

Thermodynamic Details of Pinholin S²¹68 Activation Revealed Using Alchemical Free Energy Simulations

Theodore S. Kalbfleisch,* Tanbir Ahammad, Gary A. Lorigan, and Vance W. Jaeger*



Cite This: *J. Phys. Chem. B* 2024, 128, 8762–8770



Read Online

ACCESS |



Metrics & More

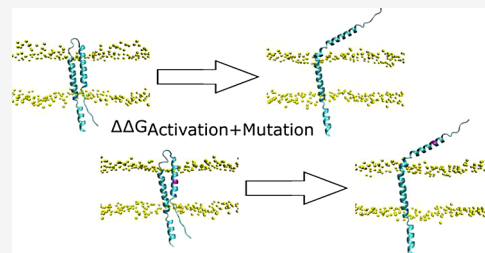


Article Recommendations



Supporting Information

ABSTRACT: Pinholin S²¹68 is a viral integral membrane protein whose function is to form nanoscopic “pinholes” in bacterial cell membranes to induce cell lysis as part of the viral replication cycle. Pinholin can transition from an inactive to an active conformation by exposing a transmembrane domain (TMD1) to the extracellular fluid. Upon activation, several copies of the protein assemble via interactions among a second transmembrane domain (TMD2) to form a single pore, thus hastening cell lysis and viral escape. The following experiments provide conformational descriptors of pinholin in active and inactive states and elucidate the molecular driving forces that control pinholin activity. In the present study, molecular dynamics (MD) simulations have been used to refine experimentally derived conformational descriptors into an atomistically detailed model of irs^{S21}68, an antiholin mutant. To provide additional details about the thermodynamics of pinholin activation and to overcome large intrinsic kinetic barriers to activation, alchemical free energy simulations have been conducted. Alchemical mutations reveal the change in folding free energy upon mutation. The results suggest that alchemical mutations are an effective tool to rationalize experimental observations and predict the effects of site mutations on conformational states for proteins integrated into lipid bilayers. S16F, A17Q, A17Q+G21Q, and A17Q+G21Q+G14Q mutants reveal how changes in hydrophilicity and disruption of the glycine zipper motif influence pinholin’s thermodynamic equilibrium, favoring the active conformation. These findings align with experimental observations from DEER spectroscopy, demonstrating that mutations increasing the hydrophilicity of TMD1 promote activation by making TMD1 more likely to exit the membrane and enter the extracellular fluid.



INTRODUCTION

Pinholin S²¹ describes a pair of proteins, namely an active (S²¹68) that forms pores within membranes and an inactive antipinholin (S²¹71) that slows the activation kinetics.¹ Viral phages balance the ratio of these two proteins to regulate the timing of bacterial cell lysis. These proteins both have two transmembrane domains, TMD1 and TMD2. TMD2 drives the assembly of several S²¹68 pinholin copies to form pin holes in the bacterial cell membrane. TMD1 inhibits the assembly of TMD2 to prevent membrane pore formation. TMD1 locks the protein in an inactive state until an allele-specific trigger unlocks the transition.² Upon an as-of-yet undiscovered stimulus, TMD1 transitions from being integrated into the membrane to being periplasm-exposed, activating individual pinholin proteins. Subsequently TMD2 is free to oligomerize with other proteins’ activated TMD2s to induce pore formation. Pores disrupt the cell’s proton gradient, initiating the lysis cascade.^{2a,3} This study will use pinholin irs^{S21}68 (Figure 1) as a model for S²¹71 to analyze antiholin mechanisms. The irs^{S21}68 antiholin differs in structure from wild type S²¹68 pinholin by the addition of a five amino acid tail to the N-terminal end of TMD1. These additional amino acids greatly inhibit TMD1 from exiting the membrane and entering the extracellular fluid.

Pinholin’s function is intricately tied to its structure, and its structure is dictated by its amino acid sequence. Hydrophobic and hydrophilic amino acids and their interactions with the local environment drive TMDs to embed in the cell membrane in specific orientations. There is a delicate balance in interactions of pinholin with its surroundings. In experiments, charged lipid heads and hydrophobic tails prevent hydrophobic moieties from entering the aqueous extracellular fluid or cytoplasm. Interactions between the two TMDs determine functionality of the protein.^{1,4} The favorable interactions between TMD1 and TMD2 that lock pinholin in its inactive state are thermodynamically balanced so the protein activates at the appropriate time. Thus, the thermodynamics and kinetics of activation are tightly controlled by evolution. One crucial structural motif that drives TMD1-TMD2 interactions and the locking of the protein into its inactive state is a glycine zipper along the interface between the domains.^{1–5} Glycine

Received: May 19, 2024

Revised: July 29, 2024

Accepted: August 19, 2024

Published: August 28, 2024



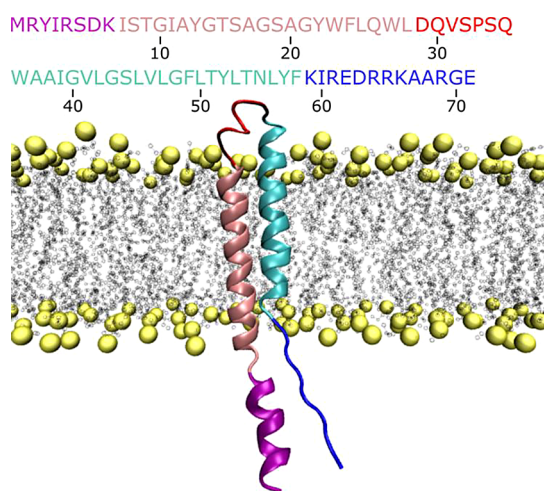


Figure 1. An illustration of the pinholin irsS²¹⁶⁸ protein and its amino acid sequence embedded in a DMPC lipid membrane. The color coding of the text in the sequence matches its corresponding color and moiety in the illustration.

zippers provide several localized stabilizing hydrogen bonds as well as favorable geometry to interlock the zipper motifs.⁶ Even small modifications to pinholin's amino acid sequence have the potential to greatly impact the protein's structure, dynamics, and function. Realizing this fact, many experimental groups have studied pinholin and antiholin as model systems for probing the structural transitions induced by mutation.^{1,4}

Young and co-workers performed a set of experiments to measure the lethality, or propensity to activate and oligomerize, of point mutations throughout the S²¹⁶⁸ pinholin. They assessed the effects of these mutations by applying them to isolated TMD2 helices, and irsS²¹⁶⁸ antiholin.^{2,3,7} These point mutations focused on modulating the strength of protein–solvent and protein–protein interactions. By recording the relative triggering time and aggregation levels of mutant holins with respect to the wild type, Young and co-workers were able to surmise the importance of specific amino acids as well as the mechanistic roles they play in pinholin activation and assembly. In the current study, we will use molecular simulations to probe some of the mutations performed by Young and co-workers and Lorigan and co-workers on the irsS²¹⁶⁸ antiholin.^{1,2b} Mutations to TMD1 of antiholin that result in a lethal phage indicate a more drastic impact on the function of pinholin, because the sole purpose of TMD1 is to limit the function of the remainder of the protein.

