Kinetic Pathway Analysis of an α-Helix in Two Protonation States: Direct Observation and Optimal Dimensionality Reduction.

Gouri S. Jas*[1], Ed W. Childs[2], Krzysztof Kuczera[3,4]

[1]Department of Pharmaceutical Chemistry, The University of Kansas, Lawrence, KS 66047

[2]Department of Surgery, Morehouse School of Medicine, Atlanta, GA, USA

[3]Department of Molecular Biosciences, The University of Kansas, Lawrence, KS 66047

[4]Department of Chemistry, The University of Kansas, Lawrence, KS 66045

Abstract. Thermodynamically stable conformers of secondary structural elements make a stable tertiary/quartenary structure that performs its proper biological function efficiently. Formation mechanism 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 from 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 than in the protonated state. Nanosecond laser temperature jump spectroscopy monitoring time-resolved tryptophan fluorescence on the protonated conformer is carried out to measure kinetics on this system. Measured relaxation rates at a final temperature between 296K - 314K generated a faster component of 11ns - 20ns and a slower component of 314ns – 198ns. All-atom molecular dynamics trajectories of these two conformers are generated to perform an atomically detailed characterization of the helix-coil kinetics pathways. Application of clustering and coarse-graining with optimum dimensionality reduction allowed 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 "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~~. The intermediate states, on the other hand, have the shortest lifetimes and lowest populations.~~ 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.

Introduction.

In order to carry out a proper biological function a protein must exist in a robust form with respect to its physiological environment. 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. Thermodynamics play 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 fast time scale, such as in several nanoseconds, to a much longer time domain. Relaxation dynamics, breaking and formation of the secondary structural elements occur in a faster time scale compared to dynamics of a protein as a whole. Improvements of instrumentation with faster time resolution have been able to monitor formation events that take place in a shorter time scale[1-3]. Laser temperature jump spectroscopy, with picosecond to nanosecond time resolution, enabled us to measure relaxation dynamics of a helix with a sequence length from five residue to a twenty-one residues[4-6]. These measurements allow us the determine the time scale associated with the breaking and formation of the global structure as well as the nature of the pathway from the folded to unfolded state. A measured kinetic trace fitted with a single exponential function will suggest that the progression follows a two-state pathway, i.e., unfolded state of structure going to a folded state without visiting other state along the pathway. 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 unfolded to folded state is achieved is yet to be fully characterized. In principle, 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, intermediates, and the rates.

Recently, attempts have been made to describe helix folding computationally, using molecular dynamics simulations[7-13]. It was suggested that helix folding progresses from multiple nucleation sites to fragmented helices[8,13]. The Milestoning approach was undertaken to describe helix folding rates and pathway from a pre-determined path[8]. Ultrafast laser temperature jump spectroscopy was

applied to measure the relaxation dynamics of a single alpha helical turn formation[4]. Five-residue homopolymers with alanine showed the formation of stable helix in solution at physiologically relevant pH[14]. 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 kinetics 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[15-17] and milestoning[9]. Recently developed optimal dimensionality reduction by Szabo and Hummer[18] 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 model based on molecular dynamics simulation. Interesting approaches to kinetic coarse graining, with the PCCA+ [16] are promising and opening important testing grounds.

To conduct an atomically detailed pathway analysis of helix-coil kinetics, in this study we combined experimental measurements with long-term all-atom molecular dynamics simulations analyzed with optimal dimensionality reduction. In the experimental part, we applied far-UV circular dichroism (CD) spectroscopy and a nanosecond laser temperature jump (t-jump) on a twenty-one residue α-helical heteropeptide. Thermodynamic characterization of the two distinct peptide forms, with a protonated and a neutral single histidine residue, is carried out with the CD spectra measured between 263K – 363K. Fraction helix content, standard free energy change, change of enthalpy, and entropy as a function of temperature are analyzed from these measurements. Unfolding and folding dynamics of this alpha helical system are measured in the nanosecond to microsecond time scale with the application of t-jump. Kinetic measurements are conducted at a final temperature between 296K – 314K. Measured t-jump kinetic trace analysis obtained two different time constants for each trace. An observed faster component lies between 11ns -22ns. This component is much slower than the instrument time resolution of 6ns. This observation strongly directs us to the presence of a non-two state processes in the relaxation pathway. The measured slower component is between 317ns – 198ns. The slower component obtained from the analysis of the kinetic measurement is associated with the global structure formation.

In the computational component of this study, a 13 μs length of molecular dynamics (MD) trajectory is generated in explicit solvent. A novel approach with simplified coarse-grained models of low dimensionality based on the kinetic matrix is applied on the MD trajectory to obtain microscopic insights into the pathway of helix folding. The conformations sampled from the MD trajectory analysis were clustered into a set of microstates, for which a lifetime-based kinetic analysis was performed[20]. Using the PCCA+ algorithm[16], these microstates were then aggregated into a small number of metastable states which were the basis for generating an optimal reduced dimensionality rate matrix[18] and transition path theory analysis[21-22]. Models of dimensionalities between 2-5 provide new insights into the response of the folding pathway of the alpha helical heteropeptide with respect to the change in the protonation state. Our modeling results show that both the helix and coil states are heterogenous 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 increase resolution of the model. Finally, 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.

Methods.

Experimental.

Materials. Twenty-one residue alpha helical heteropeptide, with i and i+4 interactions, was purchased from Gen Script USA Inc., NJ. The purity was greater than 98%., determined with 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 detail of this heteropeptide is described elsewhere[2,3].

Circular Dichroism. Peptide secondary structural elements were characterized with far-UV circular dichroism spectroscopy as a function of temperature with the protonated and deprotonated form with a JASCO 815 spectropolarimeter (Tokyo, Japan) using a cylindrical cell with a pathlength of 0.5 mm and concentrations of ~220 μM. Scans between 263-363 K were recorded

in a 10° increments. Thermodynamic parameters and fraction helix as a function of temperature were obtained from the amplitude vector of SVD component one with a two-state equilibrium model EXAM[23]. The two-state fit obtained a temperature-independent parameter of the transition enthalpy ΔH° and entropy ΔS°, and the mid-point temperature $T_m = \Delta H°/\Delta S°$. The Gibbs free energy was calculated as $\Delta G° = \Delta H° - T\Delta S°$, and the equilibrium constant K was found through $\Delta G° = -RT\ ln(K)$. With a conservative estimate, the errors in these thermodynamic parameters are estimated to be ~5%, similar to previous analyses.[2,3] Singular value decomposition (SVD) was applied to the 170− 260nm CD scans to resolve overlapping spectral components. In this study we have made use of the first and second components of the wave length dependent circular dichroism spectra, from multiple components shown elsewhere[24], 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.

Temperature Jump Kinetics. Instrumentation of the laser temperature jump apparatus used in our kinetic measurements has been described in detail elsewhere[2,3]. The initial rapid decrease in fluorescence intensity results from a decrease in the intrinsic tryptophan fluorescence due to the rise in temperature from the laser temperature jump. Temperature jumps of 8−10 K were generated by absorption of a near-infrared laser pulse (1.54 μm) by the aqueous solution. Time resolved tryptophan fluorescence intensity was monitored to measure the kinetics. Tryptophan was excited using a continuous ultraviolet probe beam (264 nm), and the emitted light was monitored using band-pass filters which transmitted wavelengths from 320 to 400 nm. The kinetic experiments were carried out in a 0.05 × 1.0 cm^2 cuvette at a sample concentration of ~100 μM in 20 mM acetate buffer in pH 4.8. Measured kinetics from the change in fluorescence of the alpha helical heteropeptide were recorded over the temperature range 296−314 K with much improved *signal-to-noise* ratio. This effort is made to capture the faster component of this kinetics process more clearly as this is a critical signal associated with the short-lived structure formation. Measured similar relaxation dynamics and the fits to the data are shown elsewhere[24]. In this this study we have made use of the relative amplitude of the fast component from a biexponential function 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 with the decrease amplitudes in kinetic measurement.

Computational.

The simulations used in this work have been described, in detail, elsewhere[24,27]. Our kinetic analysis of the neutral peptide, Ac-WA_3H-$(AAARA)_3$-A-NH_2 ($WH21^0$), is based on a 13 μs MD trajectory generated with GROMACS[25] with the CHARMM 36[19,31] force field and TIP3P water[24,30] at 300 K. For comparison, results from the 12 μs MD simulation of the protonated peptide were also used[27]. Technical details are presented in the Supplementary Information.

Hydrogen bonds were counted as they formed when the O…N distance was below 3.6 Å and a residue was considered to be in the α-helical conformation when the (φ,ψ) values were found to be within a 20° radius of an ideal helical structure (-62°,-41°). Structure clustering was done with the gromos algorithm[25]. Statistical error estimates were obtained by dividing data into ten consecutive blocks and calculating the standard error of the mean at 95% confidence level.

In the analysis of peptide kinetics, two approaches were employed. First, time autocorrelation functions $C_1(t) = \langle (x(0)-\langle x\rangle)(x(t)-\langle x\rangle)\rangle$ were calculated, and x being the number of α–helical hydrogen bonds and number of residues in the α-helical region of the Ramachandran map. These functions were fitted to two-exponential decays and the longer time scale τ_2 was used to estimate the global relaxation time of the peptide. In the second approach, we discretized the conformation space explored in the MD into a set of microstates by Cα atom RMSD and performed a lifetime-based kinetic analysis[26]. The method for obtaining transition counts between microstates, elements of the kinetic matrix **K** and transition matrix **T** are described in more detail elsewhere[27] and the in the Supplementary Information. 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 τ_2 based on the $C_1(t)$ functions agreed with λ_2 , the lowest non-zero eigenvalue of **K**, i.e. $\tau_2 \approx 1/\lambda_2$. For analysis of kinetics results we used tools from the EMMA software package – PCCA+ and TPT[28]. Finally, optimal reduced kinetic matrices were generated with the formalism of Hummer and Szabo[18], based on the coarse-graining sets obtained through PCCA+ [16].

