

# Kinetic pathway analysis of an $\alpha$ -helix in two protonation states: Direct observation and optimal dimensionality reduction

Cite as: J. Chem. Phys. **150**, 074902 (2019); <https://doi.org/10.1063/1.5082192>

Submitted: 19 November 2018 . Accepted: 16 January 2019 . Published Online: 21 February 2019

 Gouri S. Jas, Ed W. Childs, and  Krzysztof Kuczera



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Large fatty acid-derived A \$\beta\$ 42 oligomers form ring-like assemblies](#)

The Journal of Chemical Physics **150**, 075101 (2019); <https://doi.org/10.1063/1.5082659>

[The role of hydrophobicity in the cold denaturation of proteins under high pressure: A study on apomyoglobin](#)

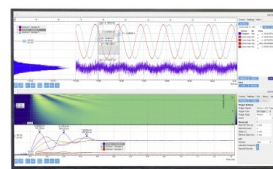
The Journal of Chemical Physics **150**, 075102 (2019); <https://doi.org/10.1063/1.5080942>

[Committers, first-passage times, fluxes, Markov states, milestones, and all that](#)

The Journal of Chemical Physics **150**, 054106 (2019); <https://doi.org/10.1063/1.5079742>

Challenge us.

What are your needs for  
periodic signal detection?



Zurich  
Instruments



# Kinetic pathway analysis of an $\alpha$ -helix in two protonation states: Direct observation and optimal dimensionality reduction

Cite as: J. Chem. Phys. 150, 074902 (2019); doi: 10.1063/1.5082192

Submitted: 19 November 2018 • Accepted: 16 January 2019 •

Published Online: 21 February 2019



View Online



Export Citation



CrossMark

Gouri S. Jas,<sup>1,a)</sup> Ed W. Childs,<sup>2</sup> and Krzysztof Kuczera<sup>3,4</sup>

## AFFILIATIONS

<sup>1</sup>Department of Pharmaceutical Chemistry, The University of Kansas, Lawrence, Kansas 66047, USA

<sup>2</sup>Department of Surgery, Morehouse School of Medicine, Atlanta, Georgia 30310, USA

<sup>3</sup>Department of Molecular Biosciences, The University of Kansas, Lawrence, Kansas 66047, USA

<sup>4</sup>Department of Chemistry, The University of Kansas, Lawrence, Kansas 66045, USA

<sup>a)</sup> Author to whom correspondence should be addressed: [gouri.jas@physicianscientists.org](mailto:gouri.jas@physicianscientists.org)

## ABSTRACT

Thermodynamically stable conformers of secondary structural elements make a stable tertiary/quaternary structure that performs its proper biological function efficiently. Formation mechanisms of secondary and tertiary/quaternary structural elements from the primary structure are driven by the kinetic properties of the respective systems. Here we have carried out thermodynamic and kinetic characterization of an alpha helical heteropeptide in two protonation states, created with the addition and removal of a proton involving a single histidine residue in the primary structure. Applying far-UV circular dichroism spectroscopy, the alpha helix is observed to be significantly more stable in the deprotonated state. Nanosecond laser temperature jump spectroscopy monitoring time-resolved tryptophan fluorescence on the protonated conformer is carried out to measure the kinetics of this system. The measured relaxation rates at a final temperature between 296K and 314 K generated a faster component of 20 ns–11 ns and a slower component of 314 ns–198 ns. Atomically detailed characterization of the helix-coil kinetic pathways is performed based on all-atom molecular dynamics trajectories of the two conformers. Application of clustering and kinetic coarse-graining with optimum dimensionality reduction produced description of the trajectories in terms of kinetic models with two to five states. These models include aggregate states corresponding to helix, coil, and intermediates. The “coil” state involves the largest number of conformations, consistent with the expected high entropy of this structural ensemble. The “helix” aggregate states are found to be mixed with the full helix and partially folded forms. The experimentally observed higher helix stability in the deprotonated form of the alpha helical heteropeptide is reflected in the nature of the “helix” aggregate state arising from the kinetic model. In the protonated form, the “coil” state exhibits the lowest free energy and longest lifetime, while in the deprotonated form, it is the “helix” that is found to be most stable. Overall, the coarse grained models suggest that the protonation of a single histidine residue in the primary structure induces significant changes in the free energy landscape and kinetic network of the studied helix-forming heteropeptide.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5082192>

## INTRODUCTION

In order to carry out a proper biological function, a protein must exist in a robust form with respect to its physiological environment. The robustness of a functional protein is governed by the composition of the stable secondary

structural conformers. A minor adjustment in the primary structure, such as the protonation state of a histidine residue, can significantly alter the stability of the secondary structural elements, as well as the overall system structure.<sup>1–8</sup> Thermodynamics plays an important role in maintaining these stable states. The formation mechanisms of secondary structural

elements, capturing the pathways of formation of a functional protein, on the other hand, are driven by the kinetic properties of these systems. These kinetic events can take place in an ultrafast time scale, such as in several nanoseconds, to a much longer time domain. Relaxation dynamics, breaking and formation of the secondary structural elements, mostly occur in a faster time scale compared to a whole protein. Improvements in instrumentation with significantly higher time resolution have enabled detecting formation events that take place in shorter time scales.<sup>9-11</sup> Laser temperature jump (t-jump) spectroscopy, with picosecond to nanosecond time resolution, has made it possible to measure the relaxation dynamics of helix forming peptides from five to twenty-one residues.<sup>12-14</sup> These measurements determine not only the time scale associated with the breaking and formation of the global structure but also the pathway associated with it. A measured kinetic trace fitted with a single exponential function will suggest that the progression follows a two-state pathway, i.e., an unfolded state of the structure going to a folded state without visiting any other states along the way. A trace fitted with a bi-exponential function will point to a pathway involving more than two states. The mechanism of how the transition from the unfolded to the folded state is achieved is yet to be fully characterized. In principle, the state-of-the-art experimental measurements combined with long-time molecular dynamics simulations and a sophisticated theoretical approach can unravel the nature of structure formation as it progresses from a disordered to a more ordered state. A molecular dynamics simulation with a very long trajectory can essentially describe the transition pathways, the intermediates, and the rates.

Recently, attempts have been made to describe helix folding computationally, using molecular dynamics simulations.<sup>15-21</sup> It was suggested that helix folding progresses from multiple nucleation sites to fragmented helices.<sup>16,21</sup> The Milestoning approach was undertaken to describe the helix folding rates and pathway from a pre-determined path.<sup>16</sup> Ultrafast laser temperature jump spectroscopy was applied to measure the relaxation dynamics of a single alpha helical turn formation.<sup>12</sup> Five-residue alanine homopolymers showed the formation of the stable helix in solution at physiologically relevant pH.<sup>22</sup> These experimental measurements provide a very important foundation and computationally easily achievable time scale for an atomically detailed description of the relaxation pathway. A full description of helix-coil kinetics with atomic detailed resolution is yet to be achieved. Fast kinetic measurements are revealing detailed features of the relaxation dynamics, and sophisticated theoretical approaches are also being tested in an attempt to characterize the measured pathways, encompassing Markov state modeling<sup>23-25</sup> and milestoning.<sup>17</sup> Recently developed optimal dimensionality reduction by Szabo and Hummer<sup>26</sup> was able to describe a folding pathway, with two to four states, in a penta-Ala system that can be compared to a master equation in the model based on molecular dynamics simulation. Interesting approaches to kinetic coarse graining, with the PCCA+,<sup>24</sup> are promising and opening important testing grounds.

We have reported circular dichroism (CD) and t-jump spectroscopic studies of the protonated form of the alpha-helical peptide (WH21<sup>+</sup>), elsewhere.<sup>27</sup> CD and fluorescence energy transfer measurements of the unprotonated (WH21<sup>0</sup>) form of the 21-residue helical heteropeptide have been reported elsewhere.<sup>28</sup> We have described long-term molecular dynamics (MD) trajectories for WH21<sup>+</sup><sup>27</sup> and WH21<sup>0</sup>,<sup>28</sup> and a kinetic network analysis of WH21<sup>+</sup>.<sup>27</sup> In this work, we extend these joint experimental and computational studies. Experimentally, we present an extensive thermodynamic analysis of the CD data and new t-jump measurements with superior signal-to-noise for WH21<sup>+</sup>. Computationally, we have performed an extensive trajectory analysis for two conformers of this peptide and a kinetic coarse-graining analysis of WH21<sup>0</sup> folding.

In the experimental component, we applied far-ultraviolet (UV) circular dichroism (CD) spectroscopy and a nanosecond laser temperature jump (t-jump) to the twenty-one residue  $\alpha$ -helical heteropeptide WH21. Thermodynamic characterization has been conducted based on the CD spectra measured between 263 K and 363 K. Extended analysis of the Singular Value Decomposition (SVD) amplitude vectors yielded equilibrium constants, as well as changes in enthalpy, entropy, and Gibbs free energy as a function of temperature for both the neutral and protonated forms of this peptide. For WH21<sup>+</sup>, relaxation dynamics are measured in the nanosecond to microsecond time scale with the application of the t-jump with significantly improved signal to noise. Kinetic measurements conducted at a final temperature between 296 K and 314 K, vividly captured two different time constants for each trace. An observed faster component lies between 22 ns and 11 ns. This observation strongly directs us to pathways of a non-two state behavior during the relaxation process. The measured slower components are found to be between 317 ns and 198 ns. The slower component obtained from the analysis of the measured kinetic data is associated with the global structure formation.