Lorigan and co-workers selected four interesting mutations from Young and co-workers' studies to explore using double electron–electron resonance (DEER) and continuous wave electron paramagnetic resonance (CW-EPR) spectroscopy.^{1,4,8} DEER and CW-EPR spectroscopy are useful to assess mutation-driven conformational changes of proteins embedded in lipid membranes. Lorigan and workers discerned the distance between spin labeled amino acids and the extent of amino acid exposure to different solvation environments (i.e., hydrophobic lipid tails, hydrophilic lipid heads, and water). Thus, these experiments confirm whether the protein is in its inactive or active conformation, and other more subtle structural differences between mutant and WT states can also be gleaned. While these experiments provide useful structural data, the results are often inferred from spectroscopic signals. The thermodynamic impact of these mutations

on structural stability and the locations of each atom in the system are difficult or impossible to explore experimentally as there are no characterization techniques that can probe aqueous membrane energetics and structure simultaneously. Molecular dynamics (MD) simulations can complement experiments by providing molecular-level details of pinholin's structure and dynamics while also elucidating energetic contributions to stability.

MD simulations have become an essential tool for studying membrane proteins, providing insights that are difficult or impossible to obtain through experimental techniques alone. These simulations offer detailed information about the structure, dynamics, and interactions of membrane proteins within their lipid bilayer environments.⁹ Of particular relevance to this study, several researchers have used MD to probe large conformational changes upon the introduction of stimuli. For instance, MD simulations have been used to investigate the conformational dynamics of the G-protein coupled receptors (GPCRs), providing crucial insights into their activation mechanisms and interactions with ligands.¹⁰ Simulations of the mechanosensitive channel of large conductance (MscL), which elucidated the gating mechanisms in response to membrane tension.¹¹ MD studies on the water channel protein aquaporin have provided a detailed understanding of its selective water permeability and gating mechanisms.¹² Through MD simulations, membrane protein behavior can be observed in various conditions, the effects of different lipid compositions can be explored, and the mechanisms of protein–lipid interactions can be understood. This comprehensive approach allows for the examination of conformational changes, stability, and functional dynamics of membrane proteins, such as pinholin, under different conditions. MD simulations can predict how mutations affect protein structure and function, providing a deeper understanding of the molecular basis of diseases and guiding the design of therapeutic interventions. The integration of MD simulations with experimental data, like DEER and CW-EPR spectroscopy, creates a synergy that enhances our ability to rationalize the behavior of membrane proteins and their roles in cellular processes.^{4b,8,9}

While MD simulations are a useful tool to investigate the structural and energetic effects of protein mutations, the time scales that simulations can probe are limited. For explicitly solvated systems containing a lipid bilayer, pinholin, and a solvent bath, it is reasonable to simulate a system for up to a few microseconds at atomic resolution using widely available computers. Coarse-grained simulations, which remove degrees of freedom from the system by combining the contributions of several atoms into one bead, can achieve time scales of tens to hundreds of microseconds. Yet, these coarse-grained models lack the detail to describe crucial interactions like those in a glycine zipper. The transition of pinholin from the inactive to the active state is much slower than can be probed by MD simulations. Young and co-workers, however, observed structural transitions on the minute time scale.^{2b,d} Kinetic limitations arise for two reasons. First, it is unknown what stimulus causes activation, and without providing that stimulus, activation is a rare event. There is no guarantee that a simulation will capture the activation cycle. Second, protein dynamics are severely hampered by the crowding of the lipid membrane environment. Activation takes place on a time scale much longer than hundreds of microseconds. Even if the

activation path were known, standard MD simulations would not be long enough to transition between the two states.

To overcome these weaknesses of MD and uncover the thermodynamics of the structural transition, alchemical mutation free energy simulations can be used. Alchemical calculations use MD to focus on the important end states of a thermodynamic cycle rather than sampling the transitions. MD simulations require a force field to describe the interatomic potential energies among the atoms in the simulation box.¹³ Common force fields (e.g., AMBER and CHARMM families) derive interaction potentials from quantum mechanical (QM) calculations and from experimental data. In an alchemical mutation simulations, interactions from one amino acid can be decoupled from the system while another amino acid is simultaneously coupled to the system as in eq 1. V is the total system potential, V_A is the system potential of state A, V_B is the system potential in state B and λ is the coupling parameter that describes the system. This means that an amino acid can be mutated out or mutated in at will. Furthermore, energetics of these mutations can be analyzed.

$$V = \lambda V_A + (1 - \lambda)V_B \quad (1)$$

Seeliger and de Groot have demonstrated that the change in thermal stability of a protein upon mutation can be accurately estimated using alchemical mutations.¹⁴ In these simulations, a four-state thermodynamic cycle is constructed containing the folded wild type (WT), folded mutant, unfolded WT, and unfolded mutant protein. In some cases, the authors used a tripeptide rather than the unfolded state and still made accurate predictions. This cycle describes the change in Gibbs free energy of folding upon mutation ($\Delta\Delta G_{\text{fold-mut}}$). The transition from unfolded to folded states is a slow process. To characterize the folding legs of the thermodynamic cycle, prohibitively long standard MD simulations would be needed. Therefore, the mutation legs of the thermodynamic cycle are explored instead using alchemical mutation simulations. In this study, we construct a similar cycle as described in the [Methods](#) section. Because free energy is a state variable, the change in unfolding free energy upon mutation can be calculated either as (a) the difference between the free energy of folded versus unfolded states for the mutant and WT proteins (ΔG_{fold}) or (b) the difference in the free energy of the mutant and WT protein in the folded state versus the mutant and WT protein in the unfolded state (ΔG_{mut}). Seeliger's method provides a value of ($\Delta\Delta G_{\text{fold-mut}}$) which gives the change in folding free energy upon mutation.¹⁴ The $\Delta\Delta G$ we calculate is similar to Seeliger. It is the change in activation energy upon mutation ($\Delta\Delta G_{\text{act-mut}}$)

