fig. 6, Fig S4) to indicate the structural reorganization the peptide undergoes as it approaches the folded state. The PC and PG systems show relatively similar projections in Fig. 6, as the data is well connected in both subspaces and the sampling broadens at negative tIC1 values (low helicity states). In contrast, the mixed system displays a narrow path in the tIC1-tIC2 subspace and is disconnected in the tIC1-tIC3 subspace. While the PC and PG systems appear to follow a single pathway, the mixed system appears forked in the tIC1-tIC3 space and that a barrier located around tIC1 = 0 may separate these paths. A potential connection between the TICA projections and the PMF for the mixed system is that the lack of a minima in the PMF around 70% helicity could be due to the system be trapped in a higher energy pathway leading to higher energy states in the middle region (50–70% helicity) of the PMFs.

Fig. 6 Projection of US data onto global tICs in the tIC1-tIC2 and tIC1-tIC3 subspaces for PC (a), PG (b), and mixed (c) bilayer systems. Representative structures are shown and denoted by red 1, 2, 3, 4 markers on the surfaces. The coloring of the surfaces are based on a pseudo free energy ($F = -k_B T \ln P_b$) in kcal/mol, where P_b are the non-reweighted probabilities from US. In all the structures, the N-terminus is oriented to the right side of the molecule and the molecules are orientated in a manner consistent with the solvent phase being in the up direction and the membrane phase being in the down direction

4 Conclusions

We have performed US sampling calculations to analyze the folding energetics and mechanisms of the FHV γ_1 peptide in homogeneous and mixed bilayer systems. We have estimated the free-energy profile and observe that the different bilayers systems influence the equilibrium probabilities of the highly folded and intermediately folded states. The heterogeneous bilayer system displays several notable differences from the homogeneous systems, including the lack of a stable intermediately folded state, and folding backtracking in the C-term segment. Additional kinetic-based analyses show differences in sampling of the slow degrees of freedom between the homogeneous and mixed systems. In particular, a much narrower pathway is observed for the mixed system in the tIC2 direction and the appearance of multiple pathways and a barrier in the tIC3 direction. While the PMFs appear to be well converged, the

TICA analysis could direct a future adaptive sampling approach which could enhance the exploration of configurations away from the initial pathway.

Acknowledgements This work has been supported by the National Institutes of Health through grant R35GM119762 to E.R.M. Computational resources have been provided through the University of Connecticut Hornet HPC cluster and NSF XSEDE program (grant number TG-MCB140016). We thank Kevin Boyd for his critical reading and providing constructive suggestions on this manuscript.

Funding This study was funded by NIH (grant number R35GM119762).