In the computational component of this study, we employ two previously described<sup>27,28</sup> molecular dynamics (MD) trajectories of WH21 in explicit solvent, generated with the CHARMM 36<sup>29</sup> protein parameter. A 13  $\mu$ s length trajectory of WH21<sup>0</sup>, from which structural and kinetic data of the neutral peptide were extracted,<sup>28</sup> is used to generate simplified kinetically coarse-grained models of low dimensionality. Additionally, a 12  $\mu$ s length trajectory, which was the basis for our conventional<sup>27</sup> and kinetic coarse-graining studies<sup>28</sup> of the protonated conformer, is used here to compare the effect of protonation of a single histidine residue on folding. Our novel kinetic network analysis approach, grounded on the kinetic matrix, is aimed at monitoring microscopic details involving the helix folding pathway. The conformations sampled from the MD trajectory are first clustered into a set of microstates, for which a lifetime-based kinetic analysis is performed.<sup>30</sup> Using the PCCA+ algorithm,<sup>24</sup> these microstates are then aggregated into a small number of kinetically metastable states. Finally, these aggregated, kinetically coarse-grained states are the basis for generating an optimal reduced dimensionality rate matrix<sup>26</sup> and transition path theory analysis.<sup>31,32</sup>

Models of dimensionalities between 2 and 5 provide new insights into the response of the folding pathway of the alpha helical heteropeptide to the change in the protonation state. Our modeling results show that both the helix and coil states are heterogeneous and allow the structural characterization of several folding intermediates. The models also indicate that helix folding occurs through multiple pathways, with the number of paths increasing with increased resolution of the model. The simulations indicate that the protonation of a single histidine residue in the primary structure induces significant changes in the free energy landscape and kinetic network of this helix-forming peptide compared to a non-protonated state.

## METHODS

### Experimental

#### Materials

The twenty-one residue alpha helical heteropeptide, with a tryptophan in position one and a histidine in position five ( $i, i + 4$ ), was purchased from Gen Script USA, Inc., NJ. The purity was greater than 98%, determined with high-performance liquid chromatography (HPLC). The N and C termini of the heteropeptide are capped with an acetyl and an amide group, respectively. There are 22 residues in this peptide including the N and C terminal caps, as acetyl and amide groups are considered half a residue each. Sequence details of this heteropeptide are shown elsewhere.<sup>10,11</sup>

#### Circular dichroism

Peptide secondary structural elements were characterized with far-UV circular dichroism spectroscopy as a function of temperature in the protonated and deprotonated form with a JASCO 815 spectropolarimeter (Tokyo, Japan) using a cylindrical cell with a path length of 0.5 mm and concentrations of  $\sim 220 \mu\text{M}$ . Scans between 263 and 363 K are recorded in  $10^\circ$  increments. Thermodynamic parameters and the fraction helix, as a function of temperature, are obtained from the amplitude vectors from the SVD components and fit with a two-state equilibrium model EXAM.<sup>33</sup> The two-state fit obtained a temperature-independent parameter of the transition enthalpy  $\Delta H^\circ$  and entropy  $\Delta S^\circ$ , and the mid-point temperature  $T_m = \Delta H^\circ / \Delta S^\circ$ . The Gibbs free energy was calculated from  $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$ , and the equilibrium constant  $K$  was found from  $\Delta G^\circ = -RT \ln(K)$ . With a conservative estimate, the errors in these thermodynamic parameters are estimated to be  $\sim 5\%$ , similar to previous analyses,<sup>10,11</sup> Singular value decomposition (SVD) was applied to the CD scans between 170 and 260 nm to resolve overlapping spectral components. In this study, we have made use of the first and second components of the wavelength dependent circular dichroism spectra, from multiple components shown elsewhere,<sup>28</sup> obtained with the application of singular value decomposition. Amplitude vectors from the component one are used to obtain a most comprehensive thermodynamic characterization of the studied peptide with respect to change in the protonation states. An additional thermodynamic analysis of

the CD data incorporating a heat capacity change of unfolding and thus temperature-dependent values of  $\Delta H^\circ$  and  $\Delta S^\circ$  is also performed, as described in the [supplementary material](#).

### Temperature jump kinetics

Instrumentation of the laser temperature jump apparatus used in these kinetic measurements has been described in detail elsewhere.<sup>10,11</sup> The initial rapid decrease in fluorescence intensity results from a decrease in the intrinsic tryptophan fluorescence from the rise in temperature due to temperature pulse. Temperature jumps of 8–10 K were generated by absorption of a near-infrared laser pulse ( $1.54 \mu\text{m}$ ) by the aqueous solution. Time resolved tryptophan fluorescence intensity change was monitored to record relaxation dynamics. Tryptophan was excited with a continuous ultraviolet probe beam (264 nm), and the emitted light was measured using band-pass filters which transmitted light between 320 and 400 nm. The kinetic experiments were carried out in a  $0.05 \times 1.0 \text{ cm}^2$  cuvette at a sample concentration of  $\sim 100 \mu\text{M}$  and in 20 mM acetate buffer at pH 4.8. The measured kinetics from the change in fluorescence of the alpha helical heteropeptide were recorded over a temperature range of 296–314 K with much improved *signal-to-noise* ratio. This effort is made to capture the faster component in this kinetic process more clearly as this is a critical signal associated with the short-lived structure formation. The measured similar relaxation dynamics and the fits to the data are shown elsewhere.<sup>28</sup> In this study, we have used relative amplitude of the fast component from a biexponential fit and from the fits with a single exponential function to show that the helix population is decreasing as a function of increasing temperature as demonstrated by the decreased amplitudes in kinetic measurement.

### Computational

The MD trajectories used in this work have been presented, elsewhere.<sup>27,28</sup> Our kinetic analysis of the neutral peptide (WH21<sup>0</sup>) is based on a  $13 \mu\text{s}$  MD trajectory generated with GROMACS<sup>34</sup> with the CHARMM 36<sup>29,35</sup> force field and TIP3P water<sup>28,36</sup> at 300 K. For comparison, results from the  $12 \mu\text{s}$  MD simulation of the protonated form of the peptide with the same force field and conditions were used previously.<sup>27</sup> Technical details are presented in the [supplementary material](#).

Hydrogen bonds were counted as formed when the  $\text{O} \cdots \text{N}$  distance was below  $3.6 \text{ \AA}$ , and a residue was considered to be in the  $\alpha$ -helical conformation when the  $(\varphi, \psi)$  values were found to be within a  $20^\circ$  radius of an ideal helical structure ( $-62^\circ, -41^\circ$ ). Structure clustering was done with the gromos algorithm.<sup>25</sup> Statistical error estimates were obtained by dividing data into ten consecutive blocks and calculating the standard error of the mean at the 95% confidence level.

In the analysis of peptide kinetics, two approaches were employed. First, time autocorrelation functions (ACFs)  $C_1(t) = \langle (x(0) - \langle x \rangle) (x(t) - \langle x \rangle) \rangle$  were calculated, and  $x$  being the number of  $\alpha$ -helical hydrogen bonds and the number of residues

in the  $\alpha$ -helical region of the Ramachandran map. These functions were fitted to two-exponential decays, and the longer time scale  $\tau_2$  was used to estimate the global relaxation time of the peptide. In the second approach, we performed a kinetic coarse graining of the MD trajectory. This involves first discretization of the conformation space explored in the MD into a set of microstates by C $\alpha$  atom root mean square deviation (RMSD), followed by a lifetime-based kinetic analysis of transitions between microstates.<sup>30</sup> The method for obtaining transition counts between microstates, elements of the kinetic matrix **K**, and the transition matrix **T** is described in more detail elsewhere<sup>27</sup> and in the [supplementary material](#). The eigenvalues of the kinetic matrix **K** were used to estimate the system relaxation times. The microstate core radius was selected such that the global relaxation time  $\tau_2$  based on the  $C_1(t)$  functions agreed with  $\lambda_2$ , the lowest non-zero eigenvalue of **K**, i.e.,  $\tau_2 \approx 1/\lambda_2$ . For analysis of kinetic results, we used tools from the EMMA software package—PCCA+ and TPT.<sup>31</sup> Finally, optimal reduced kinetic matrices were generated with the formalism of Hummer and Szabo,<sup>26</sup> based on the coarse-graining sets obtained through PCCA+.<sup>24</sup> In PCCA+, the initial set of microstates is aggregated in a smaller set of kinetically metastable states, which exhibit faster interconversions within and slower conversions between the different sets. The aggregation is based on analyzing the sign structure of the successive eigenvectors of **T** (or **K**).<sup>24</sup>