Results and Discussion.

Experimental Section.

Presented in figure 1 are the representative Far UV circular dichroism (CD) spectra of two forms of the twenty-one residue α-helical heteropeptide ($WH21^{+}$/$WH21^{0}$), with protonated and neutral histidine. A complete set of CD spectra measured between 263K – 363K have been presented elsewhere[24]. In figure 1A, are shown the CD spectra of $WH21^{+}$, where histidine has the protonated form. Measured spectral features shown here, exhibited two minima at 222nm, 210nm and a maximum at 190nm. 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)[7] on the set of temperature dependent CD spectra, shown in figure 1B. Shown in figure 1C are temperature dependent CD spectra of $WH21^{0}$ where histidine is present in the neutral form. Figure 1D shows two SVD components of $WH21^{0}$. Detailed temperature dependent CD spectra are shown elsewhere[24]. A representative CD spectrum at 263K is shown here to describe the basis of our expanded comprehensive thermodynamic analysis, presented below.

In figure 2A are shown the amplitudes vectors of component one as a function of temperature for both $WH21^{+}$ and $WH21^{0}$. Compared to $WH21^{+}$, $WH21^{0}$ vectors are shifted to the right, showing an increased α-helix stability compared to $WH21^{+}$. A two-state fit to the amplitude vectors corresponding to the first component with EXAM[23] obtained a T_m of about 310K and 295K for $WH21^{0}$ and $WH21^{+}$, respectively. Helix fraction with respect to increasing temperature is presented in figure 2B. Helix fractions at 263K for His^{0} and His^{+} are 92% and 94%, respectively. A 50% helix content are found to be at 310 K and 295K for $WH21^{0}$ and $WH21^{+}$, respectively. Figure 2C shows the standard free energy change of unfolding (ΔG^{0}) with respect to increasing temperature for $WH21^{0}$ and $WH21^{+}$. It is found in both protonation states of the helix that the free energy change decreases with increasing temperature. The equilibrium constant as a function of increasing temperature in both protonation states is shown in figure 2D. No significant deviation in the equilibrium constant for the two forms is observed up to 310K. The equilibrium constants start to deviate from each other above 310K. This suggests that above midpoint temperature, there is a significantly larger population of unfolded state in the $WH21^{+}$ compared to $WH21^{0}$. This is consistent with the overall observed thermodynamic properties of these two conformers, with the

$WH21^0$ α-helix being more stable compared to $WH21^+$. Standard enthalpy (ΔH^0) and entropy (ΔS^0) changes of unfolding with respect to temperature for $WH21^0$ and $WH21^+$ are presented in figure 2E and 2F, respectively. The enthalpy and entropy change for $WH21^0$ unfolding exhibited positive increment with temperature, corresponding to a positive heat capacity of unfolding ΔC. On the other hand, the enthalpy and entropy change for $WH21^+$ unfolding are decreasing as a function of increasing temperature, indicating a negative heat capacity of unfolding ΔC. These observations present a very consistent picture with the change in peptide properties upon protonation. A small positive heat capacity increment upon helix unfolding is observed to be associated with the solvation of non-polar groups[29]. $WH21^0$, with a neutral histidine, is less polar than $WH21^+$, where His is protonated. In the case of $WH21^+$ the negative ΔC during helix unfolding is likely due to the solvation of the more polar protonated His. The heat capacity effects appear to be consistent with our MD trajectory analysis, which indicates that the neutral form of the peptide exhibits stronger van der Waals interactions with the solvent, due to breaking of the strong Trp1…His5 interaction (see Table 1). This also suggests that the hydrogen bonds are not the only contributing factor for alpha helix formation and van der Waals interactions are also providing support for helix stabilization. More details on the thermodynamic treatment are provided in the Supplementary Information.

Table 1. Average Lennard-Jones component of peptide-solvent interaction energy. Total, Trp and His – for whole peptide, tryptophan and histidine, respectively, obtained over structures sampled every 100 ps (120,001 structures for $WH21^+$, 130,001 for $WH21^0$). Helix and coil results represent interactions of whole peptide averaged over the subsets of structures with ≥ 11 hydrogen bonds, coil with < 11 hydrogen bonds, respectively. Reported errors are 95% confidence intervals. Units: energy in kJ/mol.

	WH21	WH21n
Total	-426.0 ± 0.1	-429.6 ± 0.3
Trp	-53.2 ± 0.1	-55.2 ± 0.1
His	-18.0 ± 0.1	-24.7 ± 0.1

Helix	-399.5 ± 0.7	-415.7 ± 0.5
Coil	-430.2 ± 0.4	-435.2 ± 0.4

Relaxation dynamics of $WH21^+$ with laser t-jump in the nanosecond to microsecond time domain is measured immediately after initiating the temperature jump via absorption of infrared energy in the water to obtain information about the rate of unfolding-folding event as well as the kinetic pathway for this alpha helical system. Kinetic traces at three different final temperatures 296K, 305K, and 314K are shown in figure 3A-3C. Kinetics traces at other temperatures are presented elsewhere[27]. As shown there, the kinetics traces follow a bi-exponential nature of kinetics. In this this work, measured kinetics data at several different temperatures contain a similar profile. 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 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, is 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 figure 3D. Figure 3D summarizes the results from a bi-exponential fit to the measured data at three different final temperatures where relaxation dynamics are measured. Initial kinetic component associated with the larger rate, faster time, is presented in green with associated errors. Presented in purple are the smaller rates, longer time, with errors corresponding to the slower kinetic component. A state with faster time scale along the relaxation pathway from unfolded to a folded conformation is likely to be associated with the intermediate formation. The observed slower time is associated with the formation of global structure. Shown in figure 3E, are the relative amplitudes of the overall structure formation. In figure 3F are shown the relaxation rates obtained with a single exponential fit. Shown in the inset of 3F are the relative amplitudes. Variation in the rates and amplitudes seen 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 α-helical peptide takes longer time than what is

suggested by a single exponential fit. Fits with two exponential function significantly improved χ^2 values over single exponentials, with χ^2 values typically 2-orders of magnitude lower. A bi-exponential fit starkly points to the existence of an intermediate with the time scale and rate of formation of these rarely sampled states.

Kinetic analysis

Previously, we have reported that simulations with the CHARMM36[19] protein force field in TIP3P[30] water predicted 22-23% helix content and 600-800 ns folding relaxation for the neutral peptide, $WH21^0$ [24], while 14-17% helix content and a 360 ns relaxation was found for $WH21^+$ [27]. We have also performed a coarse-grained kinetic network analysis of dynamics of the protonated form, $WH21^+$ [27]. Here, we present the analogous kinetic analysis of $WH21^0$, 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[25], using CA atom RMSD. With a cluster radius of $r_{cluster}$=4.5 Å, N_c=194 clusters were found for $WH21^0$ (compared to 199 clusters for $WH21^+$)[27]. The lifetime-based kinetic analysis was employed, counting transitions between cluster cores and cluster residence times[26], generating a kinetic matrix **K** of order 194 for WH21n. Diagonalization of the kinetic matrix produced eigenvalues λ_i and timescales $t_i = 1/\lambda_i$. Calculations with core radius of 2.51Å for WH21n yielded models with the slowest rate in accord with the global kinetic time scales of 600-800 ns found in MD for WH21n. The slowest timescales are shown in Table 2.

Table 2. Ten slowest kinetic time scales obtained with RMSD clustering of MD trajectories with $r_{cluster}$=4.5 Å and lifetime-based kinetic modeling. Results obtained here for WH21 are compared with previously calculated lifetimes for WH21[27]. Units: ns.

WH21, N_c=199, r_{core} = 2.6 Å [27]	WH21n, N_c=194, r_{core} = 2.51 Å
75.4	122

78.1	132
87.1	137
135	152
153	231
356	287
395	600

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[28] and to perform kinetic coarse-graining using PCCA+[16]. PCCA+ uses the sign structure of the transition matrix **T** eigenvectors to find sets 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[18].

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 helix and number of i…i+4 hydrogen bonds of the fine-grained clusters making up the aggregate sets. The helix state was assigned as the set containing the cluster with central structure closest to ideal helix. The coil was assigned as the set with the highest average RMSD from helix and lowest average number of i…i+4 hydrogen bonds, with averages calculated over the component cluster centers. Any remaining states were classified as intermediates. Cluster populations were calculated as $p_i = t_i/t_{tot}$, with t_i the trajectory residence time and t_{tot} the total trajectory length, and aggregate set populations P_I equal to sum of component population p_i values.

The relative free energies for aggregate sets were $\Delta G^{o}_{I} = -RT \ln(P_I / P_{MAX})$ and the lifetimes were obtained from diagonal elements of **R**, as $\tau_I = -1/R_{II}$.

Table 3. Kinetic time scales for optimal dimensionality reduction calculation for $WH21^0$ with dimensions N=2,3,4,5 compared to lifetimes from full kinetic matrix (from Table 2). 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*

Dimensionality reduction, N=2

The reduced kinetic matrix **R** for N=2 is given in Table 4, while the properties of the states are listed in Table 5.

Table 4. Aggregate rate matrix **R** for N=2. Units: ns^{-1}. The elements R_{IJ} are the effective rates for J→I transitions, while the diagonal elements are related to state lifetimes, $\tau_I = -1/R_{II}$.