Compliance with ethical standard

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Seelig, J.: Thermodynamics of lipid–peptide interactions. *Biochim. Biophys. Acta Biomembr.* **1666**(1–2), 40–50 (2004). <https://doi.org/10.1016/j.bbamem.2004.08.004>
2. Leontiadou, H., Mark, A.E., Marrink, S.J.: Antimicrobial peptides in action. *J. Am. Chem. Soc.* **128**(37), 12156–12161 (2006). <https://doi.org/10.1021/ja062927q>
3. Marčelja, S.: Lipid-mediated protein interaction in membranes. *Biochim. Biophys. Acta Biomembr.* **455**(1), 1–7 (1976). [https://doi.org/10.1016/0005-2736\(76\)90149-8](https://doi.org/10.1016/0005-2736(76)90149-8)
4. Owicki, J.C., McConnell, H.M.: Theory of protein–lipid and protein–protein interactions in bilayer membranes. *Proc. Natl. Acad. Sci. U. S. A.* **76**(10), 4750–4754 (1979). <https://doi.org/10.1073/pnas.76.10.4750>
5. Nymeyer, H., Woolf, T.B., Garcia, A.E.: Folding is not required for bilayer insertion: Replica exchange simulations of an α -helical peptide with an explicit lipid bilayer. *Proteins. Struct. Funct. Genet.* **59**(4), 783–790 (2005). <https://doi.org/10.1002/prot.20460>
6. Cymer, F., Von, H., White, S.H.: Mechanisms of integral membrane protein insertion and folding. *J. Mol. Biol.* **427**(5), 999–1022 (2015). <https://doi.org/10.1016/j.jmb.2014.09.014>
7. Sato, H., Feix, J.B.: Peptide–membrane interactions and mechanisms of membrane destruction by amphipathic α -helical antimicrobial peptides. *Biochim. Biophys. Acta Biomembr.* **1758**(9), 1245–1256 (2006). <https://doi.org/10.1016/j.bbamem.2006.02.021>
8. Li, C., Salditt, T.: Structure of magainin and alamethicin in model membranes studied by X-ray reflectivity. *Biophys. J.* **91**(9), 3285–3300 (2006). <https://doi.org/10.1529/biophysj.106.090118>
9. Schumann, M., Dathe, M., Wierprecht, T., Beyermann, M., Bienert, M.: The tendency of magainin to associate upon binding to phospholipid bilayers. *Biochemistry* **36**(14), 4345–4351 (1997). <https://doi.org/10.1021/bi962304x>
10. Leavitt, S., Freire, E.: Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.* **11**(5), 560–566 (2001). [https://doi.org/10.1016/S0959-440X\(00\)00248-7](https://doi.org/10.1016/S0959-440X(00)00248-7)
11. Perrin, B.S., Tian, Y., Fu, R., et al.: High-resolution structures and orientations of antimicrobial peptides piscidin 1 and piscidin 3 in fluid bilayers reveal tilting, kinking, and bilayer immersion. *J. Am. Chem. Soc.* **136**(9), 3491–3504 (2014). <https://doi.org/10.1021/ja411119m>
12. Afonin, S., Grage, S.L., Ieronimo, M., Wadhwani, P., Ulrich, A.S.: Temperature-dependent transmembrane insertion of the amphiphilic peptide PGLa in lipid bilayers observed by solid state ¹⁹F NMR spectroscopy. *J. Am. Chem. Soc.* **130**(49), 16512–16514 (2008). <https://doi.org/10.1021/ja803156d>
13. Ladokhin, A.S., White, S.H.: Folding of amphipathic α -helices on membranes: Energetics of helix formation by melittin. *J. Mol. Biol.* **285**(4), 1363–1369 (1999). <https://doi.org/10.1006/jmbi.1998.2346>
14. Wierprecht, T., Seelig, J.: Isothermal titration calorimetry for studying interactions between peptides and lipid membranes. *Curr. Top. Membr.* **52**, 31–56 (2002)
15. Bechinger, B., Kim, Y., Chirlian, L.E., et al.: Orientations of amphipathic helical peptides in membrane bilayers determined by solid-state NMR spectroscopy. *J. Biomol. Nmr.* **1**(2), 167–173 (1991). <https://doi.org/10.1007/BF01877228>