METHODS

Atomistic Model Construction and Validation. An atomistic representation of the irsS²¹⁶⁸ pinholin was constructed based upon structural descriptors provided in Lorigan and co-worker's previous work.^{1,4} The protein structure was broken into five main parts, namely the N-terminal tail, TMD1, the bridging beta strand, TMD2 and the C-terminal tail. The Pro-Builder online tool⁴⁵ was used to generate a structure for each of these segments. The segments were manually reoriented in VMD, and their coordinate files were concatenated to best recreate the observed structural constraints from Lorigan's previous work. Relative TMD penetration depth and angles were used, and the glycine zipper motif was maintained for the inactive state. This process was

performed for both the inactive and active conformations because both inactive and active conformations are needed to construct the alchemical thermodynamic cycle. CHARMM GUI Membrane Bilayer Builder was used to generate the simulation box.¹⁶ The system is a $\sim 8 \times 8 \times 14$ nm box consisting of an irsS²¹⁶⁸ pinholin protein embedded in a DMPC membrane with a 0.15 M NaCl aqueous solution to match experimental conditions in Lorigan's studies. Simulations used the CHARMM36m force field¹⁷ with CHARMM TIP3P water¹⁸ (i.e., with LJ interactions on the hydrogen atoms) and ions. Input files were generated for GROMACS.¹⁹ The penetration depth of each amino acid in the TMDs was checked to ensure agreement with Lorigan and co-workers' data.^{1,4b,8} Each of the four mutant proteins were created by adding mutations (S16F, A17Q, A17Q+G21Q, and A17Q+G21Q+G14Q) to the WT structure using CHARMM GUI. This was done for both active and inactive states for the wild type and each mutant, resulting in 10 total structures. Box contents and dimensions are provided in [Table S1](#).

Generating Hybrid Topology and Equilibration. Using the pmx python library, the active and inactive starting models (before minimization and equilibration) have mutant amino acids inserted in place of the wild type amino acids.²⁰ The pmx library contains topologies for CHARMM36m. The state of these mutant amino acids is numerically represented by the λ parameter, which is anywhere between zero and one (zero being WT and one being mutated). Energy minimization used the steepest descent method for up to 5000 steps. A series of short equilibration simulations were conducted to bring the system to a stable state. Leonard-Jones interactions were cutoff at 1.2 nm with a force-switch applied from 1.0 to 1.2 nm. Particle mesh Ewald (PME) electrostatics were used for one-range interactions. During these equilibration simulations, the time step was increased from 1 to 2 fs, position restraints on the lipid headgroups were relaxed, and the Berendsen barostat was used to maintain pressure at 1 bar.²¹

Model Validation Simulation. GROMACS 2020 was used to run all simulations.¹⁹ The velocity-rescaling thermostat was used to maintain a temperature of 298.15 K.²² In production simulations, the Parrinello–Rahman barostat was used to maintain a pressure of 1 bar. An integration time step of 2 fs was used for numerical integration. After 3 ns of equilibrium simulation, the trajectory was examined using MDAnalysis²³ to mark any changes in pinholin penetration depth, TMD angles, and positions of amino acids observed in Lorigan's experiments.²⁴ This validation indicated that the starting structures were near their experimentally observed values. Validation comparisons are presented in the [Results](#) section

Generating Starting Structures. Once equilibrated, the models were simulated for 500 ns in the NPT ensemble to generate sets of unique initial frames for the mutation sample population. Identical temperature and pressure parameters were used from the model validation simulations. In order to get an accurate $\Delta\Delta G$ value we performed 50 unique transitions for each mutation. Starting structures for these mutations were obtained by taking one frame every nanosecond from the last 50 ns of the 500 ns trajectory. One nanosecond of equilibrium simulation between each starting frame ensured that each of these transitions was performed with an independent starting point.

Mutational Analysis. A fast-growth approach was used in these simulations. In fast-growth methods, the system does not

reach equilibrium at each λ value, but statistical analysis allows the thermodynamic quantities to be calculated.²⁵ For transitions between the two λ states, the longer the transitions (i.e., smaller $d(\lambda)$, and larger number of time steps) the more accurate the change in energy for a given mutation. At some point longer simulations have diminishing returns.¹³ It is important to optimize for efficiency, so we performed three different test transitions at increasingly large time scales. Transitions of the S16F mutation were run over (100, 500, and 2000 ps). It was determined that 500 ps achieved statistically identical results to 2000 ps. This analysis is presented in Figure S1 for 100 and 2000 ps transitions. For each transition there is a change in enthalpy. The derivative of enthalpy with respect to λ provides a work value described in eq 2.

$$W = \int_0^1 \frac{dH}{d\lambda} d\lambda \quad (2)$$

When we perform many of these transitions (50 forward and 50 reverse) we can check whether the work values follow a normal distribution. Some analysis methods require a Gaussian distribution for both the forward (WT to mutant) and reverse (mutant to WT) transitions. Normal distributions provide evidence that the sampling is not biased by large structural drifts or bimodal distributions. The convergent ΔG mutation lies between the two sample means, as demonstrated by Jarzynski and others.²⁵

Thermodynamic Cycle. A crucial element of the analysis is the application of a shortcut between thermodynamic end states. This cycle allows us to probe the thermodynamic impact of our mutations on the activation cycle without sampling slow transitory behavior which would be computationally impossible using MD alone.

Figure 2 illustrates the thermodynamic cycle implemented for this study. This cycle was inspired by Seeliger and co-workers who used the cycle to probe the effects of mutations on protein thermal stability.¹⁴ The difference in ΔG between the transitions of $A \rightarrow C$ and $B \rightarrow D$ is desired, but a thermodynamically identical transition between states $A \rightarrow B$ and $C \rightarrow D$ can be performed with less effort. The free energy

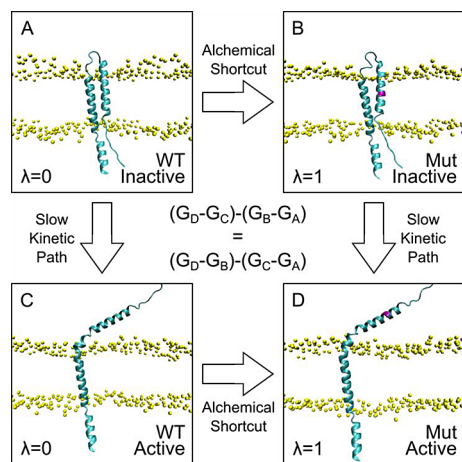


Figure 2. An illustration of the thermodynamic legs between the four states of irsS²¹⁶⁸ investigated in this study: (A) the inactive conformation of the wild type ($\lambda = 0$), (B) the inactive mutant ($\lambda = 1$), (C) the active conformation of the wild type ($\lambda = 0$), and (D) the active mutant. By performing transitions between wild type and mutant ($\lambda = 0 \rightarrow 1$) states transitions are met with low kinetic barriers.

of these four states are state values independent of the path taken. In our simulations we probe $\Delta\Delta G_{\text{act-mut}}$.