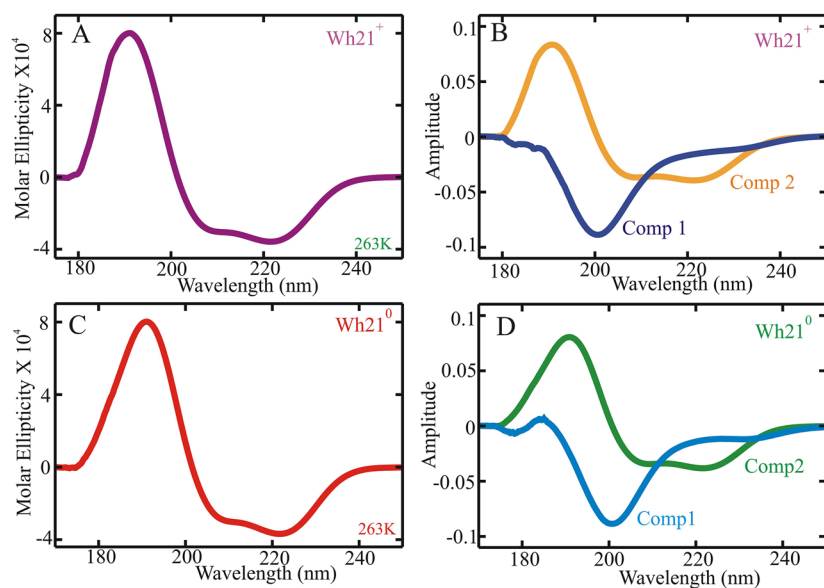
## RESULTS AND DISCUSSION

### Experimental

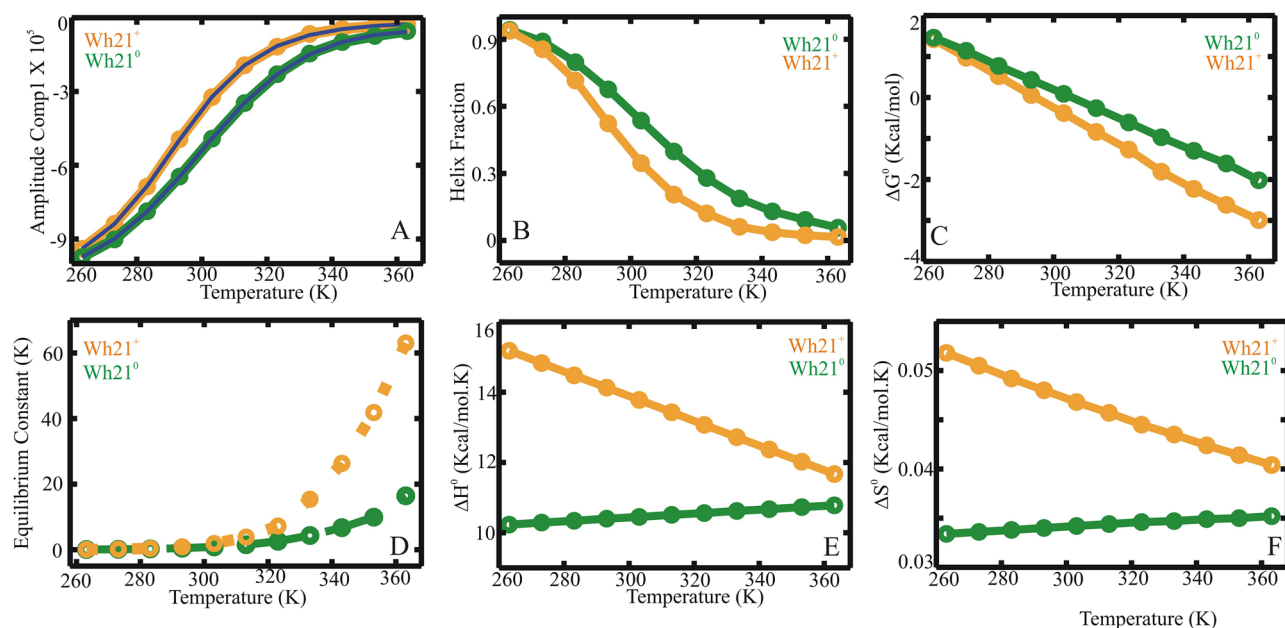
Presented in [Fig. 1](#) are the representative far UV circular dichroism (CD) spectra of two forms of the studied twenty-one residue  $\alpha$ -helical heteropeptide, with a protonated (WH21<sup>+</sup>) and a neutral (WH21<sup>0</sup>) histidine. A complete

set of CD spectra measured between 263 K and 363 K have been presented elsewhere.<sup>28</sup> In [Fig. 1\(a\)](#), is shown the CD spectrum of WH21<sup>+</sup> at 263 K, where histidine is protonated. The measured spectral features shown here exhibited two minima at 222 nm and 210 nm and the maximum at 190 nm. These spectral features represent a classic alpha helical structure, and the helicity is observed to decrease with increasing temperature. To resolve the components that are present in the spectra, we applied singular value decomposition (SVD)<sup>15</sup> on the set of temperature dependent CD spectra, shown in [Fig. 1\(b\)](#). Shown in [Fig. 1\(c\)](#) are temperature dependent CD spectra of WH21<sup>0</sup> at 263 K, where histidine is in the neutral/non-protonated state. [Figure 1\(d\)](#) shows two SVD components of WH21<sup>0</sup>. Detailed temperature dependent CD spectra are shown elsewhere.<sup>28</sup> A representative CD spectrum at 263 K is shown here to describe the basis of our new expanded comprehensive thermodynamic analysis.

In [Fig. 2\(a\)](#), shown are the amplitude vectors of component one as a function of temperature for both WH21<sup>+</sup> and WH21<sup>0</sup>. Compared to WH21<sup>+</sup>, WH21<sup>0</sup> vectors are shifted to the right, showing an increased stability of the  $\alpha$ -helix compared to WH21<sup>+</sup>. A two-state fit to the amplitude vectors corresponding to the first component with EXAM<sup>33</sup> obtained a  $T_m$  of about 310 K and 295 K for WH21<sup>0</sup> and WH21<sup>+</sup>, respectively. A helix fraction with respect to increasing temperature is presented in [Fig. 2\(b\)](#). Helix fractions at 263 K for WH21<sup>0</sup> and WH21<sup>+</sup> are 97% and 93%, respectively. A 50% helix content is found to be present at 312 K and 295 K for WH21<sup>0</sup> and WH21<sup>+</sup>, respectively. [Figure 2\(c\)](#) shows the standard free energy change of unfolding ( $\Delta G^\circ$ ) with respect to increasing temperature for WH21<sup>0</sup> and WH21<sup>+</sup>. In both protonation states of the peptide, the free energy change decreases with increasing temperature. The equilibrium constants as a function of increasing temperature in both protonation states are shown in [Fig. 2\(d\)](#). No significant deviation in the equilibrium



**FIG. 1.** Representative far UV circular dichroism (CD) spectra of an alpha-helix with two protonation states. (a) Wavelength dependent CD spectrum of the protonated form of the studied alpha-helix at 263 K. (b) Application of SVD on the wavelength dependent CD spectra between 263 K and 363 K produced classic features of an alpha-helix (orange) and coil (blue). Comp 1 and Comp 2 correspond to the two primary components obtained upon application of SVD on the measured set of CD spectra. (c) Wavelength dependent CD spectrum of the neutral form of the studied alpha-helix at 263 K. (d) SVD components, the alpha-helix (green) and coil (blue) of wavelength dependent CD spectra between 263 K and 363 K in the neutral form.



**FIG. 2.** Detailed thermodynamic analysis of the protonated and neutral form of the alpha-helical hetero peptide from the amplitude vectors of the first SVD component. (a) Amplitude vectors of the first component as a function of temperature for both protonated (orange) and neutral (green) forms of the alpha-helix. Circles are vector amplitudes and the solid-lines are the two-state fit to the data with EXAM.<sup>33</sup> (b) Obtained fraction helix content as a function of temperature of the protonated (orange) and neutral (green). (c) Change in the standard free energy of unfolding ( $\Delta G^\circ$ ) as a function of increasing temperature for protonated (orange) and neutral (green) forms. (d) Equilibrium constant as a function of increasing temperature of the protonated (orange) and neutral (green) forms. (e) Standard enthalpy ( $\Delta H^\circ$ ) of unfolding with the change in temperature of the protonated (orange) and neutral (green) forms. (f) Entropy ( $\Delta S^\circ$ ) changes of unfolding with respect to temperature for the neutral (green) and protonated (orange) forms.

constants of the two peptide forms is observed up to 310 K, while differences appear above that temperature. This suggests that above the midpoint temperature, there is a significantly larger population of the unfolded state in WH21<sup>+</sup> compared to WH21<sup>0</sup>. This is consistent with the overall observed thermodynamic properties of these two forms, with the WH21<sup>0</sup>  $\alpha$ -helix being more stable compared to WH21<sup>+</sup>. Standard enthalpy ( $\Delta H^\circ$ ) and entropy ( $\Delta S^\circ$ ) changes of unfolding with respect to temperature for WH21<sup>0</sup> and WH21<sup>+</sup> are presented in Figs. 2(e) and 2(f), respectively. The enthalpy and entropy change during unfolding of WH21<sup>0</sup> exhibited positive increment with temperature which corresponds to a positive heat capacity ( $\Delta C$ ) of unfolding. On the other hand, the enthalpy and entropy changes of unfolding for WH21<sup>+</sup> are decreasing as a function of increasing temperature, indicating a negative heat capacity change upon unfolding. These observations are consistent with the change in peptide properties due to protonation. A small positive heat capacity increment upon helix unfolding is observed to be associated with the solvation of non-polar groups.<sup>37</sup> WH21<sup>0</sup>, with a neutral histidine, is less polar than WH21<sup>+</sup>. In the case of WH21<sup>+</sup>, the negative  $\Delta C$  during helix unfolding is most likely due to solvation of the more polar protonated His. The heat capacity change appears to be consistent with our MD trajectory analysis, which indicates that the neutral form of the peptide has stronger van der Waals interactions with the solvent, which may be explained

by weaker Trp1...His5 interaction (see Table I and computational results below). This observation suggests that the hydrogen bonds are not the only contributing factor for alpha helical stability and van der Waals interactions play an important role in providing support for helix stabilization. More details on the thermodynamic treatment are provided in the [supplementary material](#).

Relaxation dynamics of WH21<sup>+</sup> with the laser t-jump in the nanosecond to microsecond time domain is measured immediately after initiating the temperature jump, via absorption of infrared energy by the water, to obtain information about

**TABLE I.** Average Lennard-Jones component of peptide-solvent interaction energy. Total, Trp and His—for the whole peptide, tryptophan and histidine, respectively, obtained over structures sampled every 100 ps (120 001 structures for WH21<sup>+</sup> and 130 001 for WH21<sup>0</sup>). Helix and coil results represent interactions of the whole peptide averaged over the subsets of structures with  $\geq 11$  hydrogen bonds and with  $< 11$  hydrogen bonds, respectively. Reported errors are 95% confidence intervals. Units: energy in kJ/mol.