16. Bechinger, B., Gierasch, L.M., Montal, M., Zasloff, M., Opella, S.J.: Orientations of helical peptides in membrane bilayers by solid state NMR spectroscopy. *Solid State Nucl. Magn. Reson.* **7**(3), 185–191 (1996). [https://doi.org/10.1016/0926-2040\(95\)01224-9](https://doi.org/10.1016/0926-2040(95)01224-9)
17. Beschiaschvili, G., Seelig, J.: Melittin Binding to Mixed Phosphatidylglycerol/Phosphatidylcholine Membranes. *Biochemistry* **29**(1), 52–58 (1990). <https://doi.org/10.1021/bi00453a007>
18. Chen, C.H., Wiedman, G., Khan, A., Ulmschneider, M.B.: Absorption and folding of melittin onto lipid bilayer membranes via unbiased atomic detail microsecond molecular dynamics simulation. *Biochim. Biophys. Acta Biomembr.* **1838**(9), 2243–2249 (2014). <https://doi.org/10.1016/j.bbamem.2014.04.012>
19. Von, D., Knecht, V.: Antimicrobial selectivity based on zwitterionic lipids and underlying balance of interactions. *Biochim. Biophys. Acta Biomembr.* **1818**(9), 2192–2201 (2012). <https://doi.org/10.1016/j.bbamem.2012.05.012>
20. Irudayam, S.J., Berkowitz, M.L.: Binding and reorientation of melittin in a POPC bilayer: Computer simulations. *Biochim. Biophys. Acta Biomembr.* **1818**(12), 2975–2981 (2012). <https://doi.org/10.1016/j.bbamem.2012.07.026>
21. Koehler, L., Ulmschneider, M.B., Gray, J.J.: Computational modeling of membrane proteins. *Proteins Struct. Funct. Bioinforma.* **83**(1), 1–24 (2015). <https://doi.org/10.1002/prot.24703>
22. Ash, W.L., Zlomislic, M.R., Oloo, E.O., Tieleman, D.P.: Computer simulations of membrane proteins. *Biochim. Biophys. Acta Biomembr.* **1666**(1–2), 158–189 (2004). <https://doi.org/10.1016/j.bbamem.2004.04.012>
23. Andersson, M., Ulmschneider, J.P., Ulmschneider, M.B., White, S.H.: Conformational states of melittin at a bilayer interface. *Biophys. J* **104**(6), L12–L14 (2013). <https://doi.org/10.1016/j.bpj.2013.02.006>
24. Bereau, T., Bennett, W.F.D., Pfäendtner, J., Desemo, M., Karttunen, M.: Folding and insertion thermodynamics of the transmembrane WALP peptide. *J. Chem. Phys.* **143**, 24 (2015). <https://doi.org/10.1063/1.4935487>
25. Im, W., Brooks III, C.L.: Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. *Proc. Natl. Acad. Sci.* **102**(19), 6771–6776 (2005). <https://doi.org/10.1073/pnas.0408135102>
26. Tieleman, D.P., Berendsen, H.J.C., Sansom, M.S.P.: Surface binding of alamethicin stabilizes its helical structure: Molecular dynamics simulations. *Biophys. J* **76**(6), 3186–3191 (1999). [https://doi.org/10.1016/S0006-3495\(99\)77470-9](https://doi.org/10.1016/S0006-3495(99)77470-9)