RESULTS

Model Validation. To begin the analysis of folding thermodynamics, the structural model was validated first. Penetration depth was calculated by measuring the distance of alpha carbons from the average center of the membrane. RMSF was calculated for alpha carbons using the GROMACS utility *rmsf*. RMSF and relative mobility are not the same measurement, but they should closely correlate to one another. Also, the depth parameter from experiments does not directly provide a length value. So, MD penetration depth is used as a strongly correlated proxy. When comparing the structure of inactive irsS²¹⁶⁸ model constructed for MD and the structural characteristics derived from CW-EPR measurements from Ahammad et al, results are qualitatively similar. Figure 3 overlays the results from the MD model and experiments.

The values around the penetration depth of TMD1 were more stretched across the x-axis due to a slight mismatch in its helical tilt (35° vs 31°). This also explains the lower mobility observed in MD simulations compared to the experimentally derived values. For TMD2, the values for both penetration depth and mobility are characteristically similar to exception of

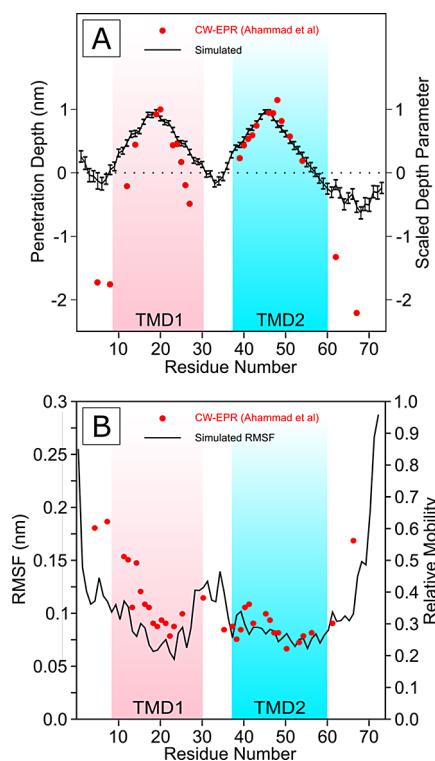


Figure 3. Comparison of simulated model (black line) vs data from CW-EPR measurements performed by Ahammad et al. (red dot).⁴ Adapted from ref 4. Copyright 2020 American Chemical Society. Amino acids in TMD1 are highlighted in pink and those in TMD2 are highlighted in cyan. (A) Average depth of penetration of amino acids within the DMPC membrane where 1 is the center, 0 is either surface of the bilayer. (B) Mobility data comparison between root mean squared fluctuation of the alpha carbons along the backbone of the simulated system (left axis) and the relative mobility of spin-labeled side chains extrapolated from the central line width of CW-EPR spectra (right) where the higher the value the more motion that residue experiences.

an outlying point at residue position G48 where the depth parameter resulted in a value above 1.0, indicating this is the closest amino acid to the center of the membrane. The residues that constitute the terminal tails or the connecting bridge between TMDs explore in the aqueous media. Differences in solvated dynamics are expected due to different time scales between simulations and experiments. Based upon these validating calculations, we proceeded with probing thermodynamics.

$\Delta\Delta G$ Calculations. $\Delta\Delta G$ calculations describing the thermodynamic impact of pinholin mutations were performed for four mutations: S16F, A17Q, A17Q+G21Q, and A17Q+G21Q+G14Q. To show one set of results, data from the S16F mutation are illustrated in Figure 4. Data for other

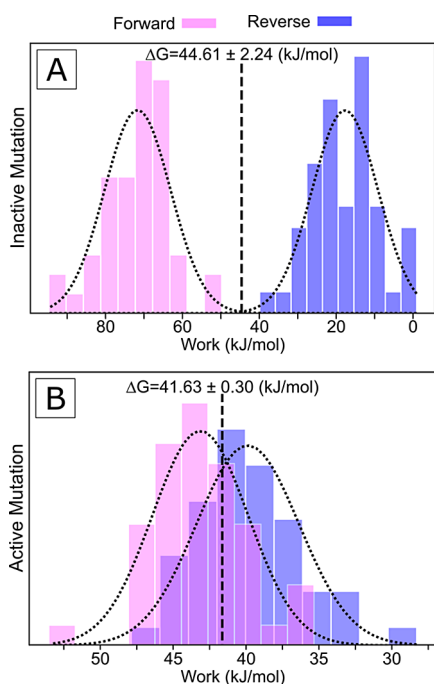


Figure 4. Work distributions calculated using the Bennet acceptance ratio (BAR)²⁶ for (A) inactive and (B) active conformations of irs^{S2168} as they transition forward from WT to mutant (pink) and in reverse from mutant to WT (blue). The Gaussian fit for each distribution is overlaid on top with a black dotted line, and the calculated ΔG value for the transition is marked by a vertical dashed line.

mutations are presented in Figure S2. S16F is an interesting mutation because despite being the most significant change in hydrophobicity out of all the mutations, it shows the most insignificant impact on activation. A17Q, G21Q, and G14Q were selected because the increase in hydrophobicity of TMD1 through these mutations were observed to increase the propensity of activation in experiments.^{1,2b}

We observe that there is much higher overlap between the forward and reverse Gaussian distributions while irs^{S2168} is in the active conformation. While within the membrane, the shift to and from a much larger, more hydrophobic amino acid is expected to have a larger difference in free energy so this comes as no surprise. Additionally, the S16 amino acid is part of a charged region that is theorized to remain partially embedded within the membrane even after activation making the shift to a larger hydrophobic amino acid less

thermodynamically impactful, allowing for a larger overlap between forward and reverse transitions.^{4b}

The S16F mutation is predicted to have little impact on pinholin activation thermodynamics ($\Delta\Delta G = -2.97 \pm 2.26$ kJ/mol). A slight trend toward preferring activation is observed, but error analysis suggests low statistical significance. Positive values of $\Delta\Delta G$ mean the mutation is unfavorable for activation, and negative values of $\Delta\Delta G$ mean it is thermodynamically favorable. A17Q, G14Q, and G21Q mutations replace the WT amino acid with a more hydrophilic mutant, which should increase TMD1's propensity toward activating and exposing itself to an aqueous environment.^{1,2b,27} In addition to the thermodynamic drive toward activation, conformationally these mutations break the important glycine zipper motif that stabilizes the two TMDs in the inactive conformation. As the hydrophilicity of TMD1 increases and the glycine zipper motif is disassembled through additive mutations, the $\Delta\Delta G$ values should steadily increase.