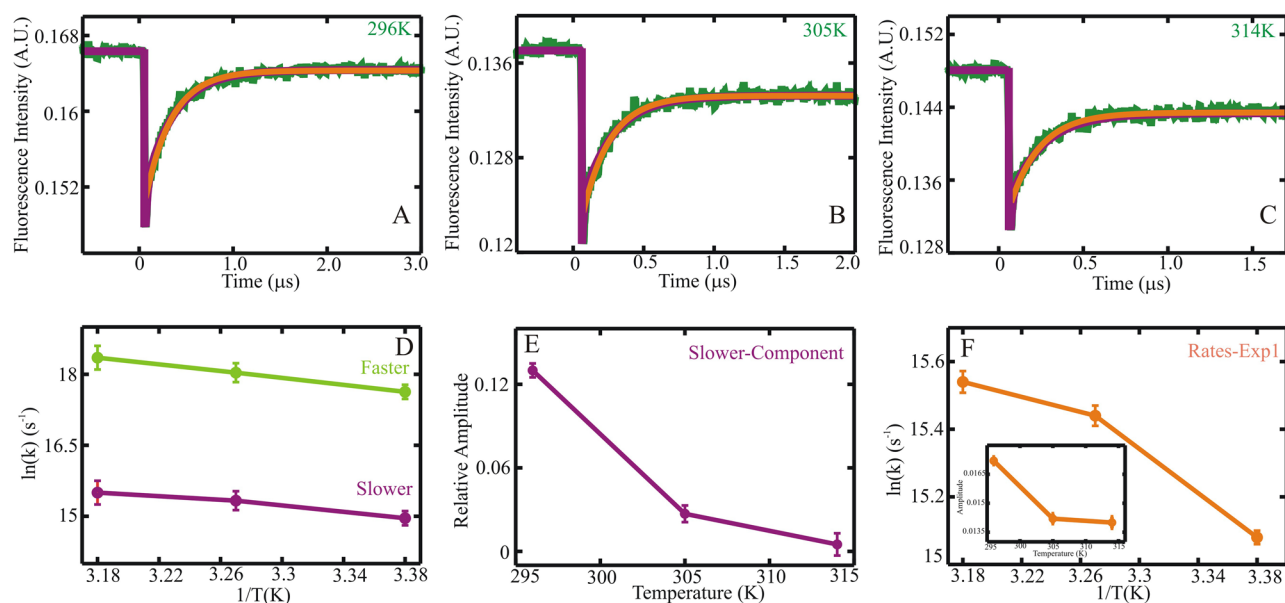
	WH21	WH21n
Total	$-426.0 \pm 0.1$	$-429.6 \pm 0.3$
Trp	$-53.2 \pm 0.1$	$-55.2 \pm 0.1$
His	$-18.0 \pm 0.1$	$-24.7 \pm 0.1$
Helix	$-399.5 \pm 0.7$	$-415.7 \pm 0.5$
Coil	$-430.2 \pm 0.4$	$-435.2 \pm 0.4$

the rate of unfolding–folding as well as the kinetic pathway for this alpha helical system. Kinetic traces at three different final temperatures 296 K, 305 K, and 314 K are shown in Figs. 3(a)–3(c). Kinetic traces at other temperatures are presented elsewhere.<sup>27</sup> As shown there, the kinetic traces follow a bi-exponential pattern of kinetics. In this work, the measured kinetic data with improved signal-to noise at several different temperatures contain a similar profile. The measured raw data are shown in forest green. Two types of fits, a single exponential in orange and a bi-exponential in purple, are shown with the measured kinetic trace. In these measurements, with *improved signal-to-noise ratio*, it is fairly evident that fits with a single exponential function, in orange, are unable to capture the entire kinetic processes involving this alpha helix. A bi-exponential fit, in purple, on the other hand, can fully describe the kinetics of this system. Results obtained based on the fits to the data with *improved signal-to-noise* are presented in Fig. 3(d). Figure 3(d) summarizes the results from a bi-exponential fit to the measured data at three different final temperatures where relaxation dynamics are measured. An initial kinetic component associated with the larger rate, faster time, is presented in green with associated errors. Presented in purple are the smaller rates, with errors corresponding to the slower kinetic component. A state with faster time scale along the relaxation pathway from an unfolded to a folded conformation is most likely associated with the formation of an intermediate. The observed slower time is associated with the formation of the global helical structure. Shown in Fig. 3(e), are the relative amplitudes of the overall structure

formation. In Fig. 3(f), are shown the relaxation rates obtained with a single exponential fit. Shown in the inset of 3(f) are the relative amplitudes. Variation in rates and amplitudes observed here with respect to increasing temperature are similar to the description with a bi-exponential fit. In spite of this similarity, the actual rate of overall structure formation in this alpha helical peptide described by a bi-exponential fit is found to be smaller than obtained with a single exponential fit. Therefore overall structure formation in this  $\alpha$ -helical peptide takes longer time than what is suggested by a single exponential fit. Fits with two exponential functions significantly improved  $\chi^2$  values, lowering them typically by 2 orders of magnitude relative to single exponential fits. A bi-exponential fit starkly points to the existence of an intermediate, also yielding the time scales and rates of formation associated with these rarely sampled states.

### Computational

We have reported simulations at 300 K with the CHARMM36<sup>29</sup> protein force field in TIP3P<sup>36</sup> water with ~25% helix content and 600–800 ns relaxation time for the neutral peptide, WH21<sup>0</sup>,<sup>28</sup> while ~20% helix content and a 360 ns relaxation time was estimated for WH21<sup>+</sup>.<sup>27</sup> The helix populations were in qualitative agreement with spectroscopic CD measurements of helix content –68% for the neutral and 44% for the protonated form at 300 K [Fig. 2(b)]. The 360 ns relaxation time of WH21<sup>+</sup> was also quite close to the value of 300 ns obtained from t-jump experiments at 296 K.<sup>27</sup> We have



**FIG. 3.** Nanosecond to microsecond relaxation dynamics measured with the laser t-jump of the protonated form of alpha-helix WH21<sup>+</sup>. The measured data at the three different final temperatures (a) 296 K, (b) 305 K, and (c) 314 K are in green. A single exponential fit to the data is shown in orange. A bi-exponential fit to the data is in purple. (d) Log of rates vs. 1/T plot. Results obtained from the biexponential fit. Faster rates are in green, and the slower rates are in purple. (e) Relative amplitudes of the slower rates as a function of temperatures. (f) Log of rates vs. 1/T plot from a single exponential fit to the measured kinetic data. Relative amplitudes of the rates associated with the single exponential fit are presented in the inset.

also performed a coarse-grained kinetic network analysis of dynamics of the protonated form, WH21<sup>+</sup>.<sup>27</sup> Here, we present the analogous kinetic analysis of WH21<sup>0</sup> as well as a comparison of the kinetic behavior of this helical peptide in its two different protonation states.

Structural clustering of MD trajectories was performed using the gromos algorithm in GROMACS,<sup>34</sup> using CA atom RMSD. With a cluster radius of  $r_{\text{cluster}} = 4.5 \text{ \AA}$ ,  $N_c = 194$  clusters were found for WH21<sup>0</sup> (compared to 199 clusters for WH21<sup>+</sup>).<sup>27</sup> The lifetime-based kinetic analysis was employed, counting transitions between cluster cores and cluster residence times,<sup>26</sup> generating a kinetic matrix **K** of order 194 for WH21<sup>0</sup>. Diagonalization of the kinetic matrix produced eigenvalues  $\lambda_i$  and time scales  $t_i = 1/\lambda_i$ . Calculations with a core radius of 2.51 Å for WH21<sup>0</sup> yielded models with the slowest rate in accord with the global kinetic time scales of 600–800 ns found in MD for WH21<sup>0</sup>. The slowest time scales are shown in Table II.

Besides the kinetic matrices, the lifetime-based analysis also yielded trajectory discretization, i.e., assignment of each trajectory frame to a cluster center (or microstate). These discretizations were used to generate transition matrices **T** with Emma 1.4 tools<sup>30</sup> and to perform kinetic coarse-graining using PCCA+.<sup>16</sup> PCCA+ uses the sign structure of the transition matrix **T** eigenvectors to find sets of microstates which form metastable meso-states, characterized by faster transitions within the sets and slower transitions between sets. Because **T** and **K** have the same eigenvectors, the meso-states from PCCA+ were the basis for calculating the long-term coarse-grained kinetic matrix **R**, using the optimal dimensionality reduction approach of Hummer and Szabo.<sup>26</sup>

A summary of the results of the optimal density reduction is presented below. Our approach reproduces the slowest kinetic lifetimes quite well, with maximum deviations of about 16%.

### Characterization of aggregate sets

The aggregate sets were characterized by state type, number of component fine-grained clusters, population, expressed as free energy, and lifetime. The state types were determined by analyzing RMSD from the helix and number of  $i \dots i + 4$  hydrogen bonds of the fine-grained clusters making up the aggregate sets. The helical state was assigned as

**TABLE II.** Ten slowest kinetic time scales obtained with RMSD clustering of MD trajectories with  $r_{\text{cluster}} = 4.5 \text{ \AA}$  and lifetime-based kinetic modeling. Results obtained here for WH21<sup>0</sup> are compared with previously calculated lifetimes for WH21<sup>+</sup>.<sup>27</sup> Units: ns.

WH21 <sup>+</sup> , $N_c = 199$ , $r_{\text{core}} = 2.6 \text{ \AA}$ <sup>27</sup>	WH21 <sup>0</sup> , $N_c = 194$ , $r_{\text{core}} = 2.51 \text{ \AA}$
75.4	122
78.1	132
87.1	137
135	152
153	231
356	287
395	600

**TABLE III.** Kinetic time scales for optimal dimensionality reduction calculation for WH21<sup>0</sup> with dimensions  $N = 2, 3, 4, 5$  compared to lifetimes from the full kinetic matrix (from Table II). Slowest lifetime italicized.