27. Lin, D., Grossfield, A.: Thermodynamics of antimicrobial lipopeptide binding to membranes: Origins of affinity and selectivity. *Biophys. J* **107**(8), 1862–1872 (2014). <https://doi.org/10.1016/j.bpj.2014.08.026>
28. Lindahl, E., Sansom, M.S.: Membrane proteins: molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **18**(4), 425–431 (2008). <https://doi.org/10.1016/j.sbi.2008.02.003>
29. Ward, M.D., Nangia, S., May, E.R.: Evaluation of the hybrid resolution PACE model for the study of folding, insertion, and pore formation of membrane associated peptides. *J. Comput. Chem.* **38**(16), 1462–1471 (2017). <https://doi.org/10.1002/jcc.24694>
30. Shai, Y.: Mode of action of membrane active antimicrobial peptides. *Biopolym. - Pept. Sci. Sect.* **66**(4), 236–248 (2002). <https://doi.org/10.1002/bip.10260>
31. Shai, Y.: Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α -helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochim. Biophys. Acta Biomembr.* **1462**(1–2), 55–70 (1999). [https://doi.org/10.1016/S0005-2736\(99\)00200-X](https://doi.org/10.1016/S0005-2736(99)00200-X)
32. Yuan, T., Zhang, X., Hu, Z., Wang, F., Lei, M.: Molecular dynamics studies of the antimicrobial peptides piscidin 1 and its mutants with a DOPC lipid bilayer. *Biopolymers* **97**(12), 998–1009 (2012). <https://doi.org/10.1002/bip.22116>
33. Rahmanpour, A., Ghahremanpour, M.M., Mehmejad, F., Moghaddam, M.E.: Interaction of Piscidin-1 with zwitterionic versus anionic membranes: A comparative molecular dynamics study. *J. Biomol. Struct. Dyn.* **31**(12), 1393–1403 (2013). <https://doi.org/10.1080/07391102.2012.737295>
34. Tieleman, D.P., Sansom, M.S.P., Berendsen, H.J.C.: Alamethicin helices in a bilayer and in solution: Molecular dynamics simulations. *Biophys. J* **76**(1 I), 40–49 (1999)
35. Perrin, B.S., Pastor, R.W.: Simulations of membrane-disrupting peptides I: alamethicin pore stability and spontaneous insertion. *Biophys. J* **111**(6), 1248–1257 (2016). <https://doi.org/10.1016/j.bpj.2016.08.014>
36. Perrin, B.S., Fu, R., Cotten, M.L., Pastor, R.W.: Simulations of membrane-disrupting peptides II: AMP piscidin 1 favors surface defects over pores. *Biophys. J* **111**(6), 1258–1266 (2016). <https://doi.org/10.1016/j.bpj.2016.08.015>
37. Nangia, S., May, E.R.: Influence of membrane composition on the binding and folding of a membrane lytic peptide from the non-enveloped flock house virus. *Biochim. Biophys. Acta Biomembr.* **1859**(7), 1190–1199 (2017). <https://doi.org/10.1016/j.bbamem.2017.04.002>

