# What Is the Time Scale for α-Helix Nucleation?

David De Sancho and Robert B. Best*

Department of Chemistry, Cambridge University, Lensfield Road, Cambridge CB2 1EW, U.K.

S *Supporting Information*

ABSTRACT: Helix formation is an elementary process in protein folding, influencing both the rate and mechanism of the global folding reaction. Yet, because helix formation is less cooperative than protein folding, the kinetics are often multiexponential, and the observed relaxation times are not straightforwardly related to the microscopic rates for helix nucleation and elongation. Recent ultrafast spectroscopic measurements on the peptide Ac-WAAAH$^+$-NH$_2$ were best fit by two relaxation modes on the ~0.1−1 ns time scale,[1] apparently much faster than had previously been experimentally inferred for helix nucleation. Here, we use replica-exchange molecular dynamics simulations with an optimized all-atom protein force field (Amber ff03w) and an accurate water model (TIP4P/2005) to study the kinetics of helix formation in this peptide. We calculate temperature-dependent microscopic rate coefficients from the simulations by treating the dynamics between helical states as a Markov process using a recently developed formalism. The fluorescence relaxation curves obtained from simulated temperature jumps are in excellent agreement with the experimentally determined results. We find that the kinetics are multiphasic but can be approximated well by a double-exponential function. The major processes contributing to the relaxation are the shrinking of helical states at the C-terminal end and a faster re-equilibration among coil states. Despite the fast observed relaxation, the helix nucleation time is estimated from our model to be 20−70 ns at 300 K, with a dependence on temperature well described by Arrhenius kinetics.
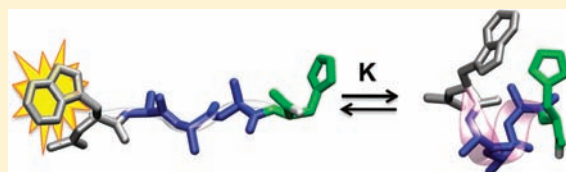
■ INTRODUCTION

Because of their importance in elementary steps of protein folding, the formation of α-helices, β-hairpins, and loops has recently attracted considerable attention,[2−4] with ultrafast spectroscopic techniques making it possible to monitor the nanosecond to microsecond time scales for formation of these fundamental units of structure.[5−9] These rates can be used to estimate the maximum rate at which proteins can fold, the so-called folding "speed limit",[10] and can also be applied to parametrize statistical mechanics models that help to interpret experiments.[11,12] The general consensus in the literature is that hairpins form on a microsecond time scale,[5,13,14] loops in tens to hundreds of nanoseconds,[8,9,15,16] and helices in around 100 ns to 1 μs[10] (although one study has suggested a much slower time scale[17]). The helix−coil transition in particular has served as a prototypical example of a biomolecular ordering transition and as such has received considerable attention from both experimentalists[1,7,17−26] and theorists.[27−42] The kinetics of helix formation can be complex to interpret due to the low cooperativity of this process, leading to very different signals and kinetics depending on the choice of molecule and experimental probe.[7,18,19,22,25,26,43] Despite this complexity, it has been possible to interpret much of this data on the basis of nucleation−elongation models for helix−coil kinetics, in which nucleation by formation of the first helical turn is followed by addition of helical residues to the end of an existing helix.[27,31,33,35] Most of the relaxation experiments, however, are primarily sensitive to re-equilibration processes between the helix and coil ensembles,[35] with helix nucleation being only indirectly observed.

The recent development of an ultrafast temperature jump instrument with picosecond time resolution[44] has made it possible to probe helix−coil kinetics in the pentapeptide Ac-WAAAH$^+$-NH$_2$ (hereafter W$_1$H$_5$). Since this peptide is barely larger than a single helical turn, it has been argued to be an ideal model for probing helix nucleation.[1] Experiments on W$_1$H$_5$ were designed such that the fluorescence of the N-terminal tryptophan is quenched on contact with the C-terminal histidine when the helix forms. The experimental signal was fitted by two exponential phases of roughly 0.8 and 5 ns at 300 K, remarkably 2 orders of magnitude faster than any process previously measured for the helix−coil transition; the two time scales were attributed to a bifurcation of pathways for helix formation.[1] A powerful method for investigating the microscopic processes involved in helix formation is molecular dynamics simulation, with increasingly accurate all-atom energy functions (force fields).[45−53] However, appropriate benchmarking against experiment is critical to validate the accuracy of the simulations.

Here, we embrace this task by examining helix nucleation in the W$_1$H$_5$ peptide using all-atom molecular dynamics simulations with explicit solvent. We use a version of the Amber ff03 force field[54] optimized against helix−coil transition data,[51] which is essential for obtaining realistic equilibrium populations, in conjunction with the highly optimized TIP4P/2005 water model.[55] In contrast to many water models often used in protein simulations, TIP4P/2005 gives realistic values for the solvent

ACS Publications   © 2011 American Chemical Society

6809

dx.doi.org/10.1021/ja200834s | *J. Am. Chem. Soc.* 2011, 133, 6809–6816

viscosity and self-diffusion coefficient,[55] important for quantitative rate calculations. We determine the rates of helix formation in the W_1H_5 peptide from long replica-exchange molecular dynamics (REMD) trajectories using a recently developed master-equation formalism.[56,57] This type of method, in conjunction with an appropriate model for the experimental signal, is ideally suited to give microscopic insight via quantitative comparison with macroscopic observables.[58−66] The fluorescence decays from simulated temperature jumps with our kinetic model are in quantitative agreement with the experimentally determined results. Although the dynamics are multiphasic, the overall decay is well-approximated by a double-exponential function. We find that the major contributions to the relaxation come from C-terminal shrinking of the helical states and a faster re-equilibration among the coil states. Despite the overall fast relaxation times, we estimate from our model that the time scale for helix nucleation (formation of three consecutive helical residues) is 20−70 ns, depending on the position in the sequence. This value is consistent with recent experimental estimates of the helix nucleation time.