$N = 2$	$N = 3$	$N = 4$	$N = 5$	Full $N_c = 194$
...	...	...	137	152
...	...	205	219	231
...	264	265	265	287
507	510	538	537	600

the set containing the cluster with the central structure closest to the ideal helix. The coil was assigned as the set with the highest average RMSD from the helix and lowest average number of  $i \dots i + 4$  hydrogen bonds, with averages calculated over the component cluster centers. Any remaining states were classified as intermediates. Cluster populations were calculated as  $p_i = t_i/t_{\text{tot}}$ , with  $t_i$  being the trajectory residence time and  $t_{\text{tot}}$  the total trajectory length, and aggregate set populations  $P_I$  equal to the sum of component population  $p_i$  values. The relative free energies for aggregate sets were  $\Delta G^{\circ}_I = -RT \ln(P_I/P_{\text{MAX}})$ , and the lifetimes were obtained from diagonal elements of **R** as  $\tau_I = -1/R_{II}$ . The kinetic time scales obtained with the application of optimal dimensionality reduction are compared with the results of the full kinetic matrix in Table III. Rate matrices are given in the supplementary material.

### Dimensionality reduction, $N = 2$

The properties of the states are listed in Table IV.

The “helix” metastable set contains 39, and the “coil” set contains 156 of the 194 total clusters. In this case, the helical state is slightly more populated than the coil ( $\Delta G^{\circ} = +0.1 \text{ kcal/mol}$  for the coil). The lifetimes are 1100 ns for the helix and 920 ns for the coil. These are also the time scales of unfolding and folding, respectively. In this simple model, existence of a single folding pathway is found. The kinetic scheme and representative structures are presented in Fig. 4.

### Dimensionality reduction, $N = 3$

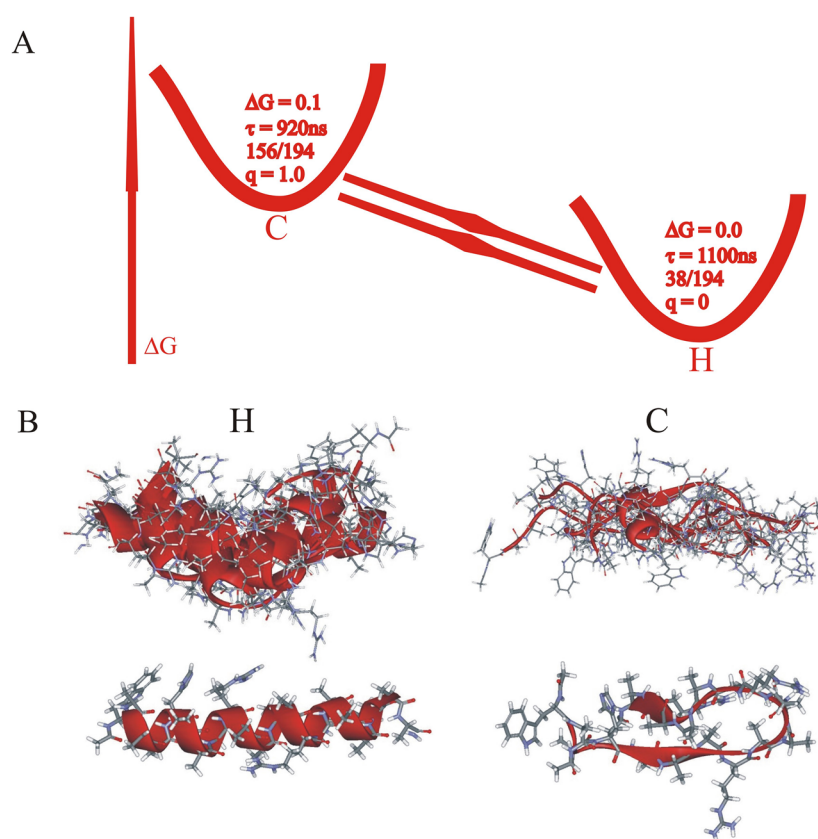
The properties of the states are listed in Table V.

The “helix” metastable set contains 34, the “coil” set contains 149, and the intermediate I contains 11 of the original clusters. The helical state is most highly populated, followed by the coil ( $\Delta G^{\circ} = +0.2 \text{ kcal/mol}$ ) and I ( $\Delta G^{\circ} = +1.7 \text{ kcal/mol}$ ). The lifetimes are 1100 ns for the helix, 760 ns for the coil, and

**TABLE IV.** Aggregate state properties for  $N = 2$ . State numbers correspond to the columns of **R**.

System	WH21 <sup>0</sup>	
State	1	2
Type	Helix	Coil
No. of clusters	38/194	156/194
$\Delta G^{\circ}$ (kcal/mol)	0.0	+0.1
Lifetime (ns)	1100	920
Committer, $q$	0.0	1.0





**FIG. 4.** Reduced dimensionality analysis for WH21<sup>0</sup> with  $N = 2$ . (a) The kinetic connectivity diagram based on the  $\mathbf{R}$  matrix. (b) Corresponding cluster center structures. The top row shows the complete ensemble of cluster centers belonging to each metastable state. The bottom row shows the central structure of the most populated cluster within each metastable state.

280 ns for I. The rates for coil  $\rightarrow$  helix and helix  $\rightarrow$  coil transitions are comparable. Rates for I  $\rightarrow$  helix and I  $\rightarrow$  coil transitions are larger than rates of the reverse processes. There are two folding paths: direct, coil  $\rightarrow$  helix and indirect, coil  $\rightarrow$  I  $\rightarrow$  helix, through the intermediate. The kinetic scheme and representative structures are presented in Fig. 5.

#### Dimensionality reduction, $N = 4$

The properties of the states are listed in Table VI.

The “helix” metastable set contains 21, the “coil” set contains 146, the intermediate I1 contains 11, and intermediate I2 contains 16 of the original clusters. The helical state is most populated, followed by the coil ( $\Delta G^\circ = +0.1$  kcal/mol), I1 ( $\Delta G^\circ = +0.9$  kcal/mol), and I2 ( $\Delta G^\circ = +1.6$  kcal/mol). The lifetimes are

**TABLE V.** Aggregate state properties for  $N = 3$ . State numbers correspond to the columns of  $\mathbf{R}$ .

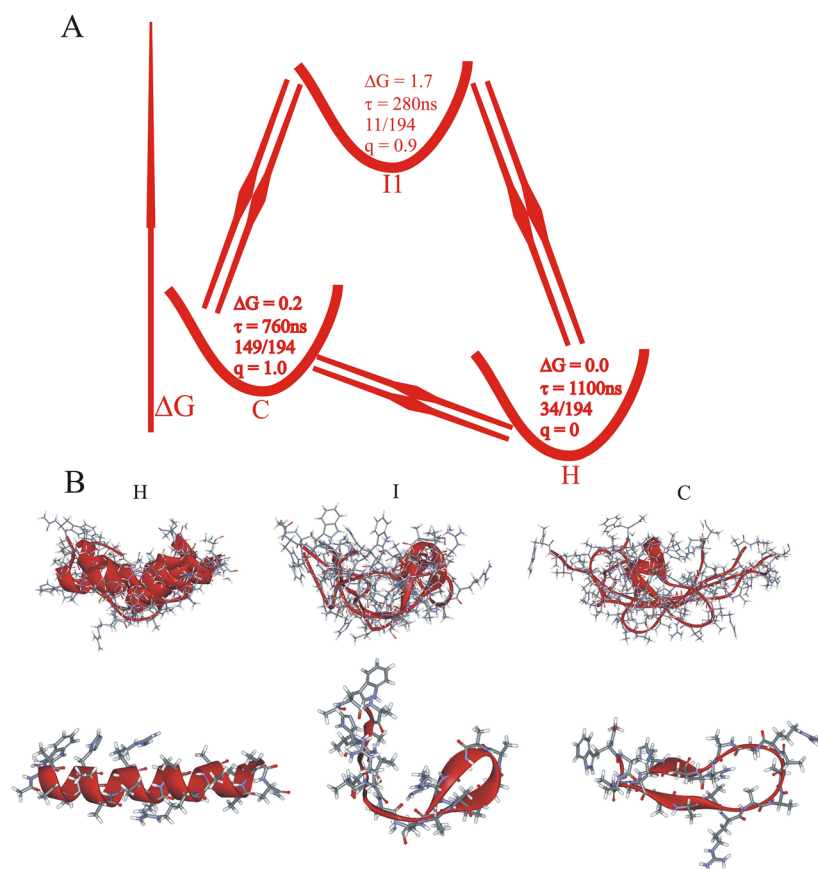
System	WH21 <sup>0</sup>		
State	1	2	3
Type	Helix	I	Coil
No. of clusters	34/194	11/194	149/194
$\Delta G^\circ$ (kcal/mol)	0.0	+1.7	+0.2
Lifetime (ns)	1100	280	760
Committer, $q$	0.0	0.9	1.0

820 ns for the helix, 720 ns for the coil, 230 ns for I1, and 280 ns for I2. The rates for coil  $\rightarrow$  helix and helix  $\rightarrow$  coil transitions are comparable. Rates for helix  $\rightarrow$  coil, I1  $\rightarrow$  helix, I1  $\rightarrow$  coil, and I2  $\rightarrow$  coil transitions are larger than rates of the reverse processes. There are five folding paths: one direct, coil  $\rightarrow$  helix and four indirect, coil  $\rightarrow$  I1  $\rightarrow$  helix, coil  $\rightarrow$  I2  $\rightarrow$  helix, coil  $\rightarrow$  I2  $\rightarrow$  I1  $\rightarrow$  helix, and coil  $\rightarrow$  I1  $\rightarrow$  I2  $\rightarrow$  helix. There are transitions between the two intermediates I1 and I2, but no significant transitions from the helix to I2. Intermediate I1, with a committer value of  $q = 0.5$ , corresponds to the helix-coil transition state. The kinetic scheme and representative structures are presented in Fig. 6.

#### Dimensionality reduction, $N = 5$

The properties of the states are listed in Table VII.