38. Bong, D.T., Steinem, C., Janshoff, A., Johnson, J.E., Ghadiri, M.R.: A highly membrane-active peptide in flock house virus: Implications for the mechanism of nodavirus infection. *Chem. Biol.* **6**(7), 473–481 (1999). [https://doi.org/10.1016/S1074-5521\(99\)80065-9](https://doi.org/10.1016/S1074-5521(99)80065-9)
39. Sugita, Y., Okamoto, Y.: Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**(1–2), 141–151 (1999)
40. Gallicchio, E., Levy, R.M., Parashar, M.: Asynchronous replica exchange for molecular simulations. *J. Comput. Chem.* **29**(5), 788–794 (2008). <https://doi.org/10.1002/jcc.20839>
41. Periole, X., Mark, A.E.: Convergence and sampling efficiency in replica exchange simulations of peptide folding in explicit solvent. *J. Chem. Phys.* **126**, 1 (2007). <https://doi.org/10.1063/1.2404954>
42. Lee, K.H., Chen, J.: Multiscale enhanced sampling of intrinsically disordered protein conformations. *J. Comput. Chem.* **37**(6), 550–557 (2016). <https://doi.org/10.1002/jcc.23957>
43. Faradjian, A.K., Elber, R.: Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.* **120**(23), 10880–10889 (2004). <https://doi.org/10.1063/1.1738640>
44. Allen, R.J., Warren, P.B., Ten, W.: Sampling rare switching events in biochemical networks. *Phys. Rev. Lett.* **94**, 1 (2005). <https://doi.org/10.1103/PhysRevLett.94.018104>
45. Dellago, C., Bolhuis, P.G., Csajka, F.S., Chandler, D.: Transition path sampling and the calculation of rate constants. *J. Chem. Phys.* **108**(5), 1964 (1998). <https://doi.org/10.1063/1.475562>
46. Grubmiller, H.: Predicting slow structural transitions in macromolecular systems: conformational flooding. *Phys. Rev. E* **52**(3), 2893–2906 (1995). <https://doi.org/10.1103/PhysRevE.52.2893>
47. Laio A, Gervasio FL. Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* 2008;71(12). <https://doi.org/10.1088/0034-4885/71/12/126601>.
48. Torrie, G.M., Valleau, J.P.: Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J. Comput. Phys.* **23**(2), 187–199 (1977). [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8)
49. Young, W.S., Brooks III, C.L.: A microscopic view of helix propagation: N and C-terminal helix growth in alanine helices. *J. Mol. Biol.* **259**(3), 560–572 (1996). <https://doi.org/10.1006/jmbi.1996.0339>
50. Bursulaya, B.D., Brooks III, C.L.: Folding free energy surface of a three-stranded β -sheet protein. *J. Am. Chem. Soc.* **121**(43), 9947–9951 (1999). <https://doi.org/10.1021/ja991764i>
51. Mahdavi, S., Kuyucak, S.: Why the *Drosophila* shaker K⁺ channel is not a good model for ligand binding to voltage-gated K⁺ channels. *Biochemistry* **52**(9), 1631–1640 (2013). <https://doi.org/10.1021/bi301257p>
52. Vijayaraj, R., Van, D., Bultinck, P., Subramanian, V.: Molecular dynamics and umbrella sampling study of stabilizing factors in cyclic peptide-based nanotubes. *J. Phys. Chem. B* **116**(33), 9922–9933 (2012). <https://doi.org/10.1021/jp303418a>
53. Yesudhas, D., Anwar, M.A., Panneerselvam, S., Kim, H.-K., Choi, S.: Evaluation of Sox2 binding affinities for distinct DNA patterns using steered molecular dynamics simulation. *FEBS Open Bio* **7**(11), 1750–1767 (2017). <https://doi.org/10.1002/2221-5463.12316>
54. Vermaas, J.V., Tajkhorshid, E.: Differential membrane binding mechanics of synaptotagmin isoforms observed in atomic detail. *Biochemistry* **56**(1), 281–293 (2017). <https://doi.org/10.1021/acs.biochem.6b00468>
55. Mascarenhas, N.M., Kästner, J.: How maltose influences structural changes to bind to maltose-binding protein: results from umbrella sampling simulation. *Proteins Struct. Funct. Bioinforma.* **81**(2), 185–198 (2013). <https://doi.org/10.1002/prot.24174>
56. Patrascu, M.B., Malek-Adamian, E., Damha, M.J., Moitessier, N.: Accurately modeling the conformational preferences of nucleosides. *J. Am. Chem. Soc.* **139**(39), 13620–13623 (2017). <https://doi.org/10.1021/jacs.7b07436>
57. Schaefer, M., Bartels, C., Karplus, M.: Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit solvation model. *J. Mol. Biol.* **284**(3), 835–848 (1998). <https://doi.org/10.1006/jmbi.1998.2172>
58. Banerjee, M., Johnson, J.E.: Activation, exposure and penetration of virally encoded, membrane-active polypeptides during non-enveloped virus entry. *Curr. Protein Pept. Sci.* **9**(1), 16–27 (2008). <https://doi.org/10.2174/138920308783565732>
59. Kumar, C., Dey, D., Ghosh, S., Banerjee, M.: Breach: Host membrane penetration and entry by nonenveloped viruses. *Cell Press Rev. (Trends In Microbiology)*. <https://doi.org/10.1016/j.tim.2017.09.010>
60. Lewis, J.R., Cafiso, D.S.: Correlation between the free energy of a channel-forming voltage-gated peptide and the spontaneous curvature of bilayer lipids. *Biochemistry* **38**(18), 5932–5938 (1999). <https://doi.org/10.1021/bi9828167>
61. Bulet, P., Hetru, C., Dimarcq, J.-L., Hoffmann, D.: Antimicrobial peptides in insects; structure and function. *Dev. Comp. Immunol.* **23**(4–5), 329–344 (1999). [https://doi.org/10.1016/S0145-305X\(99\)00015-4](https://doi.org/10.1016/S0145-305X(99)00015-4)