■ METHODS

Atomistic Molecular Dynamics Simulations. We have run simulations of the N-terminally acetylated and C-terminally amidated W_1H_5 peptide (including protonated histidine) with 790 explicit water molecules in a truncated octahedral box with 25.8 Å distance between closest walls. One Cl^- ion was added to maintain electroneutrality. We use the Amber ff03w force field[51] for the peptide and the TIP4P/2005 model for water.[55] Long-range electrostatics were treated with the particle-mesh Ewald method, using a real-space cutoff of 0.9 nm and a grid spacing 0.1 nm. Lennard-Jones terms were truncated at 1.4 nm, and a multiple time step method was used for atom pairs between 0.9 and 1.4 nm. Determining accurate equilibrium populations in explicit solvent simulations is a challenging task, even for short peptides.[49] We therefore use 300 ns of REMD simulations, as implemented in the Gromacs 4.5.1 simulation package,[67] with 32 replicas spanning the temperature range of 270−600 K, and replica swaps attempted every 10 ps. A short constant pressure run at 300 K was used to estimate the equilibrium average volume at this temperature. Its final configuration, with the peptide being in state $hhccc$ (described below), was used to initialize the subsequent constant-volume REMD runs. A leapfrog-type integrator was used to propagate Langevin dynamics[68] with a 2 fs time step and a friction of 1 ps^{-1}.

■ RESULTS AND DISCUSSION

Helix−Coil Equilibrium. In Figure 1 we show Ramachandran potentials of mean force for the different residues in this peptide, close to room temperature ($T = 302$ K), showing the location of the populated helical and extended regions. There are significant differences between the free energy surfaces for the N-terminal tryptophan, the central alanines, and the C-terminal histidine, due to the combined effects of intrinsic helix propensity, helix capping, and steric hindrance. We find that ∼50 ns is needed for the REMD simulations to equilibrate (see Supporting Information (SI)), and so we use only the last 250 ns at each temperature for subsequent analysis. To monitor the helix−coil equilibrium, we calculate the fraction of helix, $\langle h \rangle_{sim}$, from the fraction of time spent by each residue within helical segments,[48] defined as sequences of three or more consecutive helical residues (see Figure 1B). As observed for the $(AAQAA)_3$ peptide, we find that there is a sharp change in $\langle h \rangle_{sim}$ at low temperature arising from the enhanced cooperativity of the peptide with the Amber03w force field.[51] We have fitted the equilibrium distribution of helical
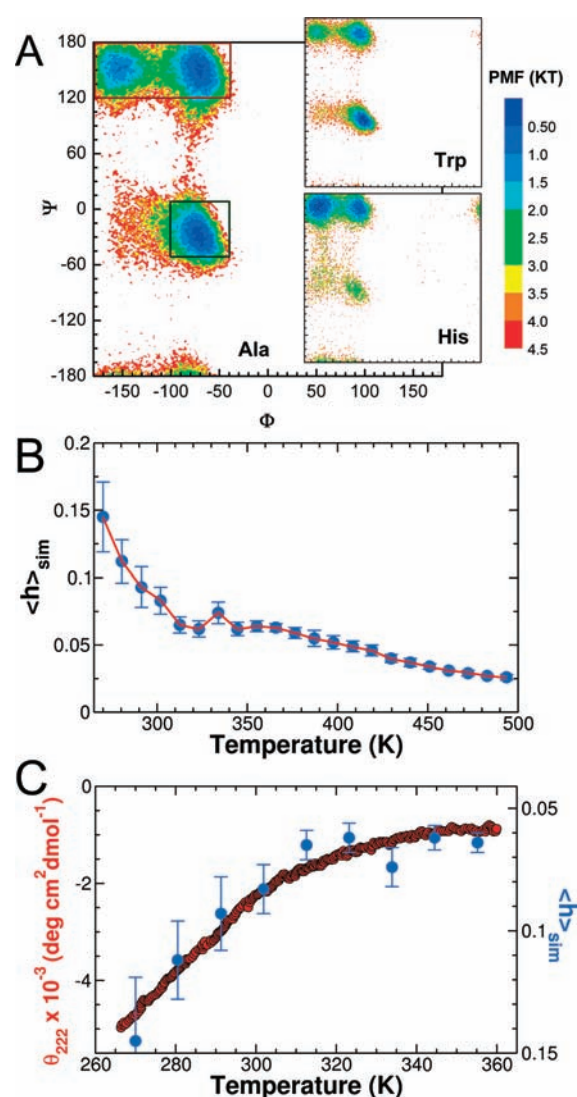
Figure 1. Thermodynamics of the W_1H_5 peptide. (A) Ramachandran potential of mean force (PMF) for the amino acid types present in the W_1H_5 peptide at 300 K. Rectangles in dark green and red in the alanine PMF mark, respectively, the regions corresponding to the helix $[(-100 < \Phi < -30), (-50 < \Psi < 10)]$ and coil $[(-180 < \Phi < -40), (120 < \Psi < 180)]$ states used in the transition-based assignment. Color contours represent $0.5RT$ energy changes. (B) Fraction of helix for the W_1H_5 peptide from the time spent by each residue in helical Ramachandran angles. The red line is a fit to the Lifson−Roig model. (C) Overlay of the fraction helix and experimental CD signal. Note that the Lifson−Roig model is fit independently to the distribution of conformations at each temperature as described in ref 48.