The “helix” metastable set contains 14, the “coil” set contains 146, the intermediate I1 contains 7, I2 contains 12, and I3 contains 15 of the original clusters. The helical state is most highly populated, followed by the coil ( $\Delta G^\circ = +0.1$  kcal/mol), I3 ( $\Delta G^\circ = +0.9$  kcal/mol), and I1 and I2 (both  $\Delta G^\circ = +1.6$  kcal/mol). The lifetimes are 670 ns for the helix, 720 ns for the coil, 150 ns for I1, 280 ns for I2, and 230 ns for I3. There are sixteen folding paths: one direct, three passing through a single intermediate, six passing through two intermediates, and six passing through all three intermediates. The kinetic network based on



**FIG. 5.** Reduced dimensionality analysis for WH21<sup>0</sup> with N = 3. (a) The kinetic connectivity diagram based on the **R** matrix. (b) Corresponding cluster center structures. The top row shows the complete ensemble of cluster centers belonging to each metastable state. The bottom row shows the central structure of the most populated cluster within each metastable state.

the **R** matrix for N = 5 in the WH21<sup>0</sup> peptide is presented in Fig. 7.

### Hydrogen bond fluctuations

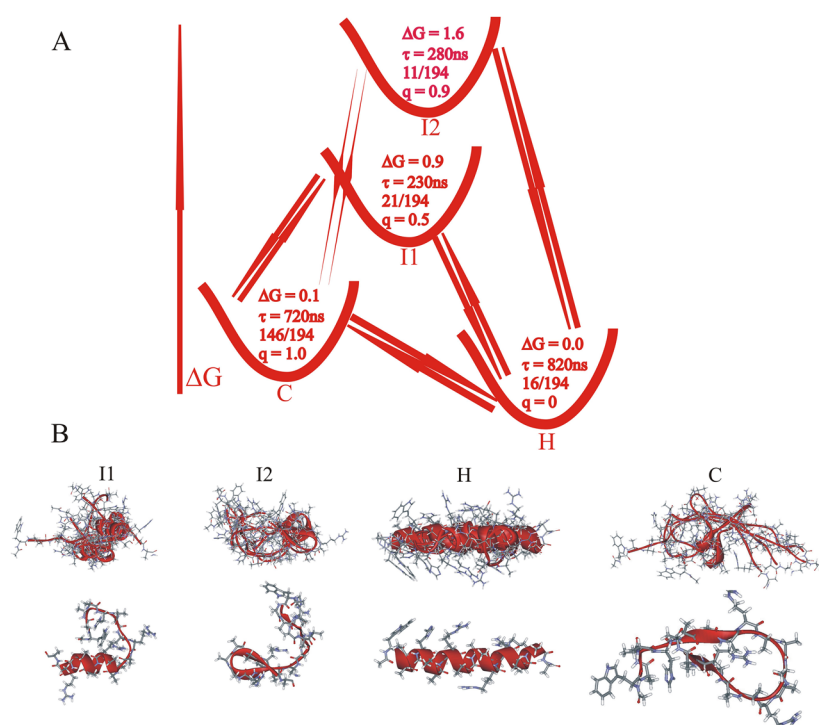
We have assigned the experimentally determined fast kinetic process in the protonated form, WH21<sup>0</sup>, with a relaxation time of 20 ns at 296 K, to correlated formation/breaking of several neighboring hydrogen bonds, as described elsewhere.<sup>28</sup> An analogous assignment may be proposed for WH21<sup>0</sup>, though its relaxation times have not yet been experimentally determined as deprotonation of the single histidine residue (His5) de-quenches the tryptophan (Trp1) fluorescence and makes the monitoring in the changes of

**TABLE VI.** Aggregate state properties for N = 4. State numbers correspond to the columns of **R**.

System	WH21 <sup>0</sup>			
State	1	2	3	4
Type	I1	I2	Helix	Coil
No. of clusters	21/194	11/194	16/194	146/194
$\Delta G^\circ$ (kcal/mol)	+0.9	+1.6	0.0	+0.1
Lifetime (ns)	230	280	820	720
Committer, $q$	0.5	0.9	0.0	1.0

tryptophan fluorescence intensity as a function of time much more challenging. The average autocorrelation function (ACF) for hydrogen bond fluctuations in WH21<sup>0</sup> is shown in Fig. 8. Fitting this average ACF to a three-exponential function yielded the relaxation times of 1, 16, and 340 ns, with weights of 0.1, 0.2, and 0.7, respectively (Fig. 8). The first time reflects local structure fluctuations. The last is shorter than the global MD relaxation time of 600–800 ns, indicating that this is a signature of a partial or segmental folding event. The middle relaxation time of 16 ns is similar to the experimentally determined 25 ns faster component found at 296 K for WH21<sup>+</sup>.<sup>27</sup> Recent triplet-triplet energy transfer studies have determined the diffusion rate of the helix-coil boundary at  $D = 2.7 \times 10^7$  (amino acids)<sup>2</sup>/s.<sup>21</sup> This boundary will move by one residue over a time of  $t = 1$  (amino acid)<sup>2</sup>/2D = 18 ns. This time is in very good agreement with our experimental and computational results of hydrogen-bond fluctuations.

Significant correlations exist between fluctuations of neighboring hydrogen bond lengths in WH21<sup>0</sup>. Average correlation coefficients are 0.84, 0.68, 0.57, and 0.52 for hydrogen bonds separated by one, two, three, and four residues, respectively (Fig. 9). Thus, it appears that helical hydrogen bond breaking and formation tends to occur in groups of 1–4 nearest neighbors, as has previously been proposed using simulations with OPLSA/AA and AMBER03 force fields,<sup>22</sup> implying



**FIG. 6.** Reduced dimensionality analysis for WH21<sup>0</sup> with  $N = 4$ . (a) The kinetic connectivity diagram based on the  $\mathbf{R}$  matrix. (b) Corresponding cluster center structures. The top row shows the complete ensemble of cluster centers belonging to each metastable state. The bottom row shows the central structure of the most populated cluster within each metastable state.

a partial cooperativity of folding. Thus, we suggest that a cooperative breaking/formation of small sections of the helix may be responsible for the observed faster component. We further suggest that these processes should be present in the helical, coil, and intermediates states of the peptide which may explain the observed relaxation time in the faster time scale.

### Peptide dynamics and folding pathways

Measured with the principal component analysis and secondary structure, the neutral peptide tends to sample the same conformational space as the protonated one, but with different weights, as described elsewhere.<sup>28</sup> The simulations reproduce the observed effect of increased helix stability in WH21<sup>0</sup> relative to WH21<sup>+</sup>. The MD results indicate that the increased helix content is local, primarily involving the central residues 6-18. This effect was correlated with a larger Trp $\cdots$ His separation and closer Arg $\cdots$ His and Arg $\cdots$ Trp interactions in the neutral form.<sup>28</sup> The Trp1 $\cdots$ His5

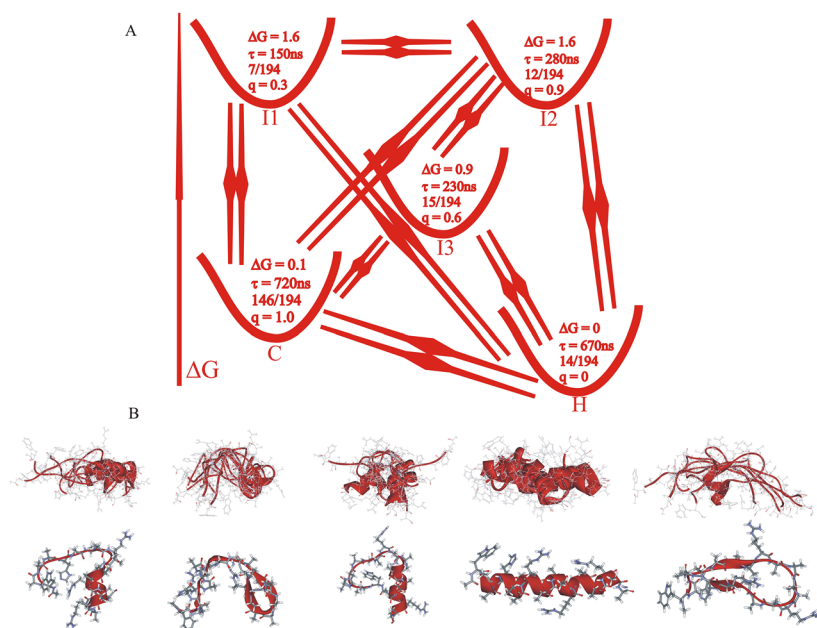
interaction is correlated with helix formation in both forms. The correlation coefficient between the fluctuations in Trp1 $\cdots$ His5 distance and the number of helical hydrogen bonds is  $-0.375$  in WH21<sup>+</sup> and  $-0.329$  in WH21<sup>0</sup>. For comparison, the correlation coefficient between the number of helical hydrogen bonds and CA RMSD from the helix is  $-0.626$  for WH21<sup>+</sup> and  $-0.829$  for WH21<sup>0</sup>. Detailed information about the correlations of different variables in the MD trajectories is presented in the [supplementary material](#).

The change in protonation state induces changes in the free energy landscapes, kinetic networks, and statistical folding paths of the peptide. Our MD trajectories of WH21<sup>0</sup> have a higher helix content and a slower global relaxation time than WH21<sup>0</sup>.<sup>27,28</sup> In the kinetic network models, the properties of the “helix” and “coil” are essentially switched between WH21<sup>0</sup> and WH21<sup>+</sup>. The biggest qualitative change is that the “helix” metastable set has the higher population and longer lifetime for WH21<sup>0</sup>, whereas in WH21<sup>+</sup> the “coil” is more stable.<sup>27</sup> In WH21<sup>+</sup>, the lifetimes of the “coil” are ca. 1000 ns and those of the “helix” are  $\sim 600$  ns for  $N = 2,3$ .<sup>27</sup> For the analogous models of WH21<sup>0</sup>, we find “helix” lifetimes of 1000–1100 ns and coil lifetimes of 800–900 ns. Interestingly, the populations and lifetimes of intermediates are comparable for the two protonation states.