$WH21^0$	Nc=194
-0.000881	-0.001092
0.000881	0.001092

Table 5. Aggregate state properties for N=2. State numbers correspond to columns of **R.**

System	$WH21^0$	

State	1	2
Type	helix	coil
#of clusters	38/194	156/194
ΔG^o , kcal/mol	0.0	+0.1
Lifetime, ns	1,100	920
Committor, q	1.0	0.0

The "helix" metastable set contains 39 and the "coil" set 156 of the 194 total clusters. In this case, the helical state is slightly more populated than the coil (ΔG^o = +0.1 kcal/mol for coil). The lifetimes are 1,100 ns for helix and 920 ns for coil. These are also the timescales of unfolding and folding, respectively. In this simple model a single folding pathway is present. The kinetic scheme and representative structures are presented in Figure 4.

Dimensionality reduction, N=3

The reduced kinetic matrix **R** for N=3 is given in Table 6, while the properties of the states are listed in Table 7.

Table 6. Aggregate rate matrix **R** for N=3. Units: ns $^{-1}$. The elements R_{IJ} are the effective rates for J→I transitions, while the diagonal elements are related to state lifetimes, $\tau_I = -1/R_{II}$.

WH21^0	Nc=194	
-0.000896	0.000684	0.001107
0.000040	-0.003535	0.000217
0.000856	0.002850	-0.001324

Table 7. Aggregate state properties for N=3. State numbers correspond to columns of **R.**

System	$WH21^0$		
State	1	2	3
Type	helix	I	coil
#of clusters	34/194	11/194	149/194
ΔG^o (kcal/mol)	0.0	+1.7	+0.2
Lifetime, ns	1,100	280	760
Commitor, q'	0.0	0.1	0.0

A very slow rate, R_{12} =0.000040 ns^{-1} is found here. Rates below 0.000100 ns^- correspond to time scales above 10,000 ns, comparable to the total trajectory lengths. Such slow rates cannot be reliably estimated from our data. In these cases we can only qualitatively state that the rate for the corresponding process (I→ helix) is slower.

The "helix" metastable set contains 34, the "coil" set 149 and the intermediate I 11 of the original clusters. The helix state is most highly populated, followed by the coil (ΔG^o = +0.2 kcal/mol) and I (ΔG^o = +1.7 kcal/mol). The lifetimes are 1,100 ns for helix, 760 ns for coil and 280 ns for I. The rates for coil→helix and helix→coil transitions are comparable. Rates for I→helix and I→coil transitions are larger than rates of the reverse processes. There are two folding paths – direct, coil→helix and indirect, coil→I→helix, through the intermediate. The kinetic scheme and representative structures are presented in Figure 5.

Dimensionality reduction, N=4

The reduced kinetic matrix **R** for N=4 is given in Table 8, and the properties of the states are listed in Table 9.

Table 8. Aggregate rate matrix **R** for N=4. Units: ns $^{-1}$. The elements R_{IJ} are the effective rates for J→I transitions, while the diagonal elements are related to state lifetimes, $\tau_I = -1/R_{II}$.

WH21^0	Nc=194		
-0.004337	0.000778	0.000526	0.000465
0.000239	-0.003564	0.000009	0.000205
0.002239	0.000118	-0.001212	0.000722
0.001859	0.002668	0.000678	-0.001393

Table 9. Aggregate state properties for N=4. State numbers correspond to columns of **R.**

System	WH21^0			
State	1	2	3	4
Type	I1	I2	helix	coil
#of clusters	21/194	11/194	16/194	146/194
ΔG° kcal/mol	+0.9	+1.6	0.0	+0.1
Lifetime, ns	230	280	820	720
Commitor, q’	0.5	0.1	1.0	0.0

Here, there is one element of R with a very low value, R_{23} =0.000009 ns^{-1}. Such slow rates cannot be reliably estimated from our data, so we can only state that based on our calculations the rate of this process (helix→I2) is slow compared to other processes studied.

The "helix" metastable set contains 21, the "coil" set 146, the intermediate I1 11 and intermediate I2 16 of the original clusters. The helical state is most populated, followed by the coil (ΔG^o = +0.1 kcal/mol), I1 (ΔG^o = +0.9 kcal/mol) and I2 (ΔG^o = +1.6 kcal/mol). The lifetimes are 820 ns for helix, 720 ns for coil, 230 ns for I1 and 280 ns for I2. The rates for coil→helix and helix→coil transitions are comparable. Rates for helix→coil, I1→helix, I1→coil and I2→coil transitions are larger than rates of the reverse processes. There are five folding paths – one direct, coil→helix and four indirect, coil→I1→helix, coil→I2→helix, coil→I2→I1→helix and coil→I1→I2→helix. There are transitions between the two intermediates I1 and I2, but no significant transitions from helix to I2. Intermediate I1, with a commitor value of q = 0.5, corresponds to the helix-coil transition state. The kinetic scheme and representative structures are presented in Figure 6.

Dimensionality reduction, N=5

The reduced kinetic matrix **R** for N=5 are given in Table 9, while the properties of the states are listed in Table 10.

Table 9. Aggregate rate matrix **R** for N=5. Units: ns $^{-1}$. The elements R_{IJ} are the effective rates for J→I transitions, while the diagonal elements are related to state lifetimes, $\tau_I = -1/R_{II}$.

WH21^0	Nc=194			
-0.006613	0.000020	0.000839	0.000456	0.000116
0.000013	-0.003570	0.000251	0.000011	0.000209
0.001639	0.000716	-0.004429	0.000351	0.000395
0.003938	0.000144	0.001555	-0.001495	0.000664
0.001023	0.002690	0.001784	0.000677	-0.001384

Table 10. Aggregate state properties for N=5. State numbers correspond to columns of **R.**

System	$WH21^0$				
State	1	2	3	4	5
Type	I1	I2	I3	helix	coil
#of clusters	7/194	12/194	15/194	14/194	146/194
ΔG^o, kcal/mol	+1.6	+1.6	0.9	+0.0	+0.1
Lifetime, ns	150	280	230	670	720
Commitor, q'	0.7	0.1	0.4	1.0	0.0

There are three elements of **R** with very low values, R_{12} =0.000020 ns^{-1}, R_{21} =0.000013 ns^{-1} and R_{24}=0.000011 ns^{-1}. Such slow rates cannot be reliably estimated from our data, so we can only state that based on our calculations the rate of these processes are slow.

The "helix" metastable set contains 14, the "coil" set 146, the intermediate I1 7, I2 12 and I3 15 of the original clusters. The helix state is most highly populated, followed by the coil (ΔG^o = +0.1 kcal/mol), I3 (ΔG^o = +0.9 kcal/mol) and I1 and I2 (both ΔG^o = +1.6 kcal/mol)). The lifetimes are 670 ns for helix. 720 ns for coil, 150 ns for I1 and 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 **R** matrix for N=5 in the $WH21^0$ peptide is presented in Figure 7.

Hydrogen bond fluctuations.

We have assigned the experimentally determined fast kinetic process in the protonated form, $WH21^0$, with a relaxation time of 20 ns at 296 K, to correlated formation/breaking of several

neighboring hydrogen bonds, as described elsewhere[24]. An analogous assignment may be proposed for $WH21^0$, though its relaxation times have not yet been experimentally assigned as deprotonation of single histidine residue (His5) de-quenches the tryptophan (Trp1) fluorescence and makes the monitoring in the changes of tryptophan fluorescence intensity as a function of time much more challenging. The average autocorrelation function (ACF) for hydrogen bond fluctuations in $WH21^0$ is shown in Figure 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 (Figure 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 very similar to the faster components of the experimentally measured time scale of 300 ns for $WH21^+$ [27]. Recent triplet-triplet energy transfer studies have determined the diffusion rate of the helix-coil boundary at $D = 2.7\times10^7$ (amino acids)^2/s[13]. This boundary will move by one residue over a time of $t = 1\ (amino\ acid)^2/2D = 18$ 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 $WH21^0$. 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 (Figure 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[14], implying a partial cooperativity of folding. *Thus, we propose that the cooperative breaking/formation of small sections of the helix are responsible for the observed faster signal.* Such a process would be present in all peptide states – helix, coil and intermediates, explaining the observation of this specific relaxation timescale.

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[24]. The simulations reproduce the observed effect of increased helix stability in $WH21^0$ relative to $WH21^+$. 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…His

separation and closer Arg…His and Arg…Trp interactions in the neutral form[24]. The Trp1…His5 interaction is correlated with helix formation in both forms. The correlation coefficient between the fluctuations in Trp1…His5 distance and the number of helical hydrogen bonds is -0.375 in $WH21^+$ and -0.329 in $WH21^0$. For comparison, the correlation coefficients between the number of helical hydrogen bonds and CA RMSD from helix is -0.626 for $WH21^+$ and -0.829 for $WH21^0$ Detailed information about the correlations of different variables in the MD trajectories is presented in the Supplementary Information.