configurations to a Lifson−Roig model[69] using a Bayesian procedure, with the likelihood of the observed configurations i being $\mathcal{L} = \Pi_i \rho(i)$, where $\rho(i)$ is the equilibrium probability of i given by the Lifson−Roig partition function. As in earlier work,[48,51] we fit independent Lifson−Roig parameters $w(T)$ and $v(T)$ at each temperature T. We sample the posterior distribution of w,v assuming a uniform prior, using Metropolis Monte Carlo as described before.[48] The fit (Figure 1B) produces elongation (w) and nucleation (v) parameters comparable in both magnitude and temperature dependence to those obtained previously from simulations of longer peptides[48,51] (see SI for further details).
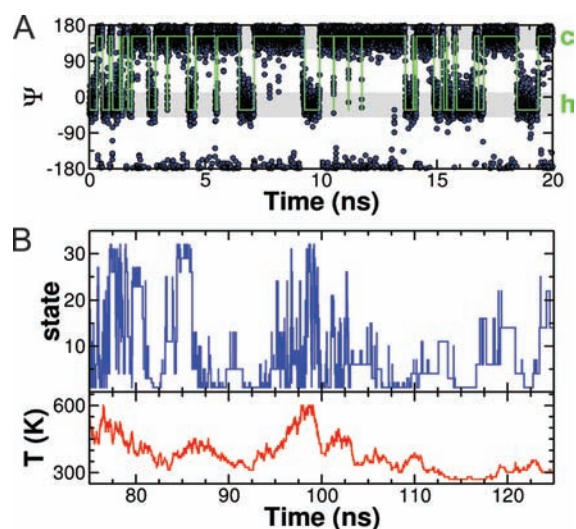
Figure 2. Transition-based assignment of helix−coil states. (A) Time series illustrating the assignment of the Ψ Ramachandran angle for a single amino acid residue (circles) to helix and coil conformational states (green line). The gray areas mark the regions corresponding to the helix and coil states. (B) Time series illustrating the transition-based assignment of protein conformations while the system diffuses through state (top) and temperature space (bottom).

While quantitative calculation of ellipticity from simulation is a challenging problem beyond the scope of this paper,[70,71] we note that the values of $\langle h \rangle_{sim}$ can be overlaid on the experimental far-UV CD data that probes helix populations (see Figure 1C). This suggests good agreement between the folding thermodynamics of W_1H_5 in our simulations, previous work with the refined force fields, and experimental data (see also below for comparison with fluorescence data).

Dynamics of Helix Formation. Although REMD is usually employed as a thermodynamic sampling method only, kinetic information may be inferred from the short trajectory segments between exchanges. We use the method developed by Buchete and Hummer to determine the helix−coil dynamics in a coarse-grained state space,[56,57] summarized briefly here. After subdividing a system into N coarse states, if the dynamics between these states can be considered Markovian (i.e., memory-less), the time evolution of the coarse states can be represented in the form of a master equation $dP_i(t)/dt = \Sigma_{j=1}^N k_{ij}P_j(t)$ or, in vector-matrix notation, $d\mathbf{P}(t)/dt = \mathbf{KP}(t)$. Here $P_i(t)$ is the population of state i at time t, and k_{ij} are the elements of the transition rate matrix, defined as $k_{ij} \geq 0$ for $i \neq j$ and $k_{ii} = -\Sigma_{j \neq i}k_{ji}$ for diagonal elements. For the W_1H_5 peptide we partition the conformational space in terms of Ramachandran dihedral angles (Φ,Ψ) (see Figure 1A) allowing for two possible states—helix and coil—for each residue, yielding a total of 32 (i.e., 2^5) coarse states in the system. Instead of assigning states considering only the instantaneous values of Φ and Ψ, we use the context of the simulation trajectory to aid in the assignment of true transitions: the alpha and extended regions are tightly defined (Figure 1A), and a transition from one to the other is only counted once the trajectory crosses completely into the product state[56] (see Figure 2A). In this way we suppress the counting of barrier recrossings as events.[56,72] After assigning the transitions for all residues in the pentapeptide, we obtain the number of transitions N_{ij} from j to i for each pair of coarse states i,j (see Figure 2B) after a given, fixed lag time Δt. The state assignment is done on the
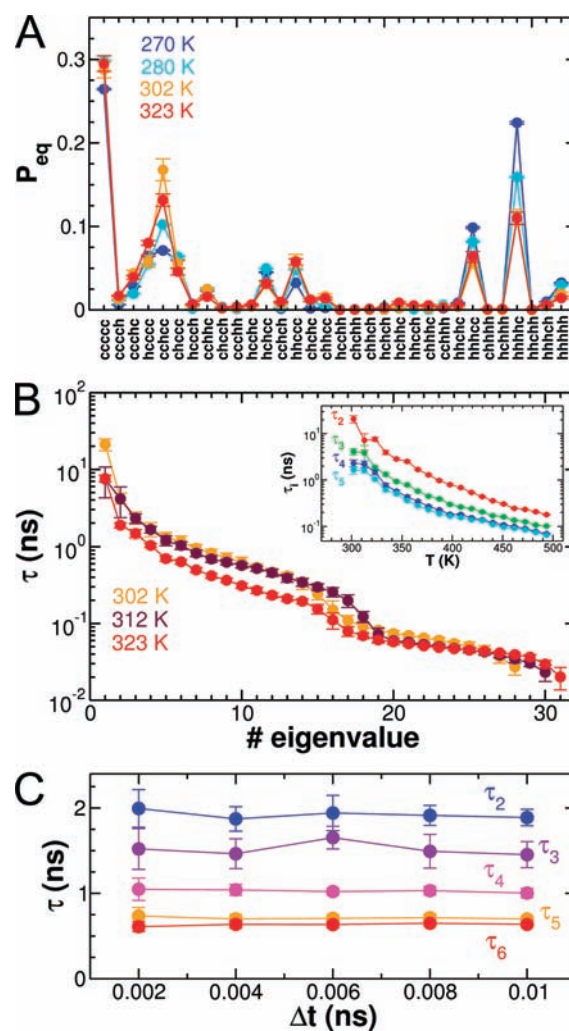
Figure 3. Results from maximum likelihood optimization of rate matrices. (A) Equilibrium population of the 32 states in the Markov model from the likelihood maximization at experimentally relevant temperatures. Error bars are calculated from bootstrap tests. (B) Eigenvalue spectrum at temperatures closest to the final temperatures in the temperature jump experiments. Inset: temperature dependence of the relaxation times for the slowest modes of the system. (C) Dependece of first five nonzero eigenvalues on lag time.