Computational Comparison to Experimental. No prior thermodynamic data are available for direct comparison between theory and experiments. Comparing the data experimentally derived from CW-EPR spectroscopy to the simulated probability values, a bridge can be formed between the two studies. A17Q, A17Q+G21Q, and A17Q+G21Q+G14Q mutations were probed with DEER to measure distance distributions between spin-labeled S8 and L53 amino acids located on TMD1 and TMD2 respectively.^{1,8} When these two spin labels are close to one another the protein is in the inactive state. When these two spin labels move apart, the protein is in the active state. Figure 5 compares the activation signature measured with DEER to alchemically calculated $\Delta\Delta G$ values. The mutations simulated in this study were selected because of the extensive kinetic data collected by both Lorigan and co-workers and Young and co-workers. The mutations are particularly interesting because of they all increase the hydrophilicity of TMD1, but S16F does not seem to increase activation at all. Additionally, the remaining mutations replace parts of the glycine zipper motif which is important to the stability of the inactive form.

As previously stated, the propensity of activation for the irs^{S2168} protein tends to increase with the addition of more hydrophilic amino acids and with the progressive degradation of the glycine zipper motif. The alchemical thermodynamic cycle provides values that agree with experimentally derived probabilities.

Differences in Analysis Methods. To validate the selection of BAR for analyzing these simulations, data were reanalyzed using the Crooks Gaussian Intersection method (CGI)²⁸ and the Jarzynski equality.²⁵ CGI is best used for systems where the work distributions of the forward and reverse processes are Gaussian,²⁸ Jarzynski is best used for calculating free energy changes from nonequilibrium work measurements,²⁵ and BAR is best used for estimating free energy differences with high statistical efficiency from minimal data.²⁶ The pinholin simulations presented here could fall into either case. Therefore, CGI and Jarzynski are used to independently validate the BAR results.²⁹ The data for S16F, A17Q, A17Q+G21Q, and A17Q+G21Q+G14Q mutations is listed in Table 1. While data are similar for all three methods (aside from the A17Q+G21Q+G14Q mutation) there are some key differences to discuss. BAR and Jarzynski methods produce very similar $\Delta\Delta G$ values, however as Jarzynski returns a bootstrap standard error for both forward and reverse

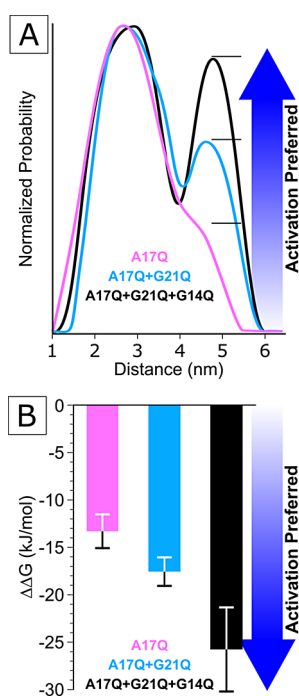


Figure 5. (A) Experimental and (B) computational trends observed when A17Q (pink), A17Q+G21Q (cyan), and A17Q+G21Q+G14Q mutations (black) are applied to irsS²¹68 protein. (A) shows the normalized probability distributions of the measured distance between spin labeled S8 and L53 amino acids. Peaks at 2.6 and 4.9 nm represent inactive and active conformations, respectively. Adapted with permission from ref 1. Copyright 2020 Elsevier. Peaks at 2.6 nm have been normalized to better illustrate the increasing probability of activation. (B) shows the $\Delta\Delta G$ of each mutation set calculated using BAR, along with the standard error. The more negative the $\Delta\Delta G$ value, the more the irsS²¹68 mutant prefers the active conformation.

mutations the magnitude of the error is much larger than BAR. CGI follows the same trend as the other two methods aside from the A17Q+G21Q+G14Q mutation where assumedly the separation between the forward and reverse Gaussian distributions is too large to get an accurate measurement. The general agreement of these three analysis methods apart from the most extreme mutation strengthens confidence in the selection of BAR.

DISCUSSION

This investigation has revealed new capabilities of non-equilibrium alchemical mutation methods. Kinetic boundaries dictate the rate of protein structural transitions. These kinetic boundaries are quite impactful in simulated systems. When kinetic barriers increase, time scales quickly exceed those accessible by standard MD simulations. Thermodynamics, however, describes the equilibrium states of the system. Calculated changes in free energy of a simulated system can be

described without reference to rare events or slow kinetics. The simulation results reported here demonstrate that theory can be effective tool for predicting real-life behavior.

Having a computational procedure that describes thermodynamic driving forces for membrane proteins saves many person-hours of experimental work. Furthermore, atomistic simulations provide atomistic structural detail which can be used to form and test mechanistic hypotheses about thermodynamics. Being able to explore and disqualify mutants through simulation before devoting efforts toward wet-lab experiments streamlines experimental design. Fortunately, the nonequilibrium approaches demonstrated in this study are highly parallelizable. Thus, many low-power computers working in concert can compete with more expensive machines.

The atomistic model produced by the simulations conformed to measurements and analysis from previous studies DEER/CW-EPR that developed a working model of pinholin.^{1–4,7,8} BAR has smaller error bars than CGI and Jarzynski, this is especially true when comparing the A17Q+G21Q+G14Q mutation where the CGI $\Delta\Delta G$ value was roughly half of that measured by BAR. Trends in $\Delta\Delta G$ from BAR agree with the conformational trends reported by Ahammad et al.^{1,4,8} as well as the kinetic trends observed by Young and co-workers.^{2,3,7}

To better understand the effects of these mutations on folding free energy, it is important to consider the chemical characteristics of each amino acid involved. First, S16F mutates serine to phenylalanine. Serine is a polar amino acid with a terminal hydroxyl group. Phenylalanine contains a large hydrophobic side chain. A few competing effects drive S16F activation thermodynamics. S16 packs tightly with the neighboring protein structure in the membrane. When S16 is exposed to the solvent, there should be little change in free energy. F16 does not pack well with its neighbors in the membrane structure. F16 can better form a helix when exposed to the solvent, and its hydrophobicity encourages better helical packing in the solvent exposed state. The relative magnitudes of these packing, folding, and hydrophobic effects drive S16F mutation-activation thermodynamics. Second, A17Q mutates alanine to a glutamine. Alanine is nonpolar while glutamine is polar. Glutamine also has a much larger side chain than alanine. This leads to a double effect. Glutamine disrupts folding in the deactivated state. Glutamine also prefers solvent exposure. This leads to preference for the activated state. G21 and G14, were mutated into glutamine. Each increased the overall hydrophilicity of TMD1 which reasonably increased the thermodynamic drive for TMD1 to become solvent exposed in the active state. Additionally, all three of these amino acids are part of a glycine zipper motif (which includes G10, G14, A17, G21, L28)^{2a,b,5a} helps maintain the inactive folded state. Glycine, being the smallest amino acid with the lowest hydrophobicity on Whimley's scale, can form divots along the