The change in protonation state induces changes in the free energy landscapes, kinetic networks and statistical folding paths of the peptide. In our MD trajectories of $WH21^0$ has higher helix content and slower global relaxation time than $WH21^0$ [24,27]. In the kinetic network models, the properties of the "helix" and "coil" are essentially switched between $WH21^0$ and $WH21^+$. The biggest qualitative change is that "helix" metastable set has the higher population and longer lifetime for $WH21^0$, whereas in $WH21^+$ the "coil" is more stable[27]. In $WH21^+$ the lifetimes of "coil" are ca. 1,000 ns and those of "helix" are ca. 600 ns for N=2,3 [27]. For the analogous models of $WH21^0$, we find "helix" lifetimes of 1,000-1,100 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 previously [24,27]. Here we compare them side-by-side in Figure 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 numbers of residues in the helical region of the Ramachandran map (NH=1,21, vertical axis). The statistical paths differ in several aspects. In the neutral peptide (Figure 10A) the nucleation probability (NH=0-5 region) is increased at the N-terminus, decreased at the C-terminus and uniform in the center. For the protonated form (Figure 10B) the pattern is similar except for a strongly decreased probability around the His5 residue. The regions of intermediates (NH=8-12) are markedly different. $WH21^+$ folding proceeds through an off-center intermediate with hydrogen bonds at residues NR=11-16 and favors N-terminal helices. However, for $WH21^0$ 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 as 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 1. 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 $WH21^0$ [24], leads to improved solvation the Trp1 and His5 residues and the peptide as a whole. The weakest solvent interaction is in the $WH21^+$ helix state, which has a higher correlation with 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 $WH21^0$ leads to stronger solute-solvent van der Waals interactions. This effect may be behind the opposite signs of ΔC 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 α-helical heteropeptide. Experimentally, we have applied far UV circular dichroism spectroscopy to characterize the thermodynamic properties of the two forms 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 $WH21^+$ and increasing for $WH21^0$. Previous studies have related such increase with exposure of hydrophobic surface area, which is consistent with $WH21^0$ being less polar than $WH21^+$. This observation may be rationalized by the stronger solute-solvent interactions found in the $WH21^0$ 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 a very likely formation of an intermediate in the progression pathway, as suggested previously[24].

Computationally, we have made use of the previously generated 13 microsecond MD trajectory of the neutral (deprotonated) peptide[24] 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 to the free energy landscape and kinetic network of the peptide. In the neutral form, $WH21^0$, the "helix" metastable states are more highly populated and exhibit longer lifetimes than the "coil", which is contrary to the phenomenon observed in the $WH21^+$ [27]. 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 state 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…His interaction in $WH21^0$ changes the nucleation pattern and leads to a preference for a central region intermediate, rather than the "off-center" intermediate seen in the $WH21^+$.

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

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 TG-MCB 16009 for computer time. 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 1807852.

Supporting information. Additional figures and tables describing details of the simulation.

References.

1. Thompson, P. A.; Eaton, W. A.; Hofrichter, J. Laser Temperature Jump Study of the Helix reversible arrow Coil Kinetics of an Alanine Peptide Interpreted with a 'Kinetic Zipper' Model. *Biochemistry* 1997, *36*, 9200-9210.
2. Thompson, P. A.; Munoz, V.; Jas, G. S.; Henry, E. R.; Eaton, W. A.; Hofrichter, J. The Helix-Coil Kinetics of a Heteropeptide. *J. Phys. Chem. B* 2000, *104*, 378−389.
3. Jas, G. S.; Eaton, W. A.; Hofrichter, J. Effect of Viscosity on the Kinetics of Alpha-Helix and Beta-Hairpin Formation. *J. Phys. Chem. B* 2001, *105*, 261−272.

4. Jas, G. S.; Middaugh, C. R.; Kuczera, K. Non-Exponential Kinetics and a Complete Folding Pathway of an α-Helical Heteropeptide: Direct Observation and Comprehensive Molecular Dynamics. *J. Phys. Chem. B* 2014, *118*, 639−647.

5. Mohammed, O.F.; Jas, G.S.; Lin, M.M.; Zewail, A. H. Primary Peptide Folding Dynamics Observed with Ultrafast Temperature Jump. *Angew. Chem. Int. Ed*. 2009, *48*, 5628−5632.

6. Lin, M. M.; Mohammed, O. F.; Jas, G. S.; Zewail, A. H. Speed Limit of Protein Folding Evidenced in Secondary Structure Dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 2011, *108*, 16622−16627.

7. Jas, G. S.; Hegefeld, W. A.; Majek, P.; Kuczera, K.; Elber, R. Experiments and Comprehensive Simulations of the Formation of a Helical Turn. *J. Phys. Chem. B* 2012, *116*, 6598−6610.

8. Kuczera, K.; Jas, G. S.; Elber, R. Kinetics of Helix Unfolding: Molecular Dynamics Simulations with Milestoning. *J. Phys. Chem. A* 2009, *113*, 7461−7473.

9. Peter Májek, P.; Elber, R. Milestoning without a Reaction Coordinate. *J. Chem. Theory Comput.* 2010, *6*, 1805-1817.

10. Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of Protein Folding: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* 1997, *48*, 545−600.

11. Dobson, C. M.; Sali, A.; Karplus, M. Protein Folding: A Perspective from Theory and Experiment. *Angew. Chem. Int. Ed.* 1998, *37*, 868−893.

12. Thirumalai, D.; Klimov, D. K. Deciphering the Timescales and Mechanisms of Protein Folding Using Minimal Off-lattice Models. *Curr. Opin. Struct. Biol.* 1999, *9*, 197− 207.

13. Sorin, E. J.; Pande, V. S. Exploring the Helix-Coil Transition via All Atom Equilibrium ensemble simulations. *Biophys. J.* 2005, *88*, 2472− 2493.

14. Hegefeld, W. A.; Chen, S. E.; DeLeon, K. Y.; Kuczera, K.; Jas, G. S. Helix Formation in a Pentapeptide: Experiment and Force-field Dependent Dynamics. *J. Phys. Chem. A* 2010, *114*, 12391−12402.

15. Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic Discovery of Metastable States for the Construction of Markov Models of Macromolecular Conformational Dynamics. *J. Chem. Phys.* 2007, *126*, 155101.

16. Kube, S.; Weber, M. A Coarse Graining Method for the Identification of Transition Rates Between Molecular Conformations. *J. Chem. Phys.* 2007, *126*, 024103.
17. Rains, E. K.; Andersen, H. C. A Bayesian Method for Construction of Markov Models to Describe Dynamics on Various Time-Scales. *J. Chem. Phys.* 2010, *133*, 144113.
18. Hummer, G.; Szabo, A. Optimal Dimensionality Reduction of Multistate Kinetic and Markov State Models. *J. Phys. Chem. B* 2015, *119*, 9029−9037.
19. Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem*. 2009, *30*, 1545−1614.
20. Buchete, N. V.; Hummer, G. Coarse Master Equations for Peptide Folding Dynamics. *J. Phys. Chem. B* 2008, *112*, 6057−6069.
21. Senne, M.; Trendelkamp-Schroer, B.; Mey A. S. J. S.; Schütte, C.; Noé, F. EMMA: A Software Package for Markov Model Building and Analysis. *J. Chem. Theor. Comput.* 2012, *8*, 2223-2238.
22. Trendelkamp-Schroer, B.; Wu, H.; Paul, F.; Noé, F. Estimation and Uncertainty of Reversible Markov Models. *J. Chem. Phys.* 2015, *143*, 174101.
23. *Kirchhoff, W. H.Exam (Coden: Ntnoef). U.S. Government Printing Office: Washington, DC,* ***1993****; Vol. 1401.*
24. Jas, G.S., Kuczera, K. Deprotonation of a Single Amino Acid Residue Induces Significant Stability in an α-Helical Heteropeptide. *J. Phys. Chem. B* 2018 [Epub ahead of print, PMID: 30339390]
25. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced and Scalable Molecular Simulation. *J. Chem. Theory. Comput.* 2008, *4*, 435−447.
26. Buchete, N. V.; Hummer, G. Coarse Master Equations for Peptide Folding Dynamics. *J. Phys. Chem. B* 2008, *112*, 6057−6069.
27. Jas, G.S., Kuczera, K. Helix-Coil Transition Courses Through Multiple Pathways and Intermediates: Fast Kinetic Measurements and Dimensionality. *J. Phys. Chem. B* 2018 [Epub ahead of print, PMID: 30395709]

28. Senne, M.; Trendelkamp-Schroer, B.; Mey A. S. J. S.; Schütte, C.; Noé, F. EMMA: A Software Package for Markov Model Building and Analysis. *J. Chem. Theor. Comput.* 2012, *8*, 2223-2238.
29. Taylor, J., Greenfield, N.J., Wu, B., Privalov, P.L. A calorimetric study of the folding-unfolding of an α-helix with covalently closed N and C-terminal loops. *JMB*, 1999, *291*, 965-976.
30. Jorgensen, W.L, Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, 79, 926-935.
31. MacKerell, Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R.L.;Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-atom empirical potential for molecular modeling and dynamics Studies of proteins. *J. of Phys. Chem. B,* **1998**, 102, 3586-3616.

Figure 1.