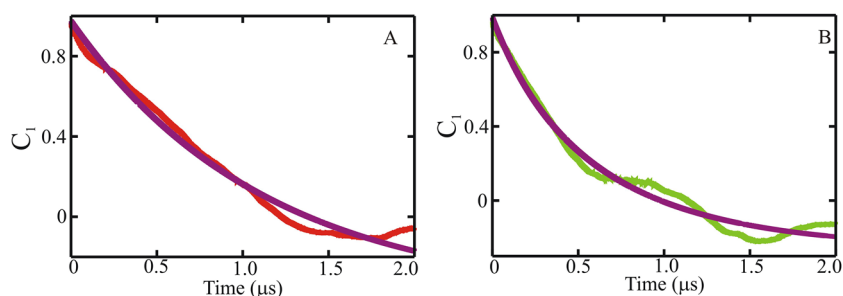
The statistical folding paths of the two peptide forms have been presented elsewhere.<sup>32,35</sup> Here we compare them with respect to each other in [Fig. 10](#), together with representative structures for coarse graining with  $N = 3$  metastable states. This figure shows the average populations of individual residues (NR = 1, 21, horizontal axis) in structures with different

**TABLE VII.** Aggregate state properties for  $N = 5$ . State numbers correspond to the columns of  $\mathbf{R}$ .

System	WH21 <sup>0</sup>				
State	1	2	3	4	5
Type	I1	I2	I3	Helix	Coil
No. of clusters	7/194	12/194	15/194	14/194	146/194
$\Delta G^\circ$ (kcal/mol)	+1.6	+1.6	0.9	+0.0	+0.1
Lifetime (ns)	150	280	230	670	720
Committer, $q$	0.3	0.9	0.6	0.0	1.0



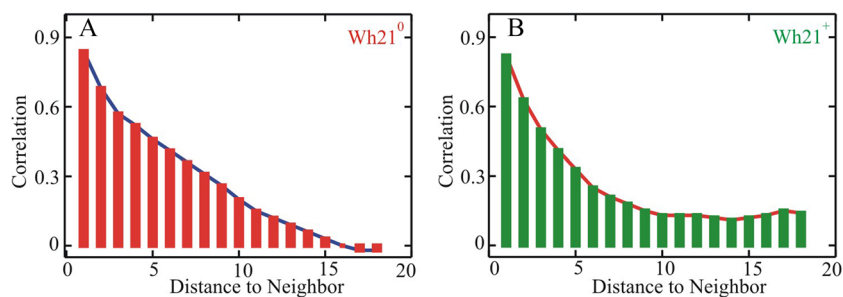
**FIG. 7.** Reduced dimensionality analysis for WH21<sup>0</sup> with  $N = 5$ . (a) The kinetic connectivity diagram based on the  $\mathbf{R}$  matrix. (b) Corresponding cluster center structures. The top row shows the complete ensemble of cluster centers belonging to each metastable state. The bottom row shows the central structure of the most populated cluster within each metastable state.



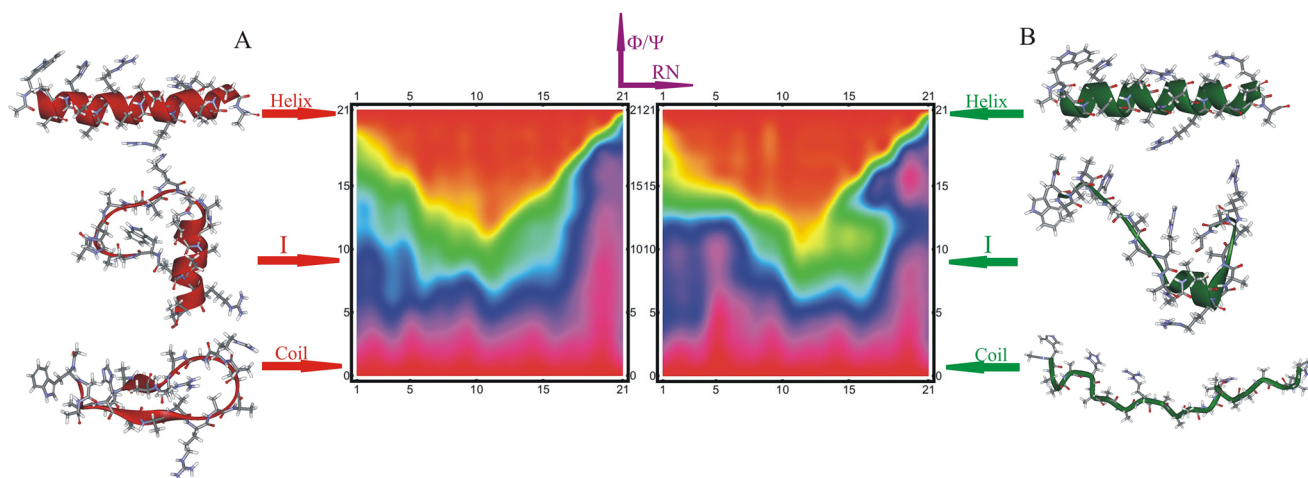
**FIG. 8.** Average autocorrelation functions for length fluctuations of the 19 helical hydrogen bonds ( $\text{C}=\text{O}_i \cdots \text{H}-\text{N}_{i+4}$ ). (a) WH21<sup>0</sup>, average autocorrelation in red, two-exponential fit in purple. (b) WH21<sup>+</sup>, average autocorrelation in green, two-exponential fit in purple.

numbers of residues in the helical region of the Ramachandran map ( $\text{NH} = 1, 21$ , vertical axis). The statistical paths differ in several aspects. In the neutral peptide [Fig. 10(a)], the nucleation probability ( $\text{NH} = 0-5$  region) is increased at the N-terminus, decreased at the C-terminus and uniform in the center.

For the protonated form [Fig. 10(b)], the pattern is similar except for a strongly decreased probability around the His5 residue. The regions of intermediates ( $\text{NH} = 8-12$ ) are markedly different. WH21<sup>+</sup> folding proceeds through an off-center intermediate with hydrogen bonds at residues  $\text{NR} = 11-16$  and



**FIG. 9.** Correlation coefficients between length fluctuations of hydrogen bonds as a function of their separation in sequence. Data averaged over hydrogen bonds with the starting residues  $i$  differing by 1, 2, 3, . . . , 18 residues. (a) Neutral form WH21<sup>0</sup> and (b) protonated form WH21<sup>+</sup>.



**FIG. 10.** The statistical path of helix folding. Data show the helix populations of individual residues (horizontal axis, NR = 1-21) averaged over samples of conformations with given number of residues in the helical region of the Ramachandran plot (vertical axis, NH = 0-21), i.e., within  $20^\circ$  radius of the ideal helix location taken from the final 500 ns of all 40 replicas of replica-exchange molecular dynamics.<sup>27,28</sup> The colors encode the average helix population of each peptide residue in the 0-1 range red: 0.0; green: 0.5; and purple: 1.0. The molecular structures are cluster centers for the N = 3 model. (a) Neutral, WH21<sup>0</sup><sup>28</sup> and (b) protonated, WH21<sup>+</sup>.<sup>27</sup>

favors N-terminal helices. However, for WH21<sup>0</sup> we find a propensity for helix formation in the peptide center, with both ends frayed.

### Peptide-solvent interactions

The experimentally observed trends in heat capacity change of unfolding have been explained due to changes in peptide polarity. To quantify this analysis, we have calculated the Lennard-Jones component of the peptide-solvent interaction in the MD trajectories. The data are presented in Table I. The energy data show that the neutral peptide is able to more strongly interact with the environment on average, and in both the helix and coil conformations. Thus, the detected weakening of the Trp1...His5 interaction in WH21<sup>0</sup><sup>28</sup> leads to improved solvation of the Trp1 and His5 residues and the peptide as a whole. The weakest solvent interaction is in the WH21<sup>+</sup> helical state, which has a higher correlation in the presence of Trp1...His5 contacts. For both forms of the peptide, the coil is more strongly solvated than the helix. Thus, the lower prevalence of the Trp1...His5 interaction in WH21<sup>0</sup> leads to stronger solute-solvent van der Waals interactions. This effect may be behind the opposite signs of  $\Delta C$  of unfolding for the two peptides.

### CONCLUSIONS

In-order to capture an atomically detailed picture of the pathway of progression from a coil to a helical state, we have performed a combined study with experiments and simulations for two protonation states of an  $\alpha$ -helical heteropeptide. Experimentally, we have applied far UV circular dichroism spectroscopy to characterize the thermodynamic properties

of the two conformers by examining the equilibrium constant, standard Gibbs free energy change, standard enthalpy change, and standard entropy change of unfolding with respect to increasing temperature. The measured midpoint temperature in the non-polar form is significantly higher than in the polar form. The unfolding enthalpy and entropy of these two systems exhibit opposite variations with increasing temperature—decreasing for WH21<sup>+</sup> and increasing for WH21<sup>0</sup>. Previous studies have related such an increase to the exposure of the hydrophobic surface area, which is consistent with WH21<sup>0</sup> being less polar than WH21<sup>+</sup>. This observation may be rationalized by the stronger solute-solvent interactions found in the WH21<sup>0</sup> simulations, explained by weakening of the Trp...His interactions in that form. The kinetic pathway of the peptide is examined with the application of a laser t-jump spectroscopy with significantly improved signal-to-noise from nanoseconds to microseconds, which not only indicates the presence of a longer time scale, associated with the formation of the global structure but also the presence of a faster time scale, related to very likely formation of an intermediate in the progression pathway, as suggested previously.<sup>27</sup>

In the previous study, we found that computational predictions on helix populations and kinetic rates were in qualitative agreement for this peptide with our experimental data.<sup>27,28</sup> Here, we have made use of the previously generated 13 microsecond MD trajectory of the neutral (deprotonated) peptide<sup>28</sup> to perform a coarse-grained kinetic analysis of that system and compare it with results for the protonated form. The deprotonation of the single histidine residue results in significant changes of the free energy landscape and kinetic network of the peptide. In the neutral form, WH21<sup>0</sup>, the “helix” metastable states are more highly populated and exhibit longer lifetimes than the “coil,” which is in contrast to the phenomenon observed in the WH21<sup>+</sup>.<sup>27</sup> However, the

“coil” states contain larger numbers of fine-grained clusters, corresponding to higher entropy, for both forms of the peptide. Also for two protonation states of the peptide, the “helix” metastable state consists of many clusters, indicating a kinetic inhomogeneity of the helical state. The intermediate states, on the other hand, contain a very few structures and have lowest populations with shortest lifetimes. As the number of considered metastable states increases, more intermediates and more folding paths appear in the coarse-grained models, clearly showing the presence of multiple folding paths and intermediates during helix folding. The statistical picture of folding is found to be markedly different in the neutral peptide compared to the protonated form. The weaker Trp $\cdots$ His interaction in WH21<sup>0</sup> changes the nucleation pattern and leads to a preference for a central region intermediate, rather than the “off-center” intermediate seen in the WH21<sup>+</sup>. An interesting question, which we will pursue in our future study, is the selection of the best reduced dimensionality N for a given system.