Table 1. Mutation $\Delta\Delta G$ Values and Trends^a

mutation (active-inactive)	BAR $\Delta\Delta G$ [kJ/mol]	CGI $\Delta\Delta G$ [kJ/mol]	Jarzynski $\Delta\Delta G$ [kJ/mol]
S16F	-2.97 ± 2.26	-4.24 ± 2.118	-3.23 ± 4.182
A17Q	-13.05 ± 0.824	-11.88 ± 2.285	-13.02 ± 1.625
A17Q+G21Q	-17.24 ± 1.778	-20.09 ± 4.026	-17.24 ± 3.298
A17Q+G21Q+G14Q	-25.31 ± 3.222	-12.71 ± 9.214	-25.28 ± 7.043

^aError bars are given as standard deviations calculated from forward and reverse variances.

sides of α helices.⁶ When paired with another glycine zipper motif along another α helix, the two divots nest together to oligomerize or fortify the secondary structure of a protein. Removing three key amino acids in the TMD1 glycine zipper motif further reduces the integrity of the interactions with TMD2 that keep pinholin stable in the inactive conformation.

A major impact of mutation on thermodynamics is derived from the contributions that the amino acids make to the protein structure. As the glycine zipper was successively deconstructed, pinholin's structure leaned further toward favoring activation. However, when S16 underwent a significant mutation the difference in free energy of activation was statistically insignificant. If this serine is not important for locking the inactive state, perhaps it is instead more crucial to other processes such as pore formation. Thus, alchemical free energy simulations can be an important tool for supporting, refuting, or developing thermodynamically informed hypotheses about membrane protein structure.

CONCLUSIONS

Within this investigation, we built and validated an accurate atomistic model of irsS68²¹ pinholin within a DMPC bilayer. We probed the thermodynamic contributions of four mutant proteins toward activation. The models provided thermodynamic data that agrees with experimental observations from DEER and CW-EPR. Experiments lack precise atomistic data and give few clues into the thermodynamic driving forces behind activation. Alchemical free energy calculations complement these experiments well by providing models for unobtainable experimental observations.

A campaign of mutations can be performed using scripted alchemical methods, thus yielding quick and valuable information. The effect of mutations on membrane-protein, protein–protein, and protein-solution interactions are complicated due to the multidimensionality of the problem.³⁰ While new methodologies are being studied to experimentally assess mutations, thermodynamic computational methods could provide the bulk free energy results that circumvent the kinetic barriers each of these interactions impose.

Alchemical mutation simulations are a straightforward way to probe the effects of point mutations that reveal themselves over normally inaccessible time scales while under massive kinetic constraints. The proof of concept provided in our work supports the ability of alchemical mutational analysis to explore mutational effects on other membrane bound proteins that undergo large conformational changes. We envision applications toward larger holin proteins, pore forming membrane oligomers, and other crucial membrane-bound proteins.

Kinetics of large conformational changes are difficult to probe using standard MD. To alleviate this shortcoming, several methods have been developed that could be applied to pinholin in future studies. We will highlight three. First, long Markov state models (MSMs) can be used to construct and explore pathways between conformational end states.³¹ Recent advances in constructing efficient MSMs allow researchers to probe events on much longer time scales than would be available to single-trajectory MD models. Second, hyperdynamics can boost a system out of energetic minima to sample rare events like barrier crossings. The boosting potential used in hyperdynamics must be carefully designed and verified to ensure valid results.³² Third, infrequent metadynamics is a related method to hyperdynamics. With infrequent metadynamics, a time-dependent potential is added

to a collective variable infrequently to allow escape out of an energy minimum.³³ After the system escapes the minimum, the simulation is halted, and additional trajectories are produced. A set of trajectories is analyzed to recover kinetics. However, there have been recent criticisms of the application of infrequent metadynamics, so researchers should use this method with caution.^{33b}

The atomistic models used here were based around putative models supported by experimental analysis. However, some mutations may cause additional low energy conformers to arise due to hydrogen or sulfur bonding between point mutations. This would cause the thermodynamic cycle to have more than four states which might not be described by experimental data.³⁴ To address this, lower cost methods like those employed by Rosetta could be used to probe the energy landscape to identify all probable conformers to analyze further possible changes in state.³⁵ End-point methods like molecular mechanics with generalized born and surface area solvation (MM/GBSA) or molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) can report the change in attractive, repulsive, hydrogen bond, solvation, and more energy values across the nonequilibrium transformation simulations.³⁶ Additionally, with robust computational and experimental resources, this alchemical method could be automated to perform alanine scanning alongside combinatorial alanine scanning.³⁷

Taken together, the simulations described here provide evidence that alchemical mutation simulations can be valuable counterparts to experiments that probe large structural changes of proteins embedded in lipid membranes. Systems like these have important implications for biophysics, metabolism, drug design, biocatalysis, and human health. A better understanding of pinholin activation would greatly improve the feasibility of phage therapy^{1,2d,38} an antibacterial method of using whole bacteriophages to kill bacteria. Phage therapy is widely believed to be an effective method to keep ahead of bacterial antibiotic resistance. It is also believed that through targeted mutations to the phage, the efficacy can be tuned and optimized to treat a wide array of bacteria. Bacteria lysing being one of the main concerns within the field, a proper mechanistic understanding of the responsible proteins is necessary.

Improved phages could be designed through predictive modeling.³⁹ Pinholins could be tuned through mutation to drive thermodynamic equilibrium toward activation. Alternatively, specialized phages could be designed to be compatible with certain membrane lipid compositions. Thus, pinholins could be designed to preferentially kill classes of bacteria in which the pinholin is activated by the bacteria's lipid profile.⁴⁰ Predictive modeling is helpful for reducing the strain on experimentalists by using distributed computing to rationally reduce the design space. Based upon the results presented in this work, we believe there is sufficient evidence to use alchemical mutations as a screening tool to test hypotheses *in silico*. For example, an initial computational alanine scan could identify the most impactful mutation sites. Those sites could then be probed for mutations to other classes of residues (e.g., positively charged, negatively charged, polar, nonpolar). Taken altogether, the results presented in this paper give researchers a new tool to understand and predict the thermodynamics of protein mutations and their effects on large conformational changes for membrane-bound systems.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.4c03302>.