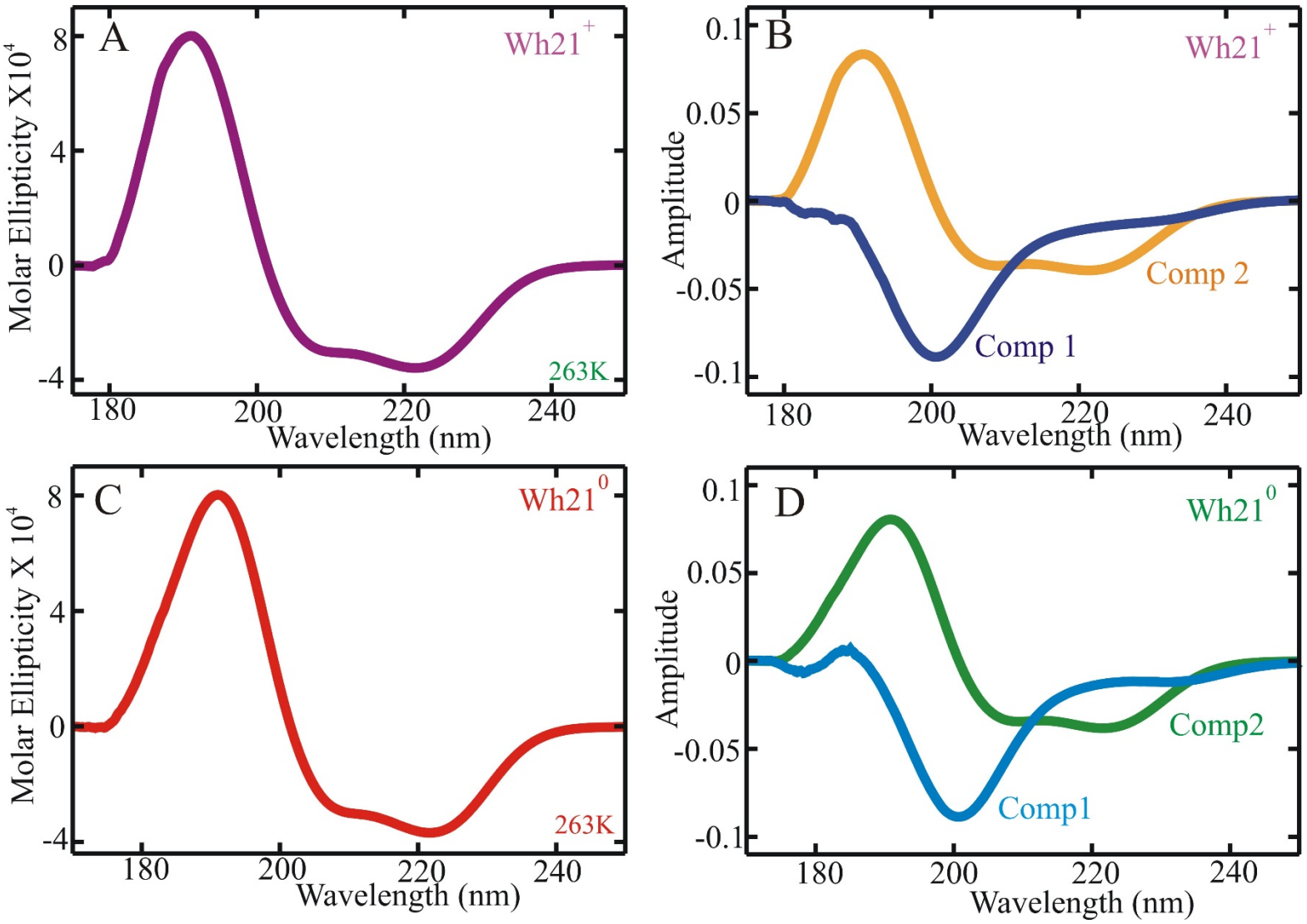
A
Wh21+
263K
Molar Ellipticity X10^4
8
4
0
-4
180
200
220
240
Wavelength (nm)
B
Wh21+
Comp 2
Comp 1
Amplitude
0.1
0.05
0
-0.05
-0.1
180
200
220
240
Wavelength (nm)
C
Wh21^0
263K
Molar Ellipticity X 10^4
8
4
0
-4
180
200
220
240
Wavelength (nm)
D
Wh21^0
Comp2
Comp1
Amplitude
0.1
0.05
0
-0.05
-0.1
180
200
220
240
Wavelength (nm)

Figure 2.

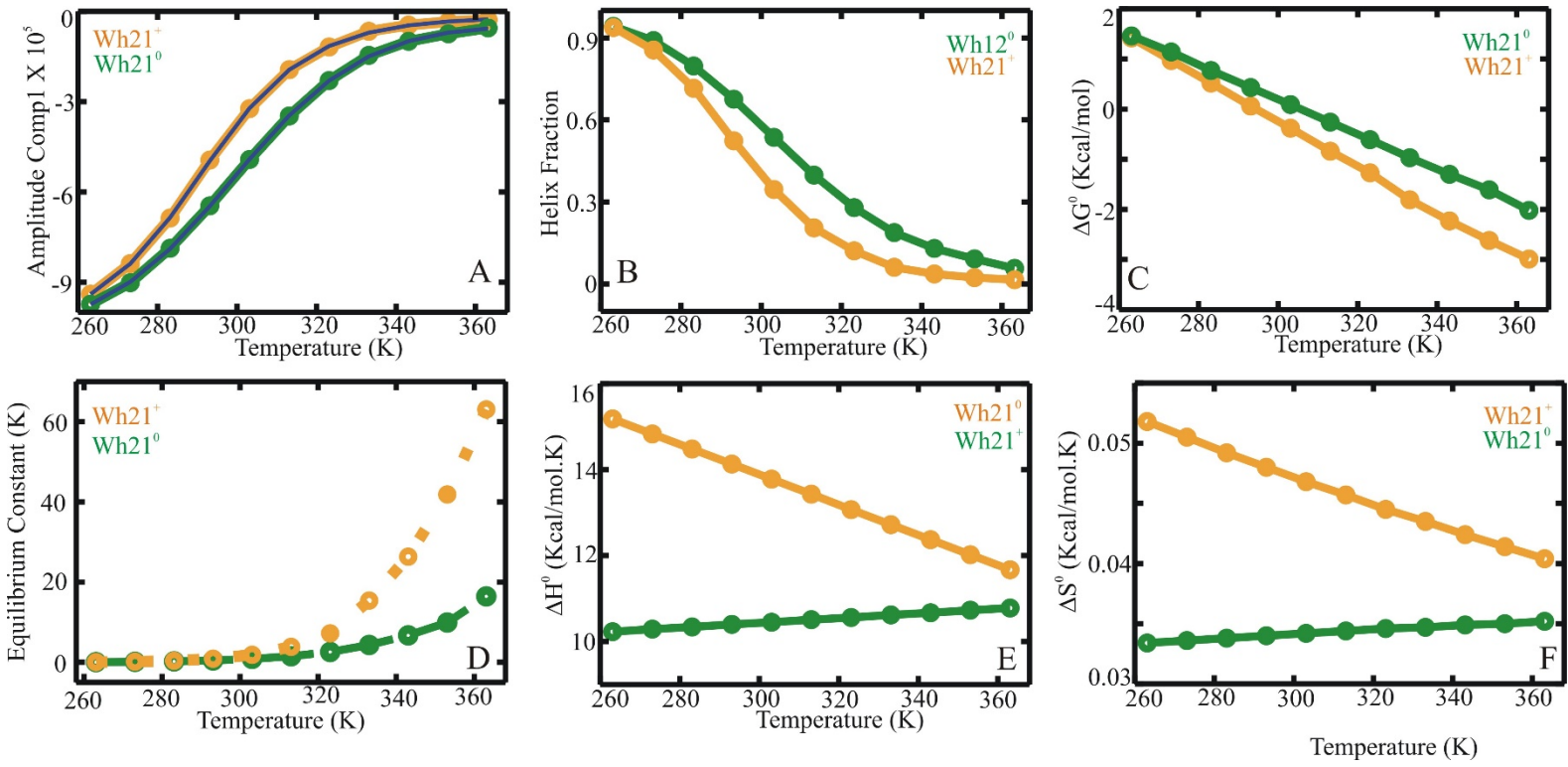

Figure 3.

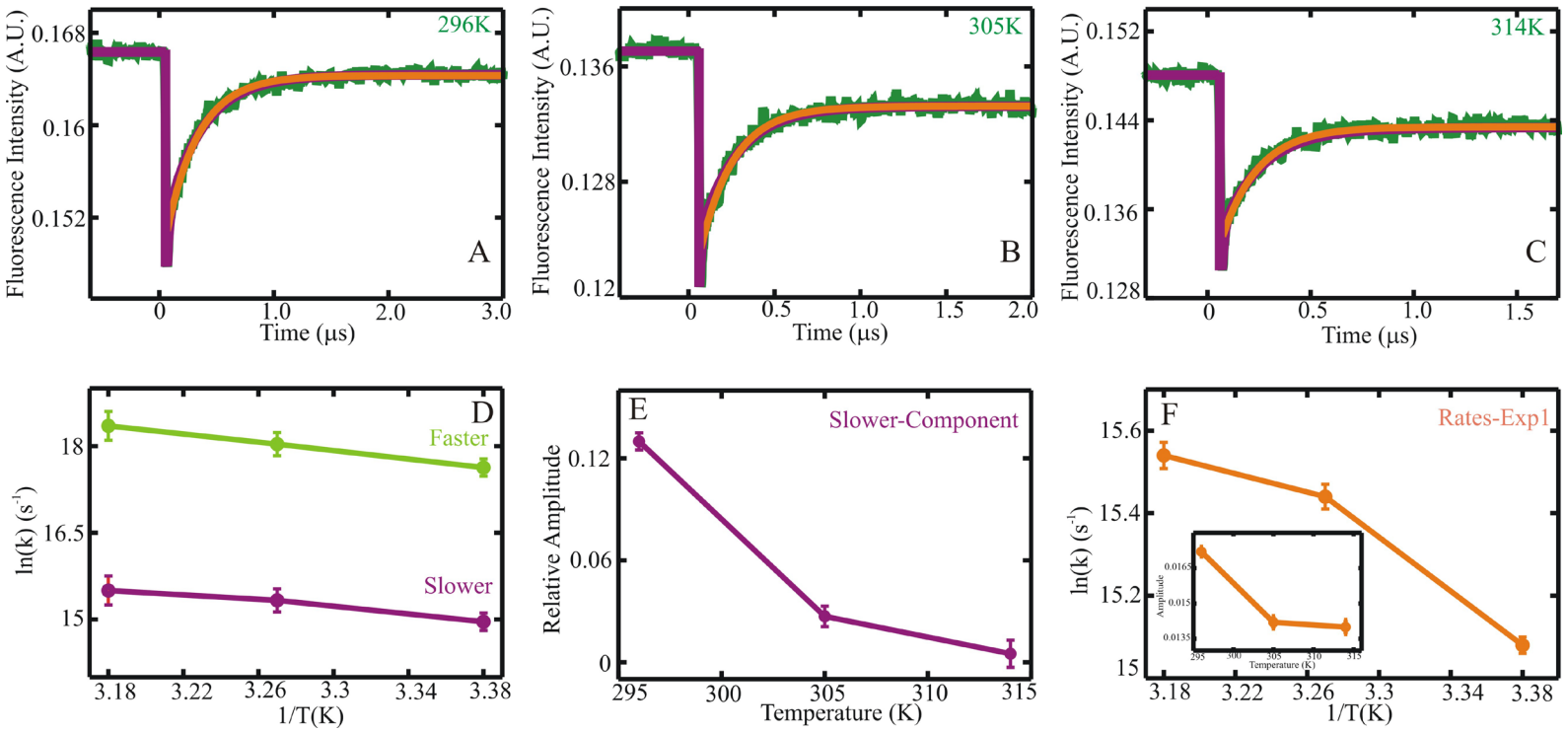

Figure 4.