continuous REMD simulation trajectories obtained by following the replicas through temperature space, with the observed transitions being assigned to the temperature at which they occur (see Figure 2B).[57] We use the transition count matrix with elements $N_{ij}(T)$ to optimize the rate matrix by maximizing the likelihood, defined as $\mathcal{L} = \Sigma_{i=1}^N\Sigma_{j=1}^N[p(i,\Delta t|j,0)]^{N_{ij}}$. In this equation the propagator $p(i,\Delta t|j,0)$ is the probability of being in state i at time Δt after being in state j, which can be obtained directly from the rate matrix exponential as $(\exp[\mathbf{K}\Delta t])_{ij}$.[56] For the optimization we run several cycles of simulated annealing Monte Carlo, with the pseudo-energy function being $-\ln\mathcal{L}$ and the simulated annealing "temperature" slowly decreasing to zero. The floating parameters in the search are the equilibrium populations and half the nondiagonal elements of the rate matrix, the remainder being determined by detailed balance.[73]

In Figure 3A we show the equilibrium populations at experimentally relevant temperatures, which can be obtained either

from the raw simulation data or from the stationary mode of the optimized rate matrix. Because of the size of the peptide, the predominant states are nonhelical even at low temperatures. Interestingly, we find that the most populated helical state is *hhhhc* rather than *hhhhh*; this may be explained by the low helix propensity of protonated histidine, as well as by stronger interactions between tryptophan and histidine side chains in *hhhhc* in comparison with *hhhhh* (see SI). Evidence for this is that the distribution of side-chain distances for these two residues is shifted to smaller distances in *hhhhc* relative to *hhhhh*, and the fact that the effect is not seen in long simulations of Ala$_5$ with the same force field. As expected, we see that the population of the most significant helical states (*hhhhh*, *hhhhc*, and *hhhcc*) decreases upon raising the temperature, while nonhelical states become more populated. This microscopic picture is consistent with the global trends shown in Figure 1. Despite the extensive sampling in our REMD simulations, with an aggregate time of almost 10 μs, we find that some rare states are not populated at low temperatures. These correspond to "broken" helices (i.e., *hhchh* or *hchhh*) which are highly disfavored in short peptides, or states where the helix propagates from the C-terminus (i.e., *chhhh*), which are unlikely due to the strong "capping" effect of the acetyl group and low helix propensity of protonated histidine. Although REMD ensures that the relative equilibrium populations of the different states should be accurate at all temperatures, direct determination of the rate matrix requires all relevant transitions to occur at each temperature of interest. We find at temperatures lower than 300 K that this condition is not satisfied, and some rare states become "disconnected" in the transition matrix. While this could be overcome by assuming a function (e.g., Arrhenius) $k_{ij}(T)$ for the rates, in this work one of our goals was to determine how the rates varied with temperature, so we did not enforce any particular temperature dependence. For temperatures above 300 K, we can obtain the fast and slow modes of the system from the eigenvalue spectrum of the rate matrix.[56,57] In Figure 3B we show the temperature dependence of the relaxation times for the slowest modes of the system. While there is a 4−5-fold difference between the second and third eigenmodes, the time scales for the next slowest modes are grouped quite closely together, a point to which we return below.