As in the case of WH21<sup>+</sup>, in WH21<sup>0</sup> we found a high correlation between fluctuations of neighboring hydrogen bonds, which suggests that the process of helix breaking and formation proceeds through an intermediate stage with correlated transitions in segments of several consecutive residues. Overall, our combined experimental-computational approach generated important fundamental data and provided an atomically detailed insight into the effects of protonation on the helical structure and folding pathway.

## SUPPLEMENTARY MATERIAL

See [supplementary material](#) for additional figures and tables describing the details of the thermodynamic analysis and the simulation.

## ACKNOWLEDGMENTS

G.S.J. would like to thank W. A. Eaton for allowing access to the nanosecond laser T-jump apparatus. We would like to acknowledge the support of XSEDE (Grant No. TG-MCB 16009) for computer time. G.S.J. would like to thank Professor Zilka Rios for encouraging comments and helpful conversation. Parts of the simulations described were conducted at the Center for Research Computing at the University of Kansas. This project was partly supported by an NSF (Grant No. 1807852).

## REFERENCES

- G. S. Jas and K. Kuczera, “Computer simulations of helix folding in homo- and heteropeptides,” *Mol. Simul.* **38**, 682–694 (2012).
- Y. Bai, J. S. Milne, L. Mayne, and S. W. Englander, “Primary structure effects on peptide group hydrogen exchange,” *Proteins: Struct., Funct., Genet.* **17**, 75–86 (1993).
- E. Y. Chi, S. Krishnan, T. W. Randolph, and J. F. Carpenter, “Physical stability of proteins in aqueous solution: Mechanism and driving forces nonnative protein aggregation,” *Pharm. Res.* **20**, 1325–1326 (2003).
- P. E. Fraser, J. T. Nguyen, W. K. Surewicz, and D. E. Kirschner, “pH-dependent structural transitions of Alzheimer amyloid peptides,” *Biophys. J.* **60**, 1190–1201 (1991).
- S. Y. Lau, A. K. Taneja, and R. S. Hodges, “Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils,” *J. Biol. Chem.* **259**, 13253–13261 (1984).
- F. J. Blanco, G. Rivas, and L. Serrano, “A short linear peptide that folds into a native stable  $\beta$ -hairpin in aqueous solution,” *Nat. Struct. Biol.* **1**, 584–590 (1994).
- P. E. Dawson, T. W. Muir, I. Clark-Lewis, and S. B. Kent, “Synthesis of proteins by native chemical ligation,” *Science* **266**, 776–779 (1994).
- C. N. Pace, S. Trevino, E. Prabhakaran, and J. M. Scholtz, “Protein structure, stability and solubility in water and other solvents,” *Philos. Trans. R. Soc., B* **359**, 1225–1235 (2004).
- P. A. Thompson, W. A. Eaton, and J. Hofrichter, “Laser temperature jump study of the helix reversible arrow coil kinetics of an alanine peptide interpreted with a ‘kinetic zipper’ model,” *Biochemistry* **36**, 9200–9210 (1997).
- P. A. Thompson, V. Munoz, G. S. Jas, E. R. Henry, W. A. Eaton, and J. Hofrichter, “The helix-coil kinetics of a heteropeptide,” *J. Phys. Chem. B* **104**, 378–389 (2000).
- G. S. Jas, W. A. Eaton, and J. Hofrichter, “Effect of viscosity on the kinetics of alpha-helix and beta-hairpin formation,” *J. Phys. Chem. B* **105**, 261–272 (2001).
- G. S. Jas, C. R. Middaugh, and K. Kuczera, “Non-exponential kinetics and a complete folding pathway of an  $\alpha$ -helical heteropeptide: Direct observation and comprehensive molecular dynamics,” *J. Phys. Chem. B* **118**, 639–647 (2014).
- O. F. Mohammed, G. S. Jas, M. M. Lin, and A. H. Zewail, “Primary peptide folding dynamics observed with ultrafast temperature jump,” *Angew. Chem., Int. Ed.* **48**, 5628–5632 (2009).
- M. M. Lin, O. F. Mohammed, G. S. Jas, and A. H. Zewail, “Speed limit of protein folding evidenced in secondary structure dynamics,” *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16622–16627 (2011).
- G. S. Jas, W. A. Hegefeld, P. Majek, K. Kuczera, and R. Elber, “Experiments and comprehensive simulations of the formation of a helical turn,” *J. Phys. Chem. B* **116**, 6598–6610 (2012).
- K. Kuczera, G. S. Jas, and R. Elber, “Kinetics of helix unfolding: Molecular dynamics simulations with milestone,” *J. Phys. Chem. A* **113**, 7461–7473 (2009).
- P. Peter Májek and R. Elber, “Milestone without a reaction coordinate,” *J. Chem. Theory Comput.* **6**, 1805–1817 (2010).
- J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, “Theory of protein folding: The energy landscape perspective,” *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
- C. M. Dobson, A. Sali, and M. Karplus, “Protein folding: A perspective from theory and experiment,” *Angew. Chem., Int. Ed.* **37**, 868–893 (1998).
- D. Thirumalai and D. K. Klimov, “Deciphering the timescales and mechanisms of protein folding using minimal off-lattice models,” *Curr. Opin. Struct. Biol.* **9**, 197–207 (1999).
- E. J. Sorin and V. S. Pande, “Exploring the helix-coil transition via all atom equilibrium ensemble simulations,” *Biophys. J.* **88**, 2472–2493 (2005).
- W. A. Hegefeld, S. E. Chen, K. Y. DeLeon, K. Kuczera, and G. S. Jas, “Helix formation in a pentapeptide: Experiment and force-field dependent dynamics,” *J. Phys. Chem. A* **114**, 12391–12402 (2010).
- J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope, “Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics,” *J. Chem. Phys.* **126**, 155101 (2007).
- S. Kube and M. Weber, “A coarse graining method for the identification of transition rates between molecular conformations,” *J. Chem. Phys.* **126**, 024103 (2007).
- E. K. Rains and H. C. Andersen, “A Bayesian method for construction of Markov models to describe dynamics on various time-scales,” *J. Chem. Phys.* **133**, 144113 (2010).

- <sup>26</sup>G. Hummer and A. Szabo, "Optimal dimensionality reduction of multi-state kinetic and Markov state models," *J. Phys. Chem. B* **119**, 9029–9037 (2015).
- <sup>27</sup>G. S. Jas and K. Kuczera, "Helix–coil transition courses through multiple pathways and intermediates: Fast kinetic measurements and dimensionality reduction," *J. Phys. Chem. B* **122**, 10806 (2018).
- <sup>28</sup>G. S. Jas and K. Kuczera, "Deprotonation of a single amino acid residue induces significant stability in an  $\alpha$ -helical heteropeptide," *J. Phys. Chem. B* **122**, 11508 (2018).
- <sup>29</sup>B. R. Brooks, C. L. Brooks, A. D. MacKerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch *et al.*, "CHARMM: The biomolecular simulation program," *J. Comput. Chem.* **30**, 1545–1614 (2009).
- <sup>30</sup>N. V. Buchete and G. Hummer, "Coarse master equations for peptide folding dynamics," *J. Phys. Chem. B* **112**, 6057–6069 (2008).
- <sup>31</sup>M. Senne, B. Trendelkamp-Schroer, A. S. J. S. Mey, C. Schütte, and F. Noé, "EMMA: A software package for Markov model building and analysis," *J. Chem. Theory Comput.* **8**, 2223–2238 (2012).
- <sup>32</sup>B. Trendelkamp-Schroer, H. Wu, F. Paul, and F. Noé, "Estimation and uncertainty of reversible Markov models," *J. Chem. Phys.* **143**, 174101 (2015).
- <sup>33</sup>W. H. Kirchhoff, Exam (Codex: Ntnoef), U.S. Government Printing Office, Washington, DC, 1993, Vol. 1401.
- <sup>34</sup>B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for highly efficient, load-balanced and scalable molecular simulation," *J. Chem. Theory Comput.* **4**, 435–447 (2008).
- <sup>35</sup>A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha *et al.*, "All-atom empirical potential for molecular modeling and dynamics studies of proteins," *J. Phys. Chem. B* **102**, 3586–3616 (1998).
- <sup>36</sup>W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.* **79**, 926–935 (1983).
- <sup>37</sup>J. Taylor, N. J. Greenfield, B. Wu, and P. L. Privalov, "A calorimetric study of the folding-unfolding of an  $\alpha$ -helix with covalently closed N and C-terminal loops," *J. Mol. Biol.* **291**, 965–976 (1999).