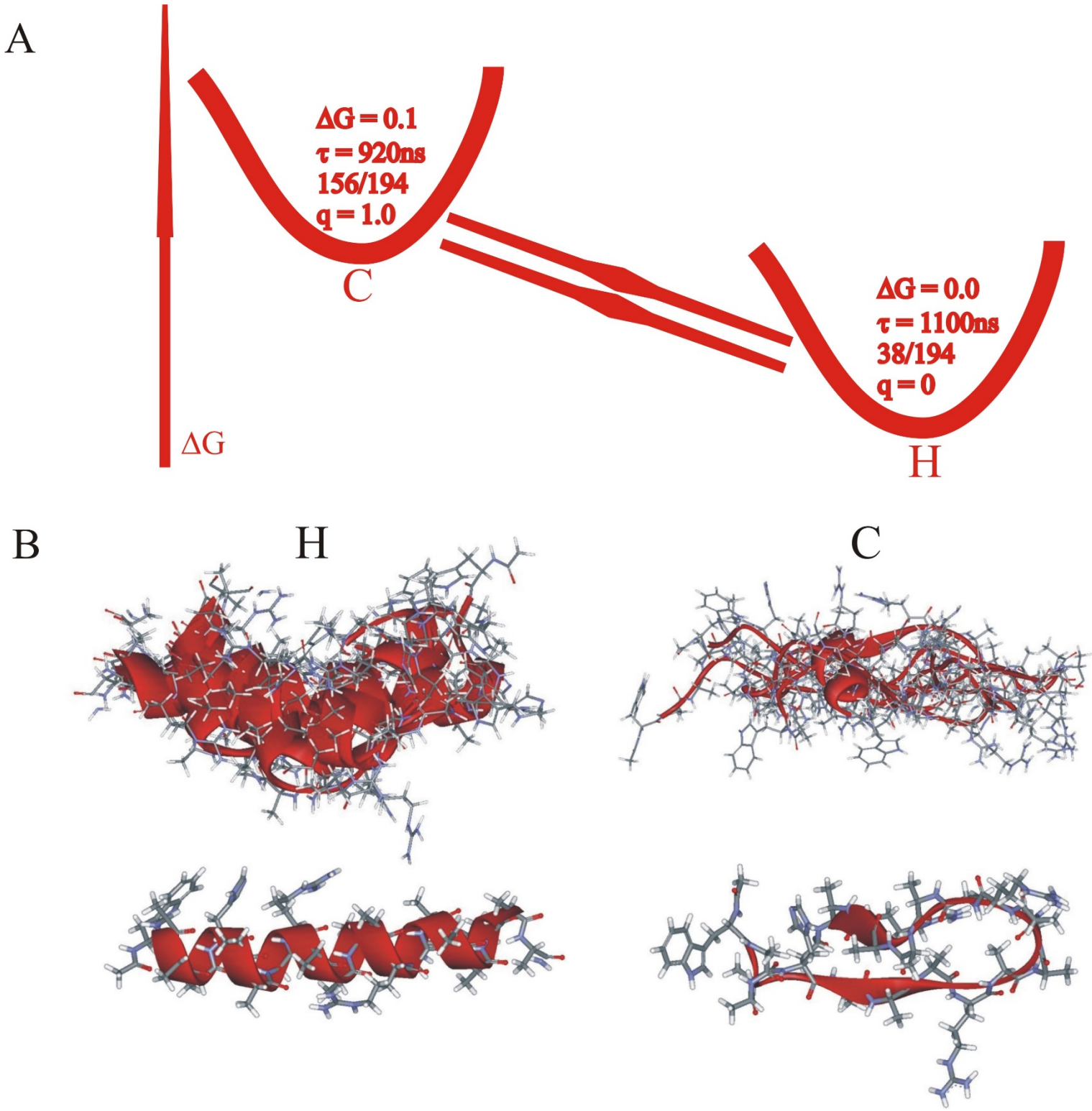

Figure 5.

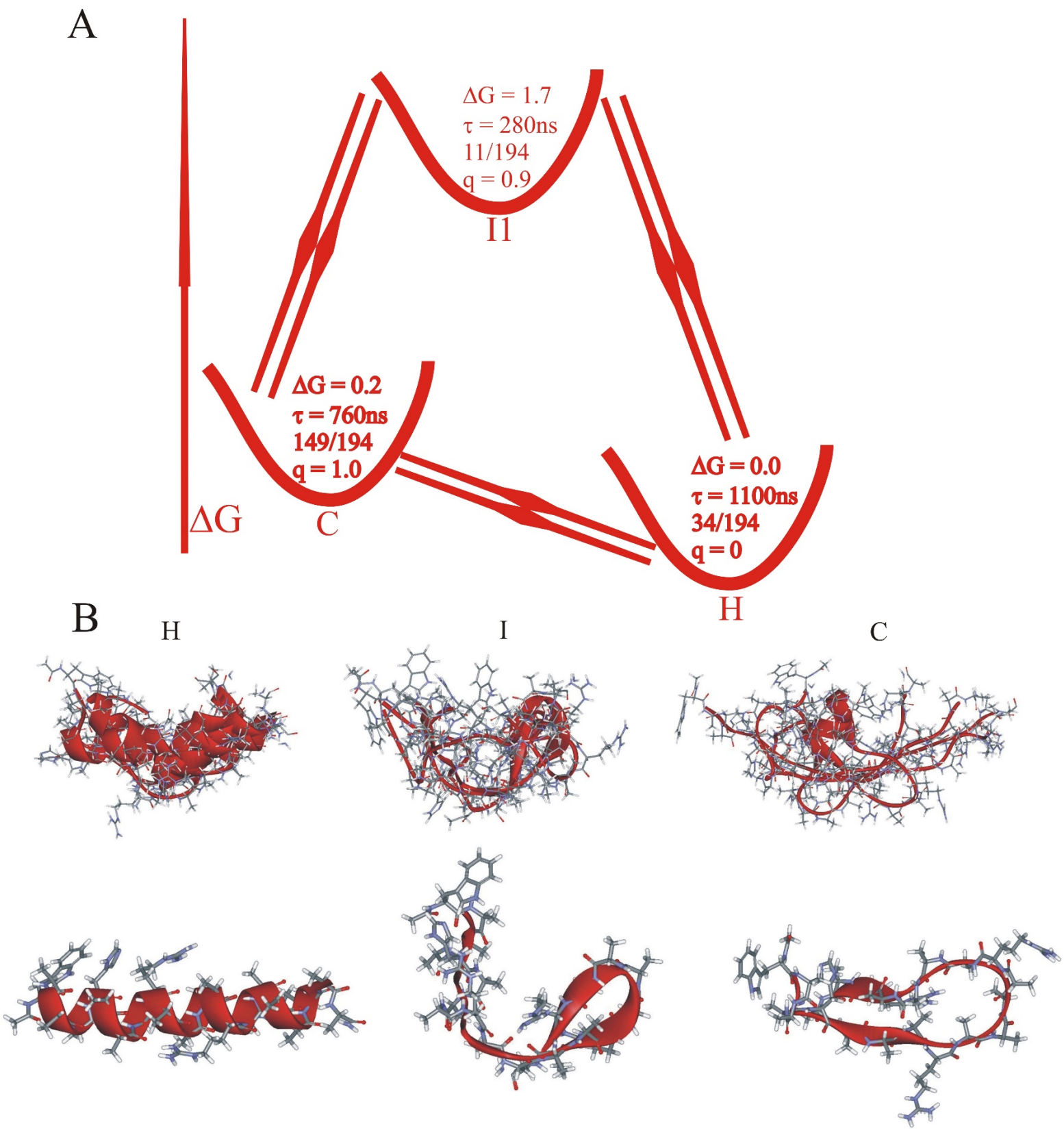
A
ΔG = 1.7
τ = 280ns
11/194
q = 0.9
I1
ΔG = 0.2
τ = 760ns
149/194
q = 1.0
ΔG = 0.0
τ = 1100ns
34/194
q = 0
ΔG
C
H
B
H
I
C

Figure 6.