When constructing Markov models for dynamics, such as the master equation model we use here, it is important to test whether non-Markovian "memory" effects have been eliminated. This can be assessed from the dependence of the eigenvalues of the inferred rate matrix **K** on the lag time Δt used in its construction: for sufficiently long Δt, the rate matrix and its eigenvalues should become independent of the lag time.[56,57] A similar eigenvalue criterion is used to assess the convergence of the closely related transition matrices, $\mathbf{T}(\Delta t) = \exp[\mathbf{K}\Delta t]$, with lag Δt.[58,60,62] In Figure 3C we show that the slowest modes in our kinetic model are independent of Δt, supporting the previous observation that the context-based assignment efficiently removes non-Markovian effects in the dynamics, even for very short lags.[56,57]

Microscopic Rates of Nucleation and Elongation. The rates estimated in our master equation use a finer description of states than commonly employed in modeling helix−coil kinetics, in which helix formation is typically described by a "nucleation" step, comprising the formation of a complete helical turn, followed by "elongation" steps in which additional residues are added to the helix. To map our results onto these more familiar processes, we can compute a "nucleation" rate, i.e., the rate for
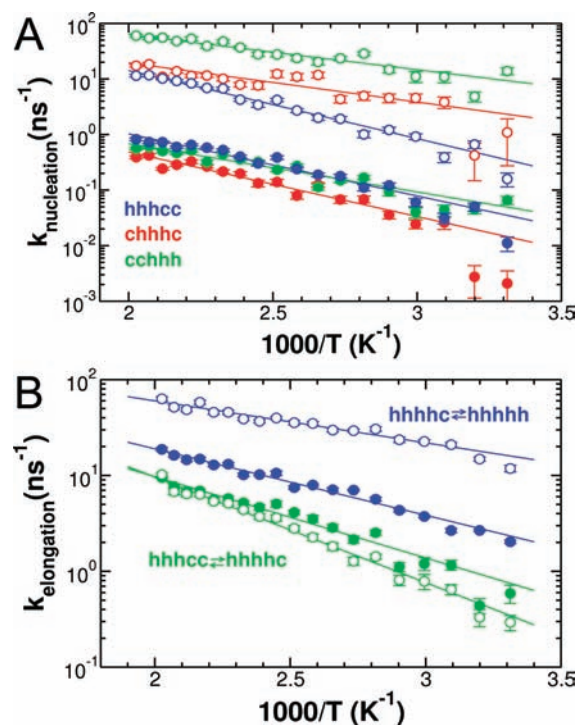
Figure 4. Rate coefficients for (A) helix nucleation and (B) helix elongation. Forward rates and reverse rates from the optimized rate matrices are shown as full and empty circles, respectively. Lines represent fits to an Arrhenius law.

formation of a helix nucleus (first helical turn) at a given point in the sequence. Note that for longer helices this will differ significantly from the rate for nucleating a helix at a specific location in the sequence. We estimate this from the total rate for conversion of coil states to a particular helix nucleus as $k_{\mathrm{nuc}} = \Sigma_{j\in\mathrm{coil}}\, p_j k_{ij}/\Sigma_{j\in\mathrm{coil}}\, p_j$, where state i is a particular helix nucleus and the states j are members of the "coil" ensemble (i.e., not containing any helical segments); the rate of the reverse process $k_{-\mathrm{nuc}}$ is given by $k_{-\mathrm{nuc}} = \Sigma_{j\in\mathrm{coil}}\, k_{ji}$. We define a helix nucleus to consist of three consecutive residues assigned as helical using the context-based assignment, surrounded by coil (e.g., ..*chhhc*.., *hhhc*.., etc.). The rates of nucleation are shown in Figure 4A as a function of temperature for the three possible nuclei in this peptide. We estimate a nucleation time of $\tau_{\mathrm{nuc}} = 1/k_{\mathrm{nuc}} \approx 20-70$ ns for a three-residue nucleus at 300 K, with nucleation at the center being slower than at the ends.

This rate is considerably slower than previous estimates based on molecular simulations, typically of the order 0.1−1 ns.[27,30] However, this is most likely due to the strong bias toward helical conformations in many older force fields,[47] reducing the nucleation barrier; even more recent force fields can result in significant differences in dynamics.[42] Estimates of nucleation times from experiment vary widely, partly because they are not probed directly, as mentioned above. On the basis of fitted helix elongation (or "zipping") rates of $k_{\mathrm{el}} \approx 2.9 \times 10^8$ s^{-1} in temperature jump experiments on alanine-based peptides,[20] a nucleation time of between 140 ns and 2.0 μs can be estimated, depending on the value of the nucleation parameter σ chosen (0.025 to 0.0017),[74] since $k_{\mathrm{nuc}} = \sigma k_{\mathrm{el}}$ in the commonly used microscopic model of helix kinetics due to Schwarz.[75] A slightly faster elongation rate of $\sim 1.3 \times 10^9$ s^{-1}, estimated from later

experiments with a Trp-His quenching pair,[7] would give estimates for the nucleation time of 30−450 ns with the same range of σ. A more recent study based on triplet−triplet energy transfer between xanthone and naphthylalanine used a nucleation time of 7.94 μs ($1/k_3$ in that work) to interpret their data,[26] and temperature jump studies of a helical peptide stabilized by side-chain salt bridges have also suggested a slower time scale.[43] A temperature jump experiment on helix formation from a cold unfolded state has also been used to probe the nucleation time, estimating a value of 300 ns.[23] Our calculated nucleation times are therefore within the range of previous experimental values, although still at the lower end. Possible reasons for the slightly accelerated nucleation rate are the stabilization of the helical state by the favorable interaction of the Trp-His pair[7] and the effects of helix capping,[24] which may lower the barrier relative to nucleation in the center of a longer peptide.