Demonstration of convergence of free energy calculations based on transition time and list of contents of simulation boxes (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Theodore S. Kalbfleisch – Department of Chemical Engineering, University of Louisville, Louisville, Kentucky 40292, United States; Email: tskalb02@louisville.edu

Vance W. Jaeger – Department of Chemical Engineering, University of Louisville, Louisville, Kentucky 40292, United States; orcid.org/0000-0003-2294-4451; Email: vwjaeg01@louisville.edu

Authors

Tanbir Ahammad – Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056, United States

Gary A. Lorigan – Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpcb.4c03302>

Author Contributions

V.W.J. and T.S.K. conceived and designed the study. T.S.K. conducted the simulations. G.A.L. and T.A. provided experimental data. T.S.K. also analyzed the results and drafted the manuscript. V.W.J., G.A.L., and T.A. validated and reviewed the manuscript. T.S.K. and V.W.J. edited the manuscript. V.W.J. supervised the project.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was conducted in part using the resources of the University of Louisville's research computing group and the Cardinal Research Cluster. G.A.L. acknowledges support from NSF CHE-2305834 grant.

■ REFERENCES

- (1) Ahammad, T.; Khan, R. H.; Sahu, I. D.; Drew, D. L.; Faul, E.; Li, T.; McCarrick, R. M.; Lorigan, G. A. Pinholin S21 mutations induce structural topology and conformational changes. *Biochim. Biophys. Acta Biomembr.* **1863**, 2021 (12), No. 183771.
- (2) (a) Pang, T.; Savva, C. G.; Fleming, K. G.; Struck, D. K.; Young, R. Structure of the lethal phage pinhole. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (45), 18966–18971. (b) Pang, T.; Park, T.; Young, R. Mutational analysis of the S21 pinholin. *Mol. Microbiol.* **2010**, *76* (1), 68–77. (c) Young, R. Bacteriophage holins: deadly diversity. *J. Mol. Microbiol. Biotechnol.* **2002**, *4* (1), 21–36. (d) Young, R. Bacteriophage lysis: mechanism and regulation. *Microbiol. Rev.* **1992**, *56* (3), 430–481.
- (3) Park, T.; Struck, D. K.; Dankenbring, C. A.; Young, R. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. *J. Bacteriol.* **2007**, *189* (24), 9135–9139.
- (4) (a) Ahammad, T.; Drew, D. L.; Khan, R. H.; Sahu, I. D.; Faul, E.; Li, T.; Lorigan, G. A. Structural Dynamics and Topology of the Inactive Form of S21 Holin in a Lipid Bilayer Using Continuous-

Wave Electron Paramagnetic Resonance Spectroscopy. *J. Phys. Chem. B* **2020**, *124* (26), 5370–5379. (b) Ahammad, T.; Drew, D. L.; Sahu, I. D.; Serafin, R. A.; Clowes, K. R.; Lorigan, G. A. Continuous Wave Electron Paramagnetic Resonance Spectroscopy Reveals the Structural Topology and Dynamic Properties of Active Pinholin S2168 in a Lipid Bilayer. *J. Phys. Chem. B* **2019**, *123* (38), 8048–8056.

(5) (a) Steger, L. M. E.; Kohlmeyer, A.; Wadhwani, P.; Bürck, J.; Strandberg, E.; Reichert, J.; Grage, S. L.; Afonin, S.; Kempfer, M.; Görner, A. C.; Koch, J.; Walther, T. H.; Ulrich, A. S. Structural and functional characterization of the pore-forming domain of pinholin S(21)68. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (47), 29637–29646. (b) Park, T.; Struck, D. K.; Dankenbring, C. A.; Young, R. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. *J. Bacteriol.* **2007**, *189* (24), 9135–9.

(6) Kim, S.; Jeon, T.-J.; Oberai, A.; Yang, D.; Schmidt, J. J.; Bowie, J. U. Transmembrane glycine zippers: Physiological and pathological roles in membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (40), 14278–14283.

(7) Pang, T.; Park, T.; Young, R. Mapping the pinhole formation pathway of S21. *Mol. Microbiol.* **2010**, *78* (3), 710–9.

(8) Ahammad, T.; Drew, D. L.; Sahu, I. D.; Khan, R. H.; Butcher, B. J.; Serafin, R. A.; Galende, A. P.; McCarrick, R. M.; Lorigan, G. A. Conformational Differences Are Observed for the Active and Inactive Forms of Pinholin S21 Using DEER Spectroscopy. *J. Phys. Chem. B* **2020**, *124* (50), 11396–11405.

(9) Lindahl, E.; Sansom, M. S. P. Membrane proteins: molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2008**, *18* (4), 425–431.

(10) Zou, Y.; Ewalt, J.; Ng, H. L. Recent Insights from Molecular Dynamics Simulations for G Protein-Coupled Receptor Drug Discovery. *Int. J. Mol. Sci.* **2019**, *20* (17), 4237.

(11) Bilston, L. E.; Mylvaganam, K. Molecular simulations of the large conductance mechanosensitive (MscL) channel under mechanical loading. *FEBS Lett.* **2002**, *512* (1–3), 185–90.

(12) Kozono, D.; Yasui, M.; King, L. S.; Agre, P. Aquaporin water channels: atomic structure molecular dynamics meet clinical medicine. *J. Clin. Invest.* **2002**, *109* (11), 1395–9.

(13) Mey, A. S. J. S.; Allen, B. K.; Bruce McDonald, H. E.; Chodera, J. D.; Hahn, D. F.; Kuhn, M.; Michel, J.; Mobley, D. L.; Naden, L. N.; Prasad, S.; Rizzi, A.; Scheen, J.; Shirts, M. R.; Tresadern, G.; Xu, H. Best Practices for Alchemical Free Energy Calculations [Article v1.0]. *Living J. Comput. Mol. Sci.* **2020**, *2* (1), No. 18378.

(14) Seeliger, D.; de Groot, B. L. Protein Thermostability Calculations Using Alchemical Free Energy Simulations. *Biophys. J.* **2010**, *98* (10), 2309–2316.

(15) Pedretti, A.; Mazzolari, A.; Gervasoni, S.; Fumagalli, L.; Vistoli, G. The VEGA suite of programs: an versatile platform for cheminformatics and drug design projects. *Bioinformatics* **2020**, *37* (8), 1174–1175.

(16) (a) Jo, S.; Kim, T.; Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS One* **2007**, *2* (9), No. e880. (b) Lee, J.; Patel, D. S.; Stähle, J.; Park, S.-J.; Kern, N. R.; Kim, S.; Lee, J.; Cheng, X.; Valvano, M. A.; Holst, O.; Knirel, Y. A.; Qi, Y.; Jo, S.; Klauda, J. B.; Widmalm, G.; Im, W. CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput.* **2019**, *15* (1), 775–786. (c) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* **2009**, *97* (1), 50–58. (d) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **2014**, *35* (27), 1997–2004. (e) 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), 1859–1865.