62. Banerjee, M., Khayat, R., Walukiewicz, H.E., Odegard, A.L., Schneemann, A., Johnson, J.E.: Dissecting the functional domains of a nonenveloped virus membrane penetration peptide. *J. Virol.* **83**(13), 6929–6933 (2009). <https://doi.org/10.1128/JVI.02299-08>
63. Bajaj, S., Dey, D., Bhukar, R., Kumar, M., Banerjee, M.: Non-enveloped virus entry: structural determinants and mechanism of functioning of a viral lytic peptide. *J. Mol. Biol.* **428**(17), 3540–3556 (2016). <https://doi.org/10.1016/j.jmb.2016.06.006>
64. Abraham, M.J., Murtola, T., Schulz, R., et al.: Gromacs: High-performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015). <https://doi.org/10.1016/j.softx.2015.06.001>
65. Best, R.B., Zhu, X., Shim, J., et al.: Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ 1 and χ 2 Dihedral Angles. *J. Chem. Theory Comput.* **8**(9), 3257–3273 (2012). <https://doi.org/10.1021/ct300400x>
66. Klauda, J.B., Venable, R.M., Freites, J.A., et al.: Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**(23), 7830–7843 (2010). <https://doi.org/10.1021/jp101759q>
67. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G.: A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**(19), 8577–8593 (1995)
68. Humphrey, W., Dalke, A., Schulten, K.: VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**(1), 33–38 (1996). [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
69. Tribello, G.A., Bonomi, M., Branduardi, D., Camilloni, C., Bussi, G.: PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**(2), 604–613 (2014). <https://doi.org/10.1016/j.cpc.2013.09.018>
70. Bonomi, M., Branduardi, D., Bussi, G., et al.: PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **180**(10), 1961–1972 (2009). <https://doi.org/10.1016/j.cpc.2009.05.011>
71. Pietrucci, F., Laio, A.: A collective variable for the efficient exploration of protein beta-sheet structures: application to SH3 and GB1. *J. Chem. Theory Comput.* **5**(9), 2197–2201 (2009). <https://doi.org/10.1021/ct900202f>
72. Roux, B.: The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **91**(1–3), 275–282 (1995). [https://doi.org/10.1016/0010-4655\(95\)00053-1](https://doi.org/10.1016/0010-4655(95)00053-1)
73. Grossfield A. Grossfield, Alan, “WHAM: the weighted histogram analysis method”, version.
74. McGibbon, R.T., Beauchamp, K.A., Harrigan, M.P., et al.: MDTraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**(8), 1528–1532 (2015). <https://doi.org/10.1016/j.bpj.2015.08.015>
75. Husic BE, McGibbon RT, Sultan MM, Pande VS. Optimized parameter selection reveals trends in Markov state models for protein folding. *J. Chem. Phys.* 2016;145(19). <https://doi.org/10.1063/1.4967809>.
76. 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(1). <https://doi.org/10.1063/1.4811489>.
77. Schwantes, C.R., Pande, V.S.: Improvements in Markov state model construction reveal many non-native interactions in the folding of NTL9. *J. Chem. Theory Comput.* **9**(4), 2000–2009 (2013). <https://doi.org/10.1021/ct300878a>
78. Wu H, Mey ASJS, Rosta E, Noé F. Statistically optimal analysis of state-discretized trajectory data from multiple thermodynamic states. *J. Chem. Phys.* 2014;141(21). <https://doi.org/10.1063/1.4902240>.
79. Prinz JH, Wu H, Sarich M, et al. Markov models of molecular kinetics: generation and validation. *J. Chem. Phys.* 2011;134(17). <https://doi.org/10.1063/1.3565032>.
80. Hills, R.D., Brooks III, C.L.: Subdomain competition, cooperativity, and topological frustration in the folding of CheY. *J. Mol. Biol.* (2008). <https://doi.org/10.1016/j.jmb.2008.07.007>
81. Wu H, Paul F, Wehmeyer C, Noé F. Multiensemble Markov models of molecular thermodynamics and kinetics. 2016. <https://doi.org/10.1073/pnas.1525092113>.
82. Jo, S., Suh, D., He, Z., Chipot, C., Roux, B.: Leveraging the information from Markov state models to improve the convergence of umbrella sampling simulations. *J. Phys. Chem. B* **120**(33), 8733–8742 (2016). <https://doi.org/10.1021/acs.jpcc.6b05125>