We can also estimate time constants for adding a single residue to a helix directly from our rate matrices. The helix elongation rates at 300 K are in the range of $(0.6-2) \times 10^9$ s^{-1}, consistent with the experimental estimates given above (see Figure 4B). We find that the temperature dependence of helix nucleation and elongation processes can be reasonably fit to an Arrhenius form, with only small deviations at low temperatures—this is most likely the result of the master equation being less well determined at these temperatures due to the fewer transitions. The Arrhenius approximation will be useful if this formalism is to be adapted for longer helical peptides.[56,73]

Comparison with Contact-Based Fluorescence Quenching. To validate the dynamics in our simulations, we compare our kinetic model with the ultrafast kinetic experiments on the W_1H_5 peptide previously reported.[1] As a probe for helix nucleation upon temperature jumps they measured the tryptophan fluorescence in the 310−500 nm window. To interpret the results they assumed that tryptophan and histidine interact (resulting in quenching of fluorescence) only when the entire peptide adopts helical Ramachandran angles. Since in the simulations we can directly monitor interactions between the Trp and His, we do not predetermine which states are quenched, but instead we compute a "quantum yield" for each state (Φ_s) using a simple approximation, relative to the situation where the His is absent. We assume that when the minimum Trp-His distance in a given conformation is smaller than a cutoff distance (r_{cut}), the fluorescence is completely quenched; otherwise, it is completely unquenched. Thus, the simulated quantum yield for each state is calculated as $\Phi_s = \int_{r_{cut}}^{\infty} p(r_{WH}|s)\,dr_{WH}$, where $p(r_{WH}|s)$ is the probability density for the minimum interatomic Trp-His distance (r_{WH}) given that the peptide is in state s. Here we compute Φ using r_{cut} = 4 Å from data at all temperatures to guarantee extensive sampling of the distances for each state. We note, however, that the estimate of the simulated quantum yield is robust and, for a short peptide like W_1H_5, the results do not depend significantly of the range of temperatures from which conformations are drawn or the choice of the cutoff distance (see SI). Our choice of a simple cutoff to define quenched configurations is motivated by the short-range nature of the quenching of the tryptophan by histidine and is the approximation that is most often used to interpret experiments. However, we have also considered an explicitly distance-dependent quenching rate, $k(r) = k_1 \exp[-(r - a)/b]$, motivated by the evidence that histidine quenches the tryptophan singlet excited state by an electron-transfer mechanism.[7,76] In this rate expression, k_1 is a prefactor, a is the distance of closest approach (contact), and b is
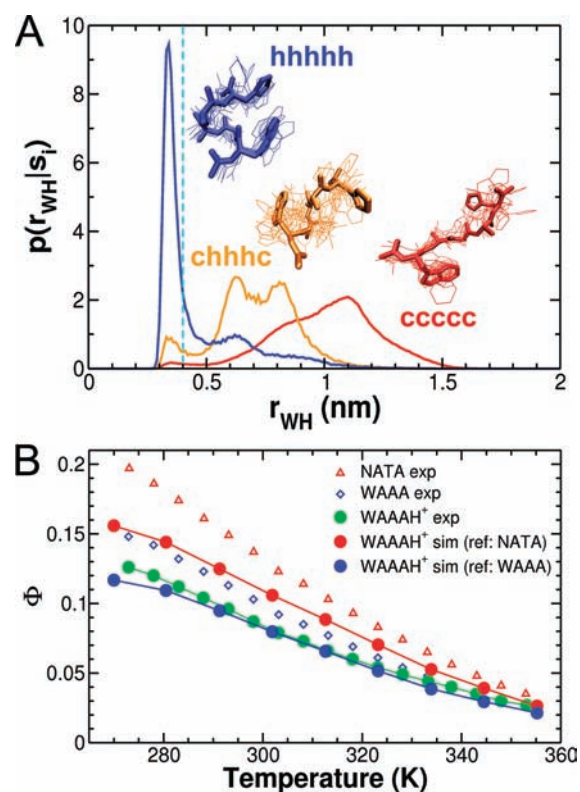

Figure 5. Simulation of the tryptophan fluorescence signal. (A) Probability density functions of the minimum distance between His and Trp for three representative states of the system. The dashed cyan line shows the cutoff distance for fluorescence quenching. (B) Comparison between experimental W_1H_5 fluorescence quantum yield (green circles)[7] and our estimates (full red and blue circles, respectively). These estimates are derived respectively by scaling the temperature-dependent fluorescence quantum yield of NATA (red triangles) and WAAA (blue diamonds)—representing the quantum yield of the Trp in the absence of the quenching His—by the relative quantum yield determined from the simulations.

the range of quenching. We also consider that the quantum yield relative to that in the absence of the quencher will depend on the dynamics within each helical state. We find that, for reasonable choices of the parameters k_1, a, and b, and in both extremes of fast and slow intrastate sampling, we obtain results comparable to those obtained with the simple cutoff (see SI).