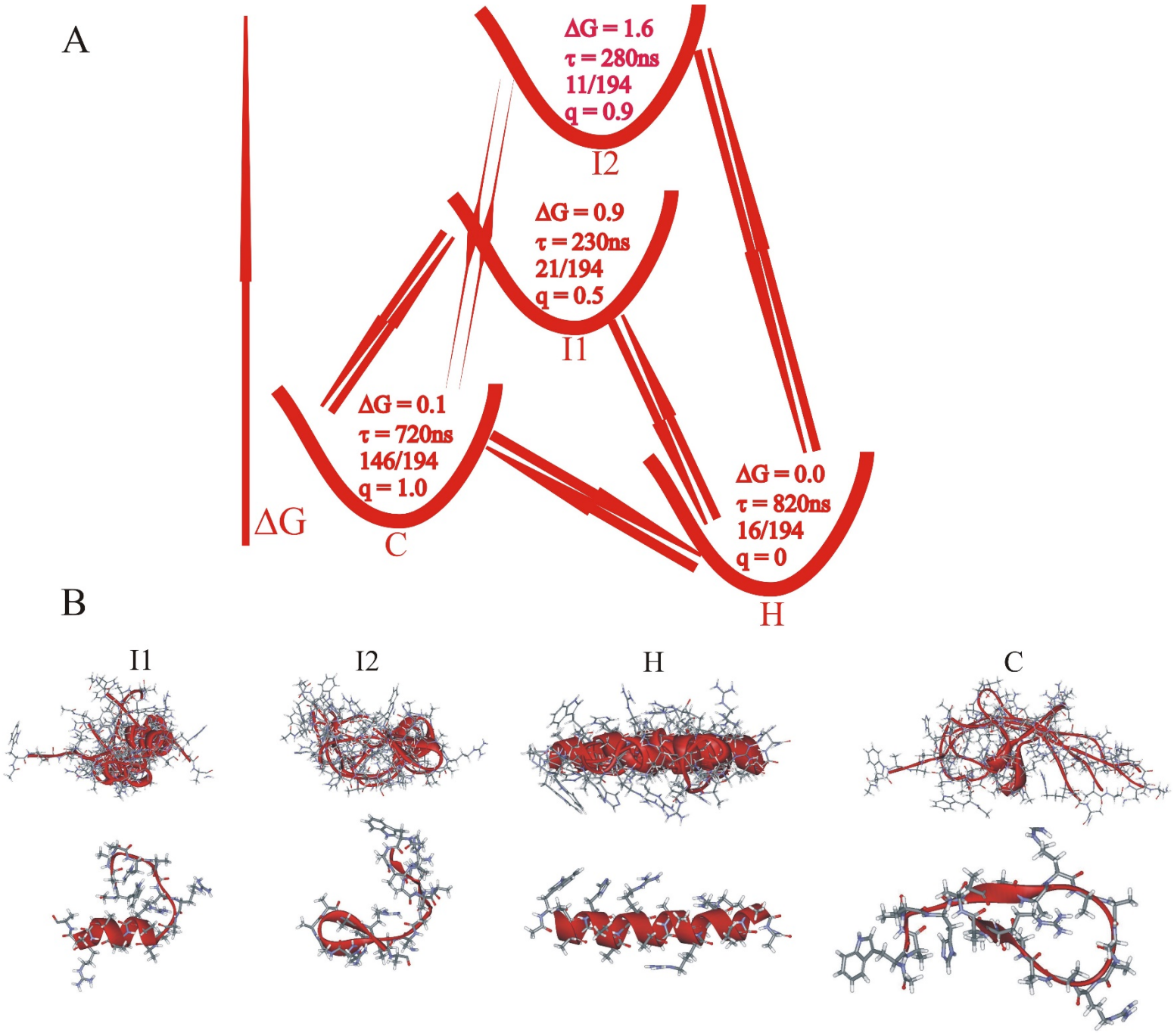
A
ΔG = 1.6
τ = 280ns
11/194
q = 0.9
I2
ΔG = 0.9
τ = 230ns
21/194
q = 0.5
I1
ΔG = 0.1
τ = 720ns
146/194
q = 1.0
C
ΔG = 0.0
τ = 820ns
16/194
q = 0
H
ΔG
B
I1
I2
H
C

Figure 7.

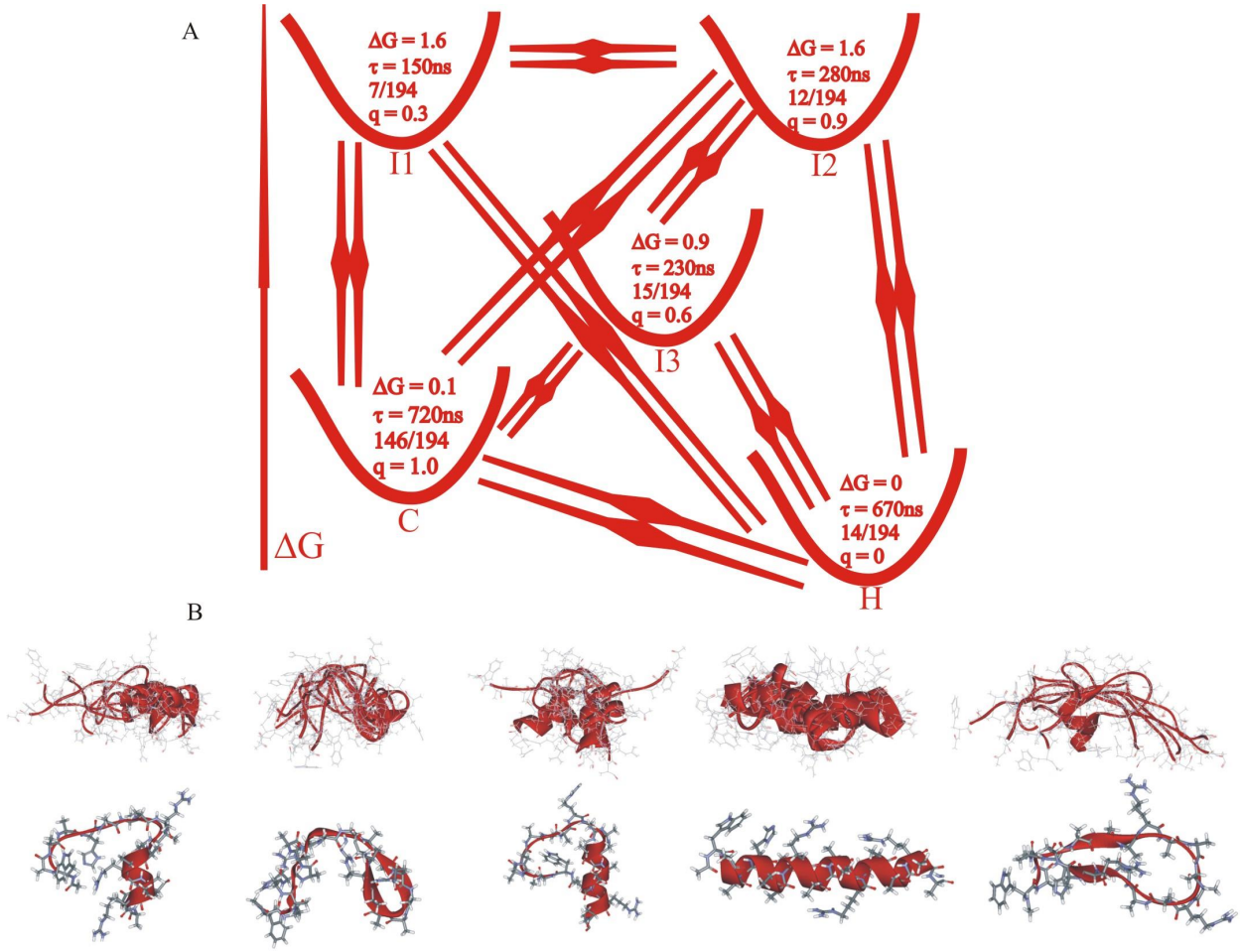
A
ΔG = 1.6
τ = 150ns
7/194
q = 0.3
I1
ΔG = 1.6
τ = 280ns
12/194
q = 0.9
I2
ΔG = 0.9
τ = 230ns
15/194
q = 0.6
I3
ΔG = 0.1
τ = 720ns
146/194
q = 1.0
C
ΔG = 0
τ = 670ns
14/194
q = 0
H
ΔG
B

Figure 8.

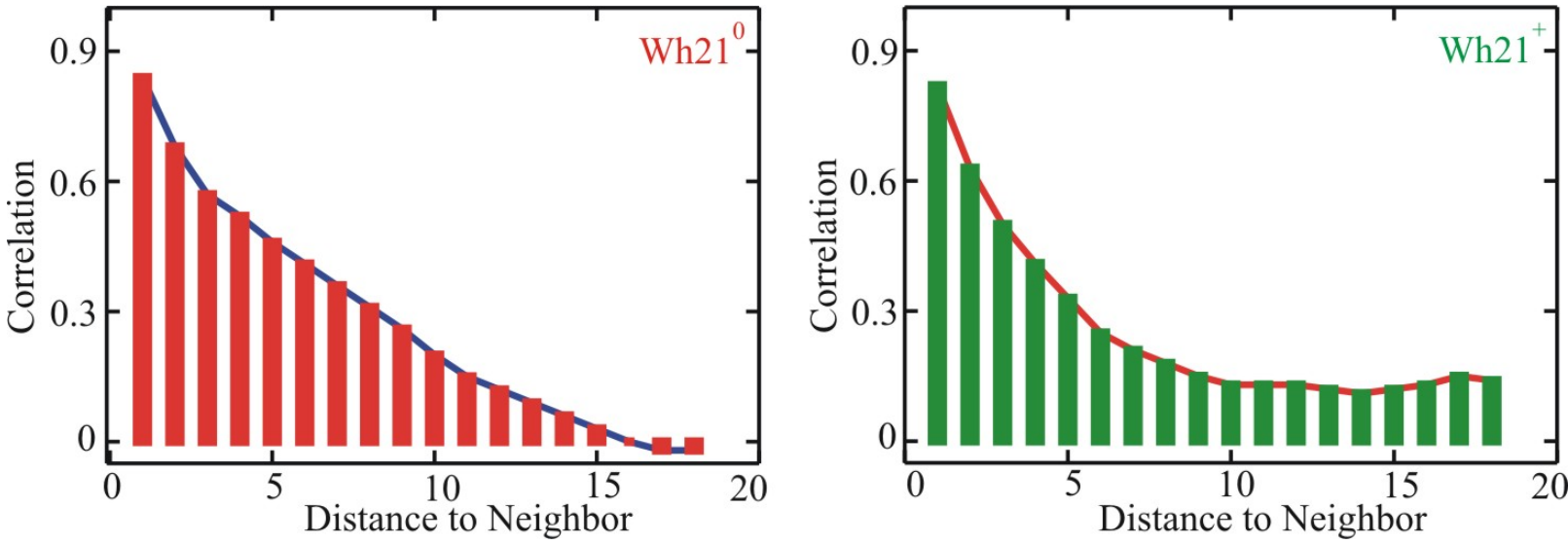

Figure 9.