(17) (a) Bjellmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E. Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *J. Chem. Theory Comput.* **2010**, *6*

- (2), 459–466. (b) Huang, J.; MacKerell, A. D., Jr CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34* (25), 2135–2145.
- (18) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (19) (a) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. (b) Páll, S.; Zhmurov, A.; Bauer, P.; Abraham, M.; Lundborg, M.; Gray, A.; Hess, B.; Lindahl, E. Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. *J. Chem. Phys.* **2020**, *153* (13), No. 134110.
- (20) Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B. L. pmx: Automated protein structure and topology generation for alchemical perturbations. *J. Comput. Chem.* **2015**, *36* (5), 348–354.
- (21) Berendsen, H. J.; Postma, J. v.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.
- (22) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126* (1), No. 014101.
- (23) (a) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domanski, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O., MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. In *Conference: PROC. OF THE 15th PYTHON IN SCIENCE CONF. (SCIPY 2016)*; 2016–07–11 - 2016–07–11; United States, 2019; Medium: ED; Size: 98;. (b) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32* (10), 2319–27.
- (24) Khan, R. H.; Ahammad, T.; Sahu, I. D.; Rotich, N. C.; Daufel, A.; Lorigan, G. A. Determining the helical tilt angle and dynamic properties of the transmembrane domains of pinholin S2168 using mechanical alignment EPR spectroscopy. *Biochim. Biophys. Acta Biomembr.* **1865**, 2023 (5), No. 184154.
- (25) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78* (14), 2690–2693.
- (26) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22* (2), 245–268.
- (27) Wimley, W. C.; White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* **1996**, *3* (10), 842–848.
- (28) Crooks, G. E. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Stat. Phys.* **1998**, *90* (5), 1481–1487.
- (29) (a) Goette, M.; Grubmüller, H. Accuracy and convergence of free energy differences calculated from nonequilibrium switching processes. *J. Comput. Chem.* **2009**, *30* (3), 447–456. (b) Aldeghi, M.; de Groot, B. L.; Gapsys, V., Accurate Calculation of Free Energy Changes upon Amino Acid Mutation. In *Computational Methods in Protein Evolution*, Sikosek, T., Ed.; Springer: New York: New York, NY, 2019; 19–47.
- (30) Patel, D.; Patel, J. S.; Ytreberg, F. M. Implementing and Assessing an Alchemical Method for Calculating Protein–Protein Binding Free Energy. *J. Chem. Theory Comput.* **2021**, *17* (4), 2457–2464.
- (31) Pande, V. S.; Beauchamp, K.; Bowman, G. R. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **2010**, *52* (1), 99–105.
- (32) (a) Voter, A. F. Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events. *Phys. Rev. Lett.* **1997**, *78* (20), 3908–3911. (b) Bal, K. M.; Neyts, E. C. Merging Metadynamics into Hyperdynamics: Accelerated Molecular Simulations Reaching Time Scales from Microseconds to Seconds. *J. Chem. Theory Comput.* **2015**, *11* (10), 4545–4554. (c) Vigil, G.; Xu, Z.; Steinberg, S.; Israelachvili, J. Interactions of Silica Surfaces. *J. Colloid Interface Sci.* **1994**, *165* (2), 367–385.
- (33) (a) Blumer, O.; Reuveni, S.; Hirshberg, B. Short-Time Infrequent Metadynamics for Improved Kinetics Inference. *J. Chem. Theory Comput.* **2024**, *20* (9), 3484–3491. (b) Ray, D.; Parrinello, M. Kinetics from Metadynamics: Principles, Applications, and Outlook. *J. Chem. Theory Comput.* **2023**, *19* (17), 5649–5670.
- (34) Lorch, M.; Mason, J. M.; Sessions, R. B.; Clarke, A. R. Effects of Mutations on the Thermodynamics of a Protein Folding Reaction: Implications for the Mechanism of Formation of the Intermediate and Transition States. *Biochemistry* **2000**, *39* (12), 3480–3485.
- (35) Pietrucci, F. Strategies for the exploration of free energy landscapes: Unity in diversity and challenges ahead. *Rev. Phys.* **2017**, *2*, 32–45.
- (36) Alford, R. F.; Leaver-Fay, A.; Jeliazkov, J. R.; O’Meara, M. J.; DiMaio, F. P.; Park, H.; Shapovalov, M. V.; Renfrew, P. D.; Mulligan, V. K.; Kappel, K.; Labonte, J. W.; Pacella, M. S.; Bonneau, R.; Bradley, P.; Dunbrack, R. L.; Das, R.; Baker, D.; Kuhlman, B.; Kortemme, T.; Gray, J. J. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.* **2017**, *13* (6), 3031–3048.
- (37) (a) Kortemme, T.; Kim, D. E.; Baker, D. Computational Alanine Scanning of Protein–Protein Interfaces. *Sci. STKE* **2004**, 219, pl2. (b) Morrison, K. L.; Weiss, G. A. Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* **2001**, *5* (3), 302–307.
- (38) (a) Goodridge, D. L. Designing Phage Therapeutics. *Curr. Pharm. Biotechnol.* **2010**, *11* (1), 15–27. (b) Adamczyk-Popławska, M.; Tracz-Gaszewska, Z.; Lasota, P.; Kwiatek, A.; Piekarowicz, A. Haemophilus influenzae HP1 Bacteriophage Encodes a Lytic Cassette with a Pinholin and a Signal-Arrest-Release Endolysin. *Int. J. Mol. Sci.* **2020**, *21* (11), 4013. (c) Chan, B. K.; Abedon, S. T.; Loc-Carrillo, C. Phage cocktails and the future of phage therapy. *Future Microbiol.* **2013**, *8* (6), 769–783. (d) Curtright, A. J.; Abedon, S. T. Phage Therapy: Emergent Property Pharmacology. *J. Bioanal. Biomed.* **2012**, *2012*, 1–14.
- (39) Lood, C.; Boeckaerts, D.; Stock, M.; De Baets, B.; Lavigne, R.; van Noort, V.; Briers, Y. Digital phagograms: predicting phage infectivity through a multilayer machine learning approach. *Curr. Opin. Virol.* **2022**, *52*, 174–181.
- (40) Hyman, P. Phages for Phage Therapy: Isolation, Characterization, and Host Range Breadth. *Pharmaceuticals (Basel)* **2019**, *12* (1), 35.