In Figure 5 we show the probability density $p(r_{WH}|s)$ for three representative states (see also the signal for each of the 32 coarse states in Figure 6E). In good accord with the experimental design, most of the coil-like states populate long Trp-His distances, while the more helical states cluster at distances close to the cutoff. However, we find that some nonhelical states are significantly "quenched" (see SI), stressing the importance of a detailed simulation of the experimental probes. Comparison of our simulated quantum yield with experiment is complicated by the strong temperature dependence of the tryptophan quantum yield in the absence of histidine, due to the range of other processes which can quench the fluorescence. We correct for this empirically, by scaling our relative quantum yield by the measured quantum yield of suitable reference molecules lacking histidine, measured by Thompson et al.[7] The first reference is N-acetyl-L-tryptophanamide (NATA), and the second is the WAAA control peptide (see Figure 5B). We find that the
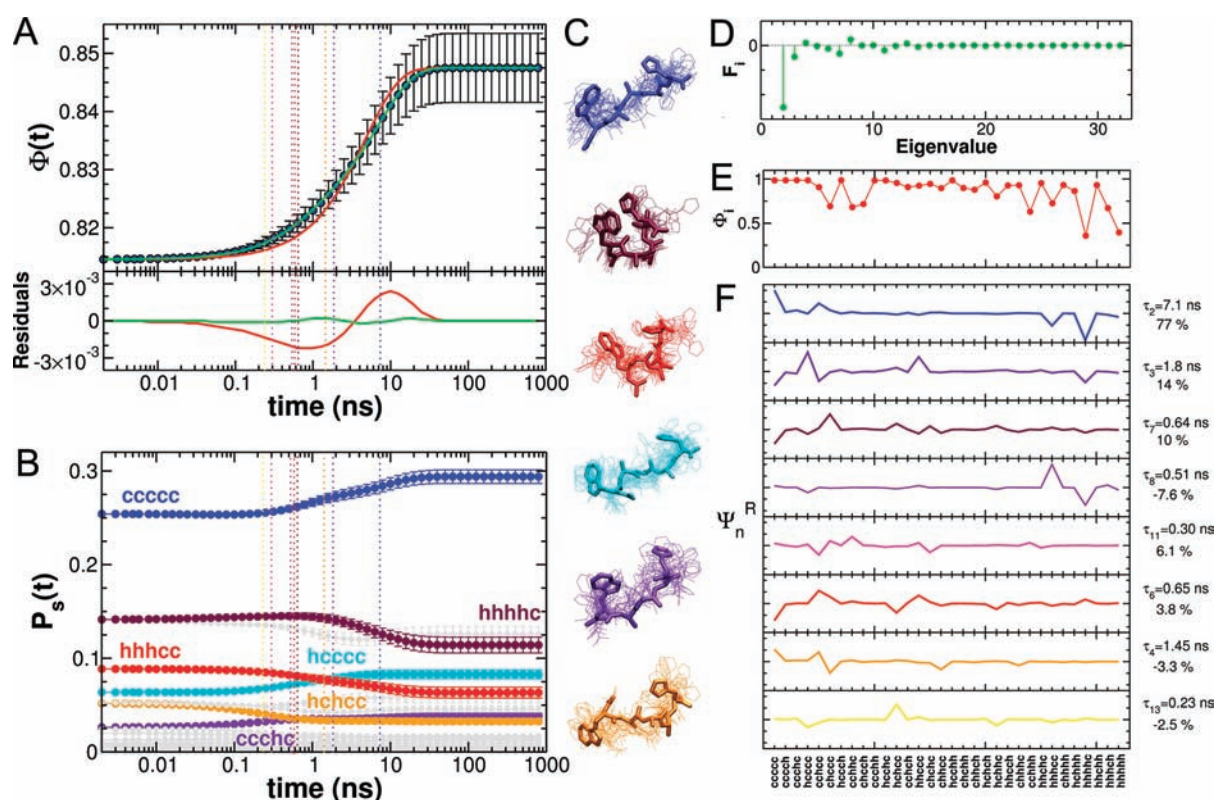
Figure 6. Simulation of the fluorescence quenching temperature jump experiments. (A) Top: change in the calculated fluorescence quantum yield from jumps with T_1 = 290 K and T_2 = 323 K. Red and green lines show fits to single and double exponentials. Bottom: residuals from fits. (B) Contribution to the calculated fluorescence signal for all the states in the Markov state model. In color we show some states whose contributions are especially important for the increase and decrease of the global signal. Dashed vertical lines in (A) and (B) mark the time scales corresponding to the modes that contribute with largest amplitudes to the signal. (C) States whose population changes most significantly upon temperature jumps, with colors as highlighted in (B). (D) Amplitude of the signal contributed by each of the eigenmodes. (E) Signal vector with the relative quantum yield of each of the states calculated using a simple cutoff. (F) Right eigenvectors of the eigenmodes contributing most to the signal change. Colors correspond to those of the dashed lines in (A) and (B).

calculated fluorescence quantum yield from simulations compares extremely well with the experimental measurements for WAAAH[+],[7] particularly when the more realistic WAAA peptide is used as reference.