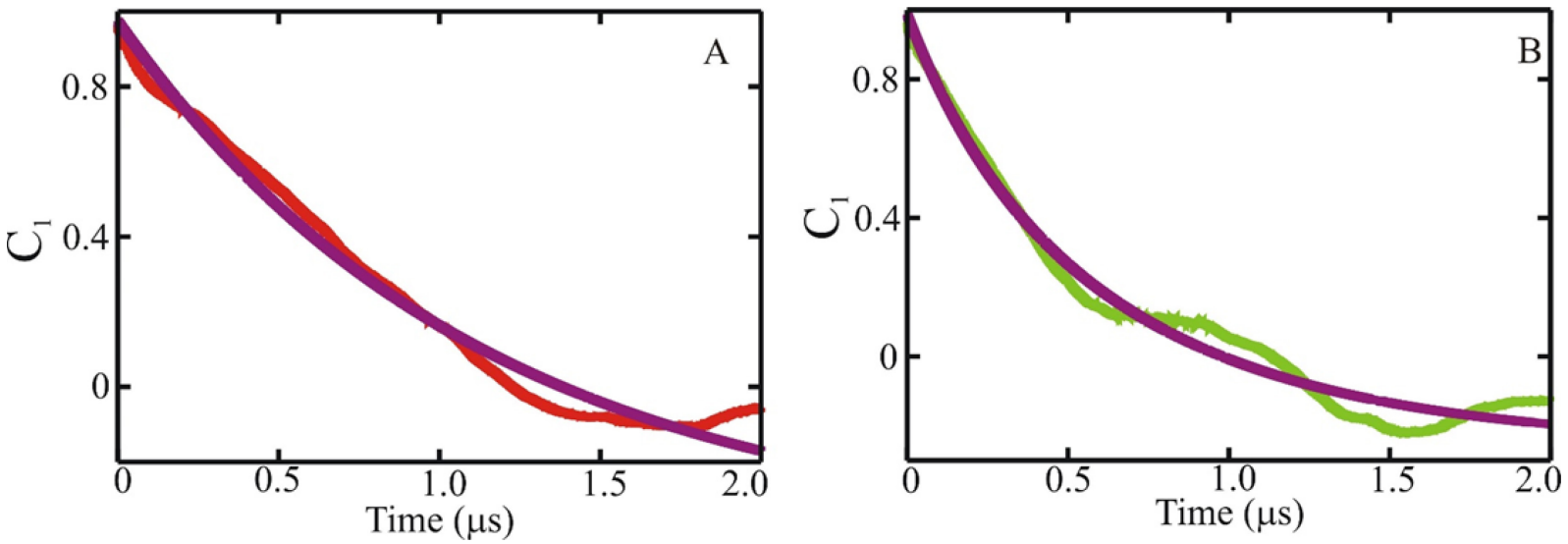

Figure 10.

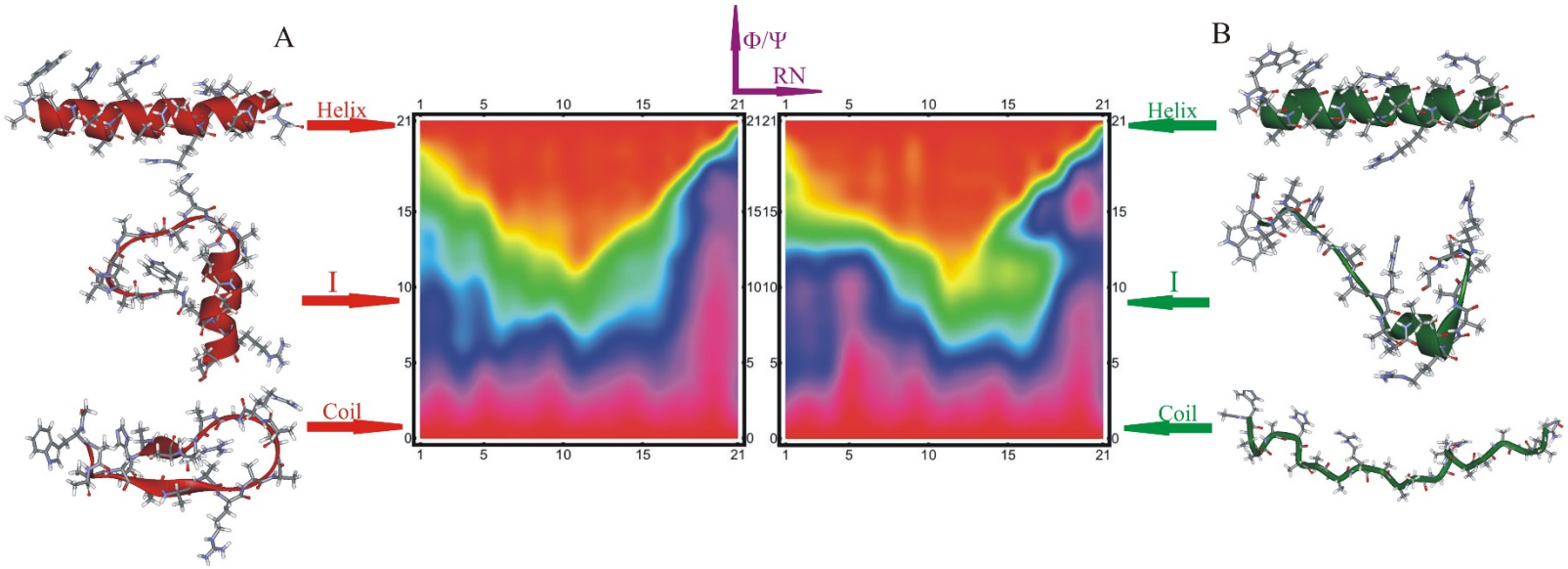

Figure Captions.

Figure 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 studied alpha-helix at 263K. (B) Application of SVD on the wavelength dependent CD spectra between 263K-363K produced classic features of an alpha-helix (orange) and coil (blue). (C) Wavelength dependent CD spectrum of the neutral form of studied alpha-helix at 263K. (D) SVD components, alpha-helix (green) and coil (blue) of wavelength dependent CD spectra between 263K-363K in the neutral form.

Figure 2.

Detailed thermodynamic analysis of the protonated and neutral form the alpha-helical hetero peptide from the amplitude vectors of the fist SVD component. (A) Amplitude vectors of the first component as a function of temperature for both protonated (orange) and neutral (green) form of the alpha-helix. Circles are vector amplitudes and the solid-lines are the two-state fit to the data with EXAM[23] (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 (ΔG^0) as a function of increasing temperature for protonated (orange) and neutral (green) . (D) Equilibrium constant as a function of increasing temperature of the protonated (orange) and neutral (green) forms. (E) Standard enthalpy (ΔH^0) of unfolding with the change in temperature of the protonated (orange) and neutral (green) form. (F) Entropy (ΔS^0) changes of unfolding with respect to temperature for the neutral (green) and protonated (orange) forms.

Figure 3. Nanosecond to microsecond relaxation dynamics measured with laser t-jump of the protonated form of alpha-helix $WH21^+$. Measured data at the three different final temperatures (A) 296K, (B) 305K and (C) 314K are in green. A single exponential fit to the data are shown in orange. A bi-exponential fit to the data are 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 form a single exponential fit to the measured kinetic data. Relative amplitudes of the rates associated with the single exponential fit are presented in the in-set.

Figure 4. Reduced dimensionality analysis for $WH21^0$ with N=2. (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.

Figure 5. Reduced dimensionality analysis for $WH21^0$ 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.

Figure 6. Reduced dimensionality analysis for $WH21^0$ with N=4. (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.

Figure 7. 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 $WH21^0$ (B) Protonated form $WH21^+$.

Figure 8. Reduced dimensionality analysis for $WH21^0$ with N=5. (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.

Figure 9. Average autocorrelation functions for length fluctuations of the 19 helical hydrogen bonds ($C{=}O_i\ldots H\text{-}N_{i+4}$). (A) $WH21^0$, average autocorrelation in red, two-exponential fit in purple. (B) $WH21^+$, average autocorrelation in green, two-exponential fit in purple.

Figure 10. The statistical path of helix folding. Data show the helix populations of individual residues (horizontal axis, NR=1-21) averaged aver 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^o radius of the ideal helix location taken from the final 500 ns of all 40 replicas of replica-exchange molecular dynamics [24,27]. The colors encode the average helix population of each peptide residue in the 0-1 range; red – 0.0, green – 0.5, purple – 1.0 . The molecular structures are cluster centers for the N=3 model. (A) Neutral, $WH21^0$ [24] (B) Protonated, $WH21^+$ [27].