Simulated Temperature Jumps. We can use the optimized rate matrices and the calculated fluorescence vector to simulate temperature jumps as in the experiments on W_1H_5.[1] Here, we exploit the fact that, from the REMD simulations, we know both equilibrium populations and a master equation describing the dynamics at each temperature. For a temperature jump with initial and final temperatures T_1 and T_2, respectively, the population of each state at time t can be obtained from the matrix exponential:

$$\mathbf{P}(t) = \exp(\mathbf{K}(T_2)t)\mathbf{P}_{eq}(T_1)$$

$$= \sum_{n=1}^{N} (\Psi_n^R[\Psi_n^L \cdot \mathbf{P}_{eq}(T_1)] \, e^{\lambda_n t}) \qquad (1)$$

Here, Ψ_n^R, Ψ_n^L, and λ_n are the right and left eigenvectors and eigenvalues, respectively, of $\mathbf{K}(T_2)$. The simulated fluorescence signal is then obtained straightforwardly as $F(t) = \mathbf{\Phi} \cdot \mathbf{P}(t)$; in the following we use this relative quantum yield without scaling it using the intrinsic fluorescence of the WAAA peptide, since that would not alter the observed rates or relative amplitudes. We simulate temperature jumps for three final temperatures very

Table 1. Double-Exponential Fits to the Simulated Fluorescence Decay[a]

			simulation		experiment	
T_1 (K)	T_2 (K)	A_1 (%)	τ_1 (ns)	τ_2 (ns)	τ_1 (ns)	τ_2 (ns)
270	302	6.9 (2.1)	0.32 (0.09)	21.0 (3.7)	0.85	5.3
280	312	26 (13)	1.3 (1.2)	5.2 (1.1)	0.65	4.7
290	323	21 (8)	0.76 (0.28)	7.2 (0.8)	0.45	3.6

[a] Data correspond to ∼30 K jumps to the final temperatures (T_2) shown. At each temperature, we give the relaxation times and relative amplitude of the fast phase (A_1) from a double-exponential fit. The errors, corresponding to standard deviations from bootstrap tests, are shown in parentheses. Note that the final temperatures for the experiments were 300 and 310 K, rather than 302 and 312 K, respectively.

close to those in the experiments (see Table 1).[1] We use the maximum likelihood rate matrices at these final temperatures and the calculated quantum yield as described above. To maximize the amplitudes we jump from populations at temperatures ∼30 K lower, instead of the ∼10 K jumps in the experiments (note, however, that comparable results are obtained with smaller temperature jumps, see SI). In all cases we find that the response to the simulated temperature jumps is an increase in the fluorescence corresponding to the decrease in the population

of helical states and an increase in the population of coil states (see Figure 6). The transients compare very well with those obtained in the experiments and can be fitted to double-exponential traces with positive amplitudes (see Table 1). At 312 and 323 K the rates for the two phases are in good agreement with those obtained by Mohammed et al.,[1] while at 302 K the fitted relaxations are slower, most likely related to the master equation being less well determined at these lowest temperatures, as discussed above.

In addition to reproducing the same macroscopic behavior observed in the ultrafast kinetic experiments from the simulated temperature jumps, we can use our model to propose an explanation for the microscopic origin of the observed signals. In Figure 6B we show the changes in the populations upon temperature jumps for each state n in the kinetic model that contributes to the signal as $\Phi_n P_n(t)$. We see that the species whose population changes most upon the temperature jumps are the fully coil state, $ccccc$, and the helices $hhhhc$ and $hhhcc$ (see Figure 6C), consistent with the differences in the equilibrium populations of initial and final states (see Figure 3A). As can be observed, although the global fluorescence can be explained using a double exponential, we see that the origin of the signal is actually multiphasic.

We study this in more detail by looking at the contributions of the individual eigenmodes, whose amplitude can be calculated from the signal vector Φ, the initial population $P_{eq}(T_1)$, and the right and left eigenvectors for the rate matrix $K(T_2)$. We can rewrite the expression for the population of all species at each time t as $P(t) = P_{eq}(T_2) + \sum_{n=2}^{N} \Psi_n^R [\Psi_n^L \cdot P_{eq}(T_1)] e^{\lambda_n t}$. Each term in this sum corresponds to the perturbation from the final population contributed by each mode of the rate matrix. Therefore, we can calculate the actual amplitude contributed by the nth mode as $F_n = [\Phi \cdot \Psi_n^R][\Psi_n^L \cdot P_{eq}(T_1)]$ (see Figure 6D). We note that these calculated amplitudes are analogous to what others have termed "dynamical fingerprints".[66] We see that the largest amplitude corresponds to the slowest mode of the system λ_2 but that many others contribute to the overall signal. This multimodality, however, is not discernible in the macroscopic signal due to the close spacing of time scales for the remaining eigenmodes. For the modes with the largest amplitudes F_n, we show the right eigenvectors Ψ_n^R, reflecting the change in population accounted for by mode n[56] (see Figure 6F). We see that the slower modes explain the observed shrinking of helical states at the C-terminus and an increase in the population of coil states, while the faster modes determine the interconversion of states within the nonhelical "coil" ensemble. Importantly, although the overall signal change is positive, we find that some of the minor modes contribute a negative component to the overall amplitude (i.e., populating more strongly quenched species).

Therefore, from the analysis of our kinetic model, we find that the origin of the observed phases does not appear to be separate collapse and folding of the helix, as originally proposed.[1] Rather, we suggest that these experiments represent the first direct measurement of helix melting and coil dynamics, with time scales which were inaccessible to earlier temperature jump methods.

CONCLUSIONS

Here we show that, using molecular dynamics simulations, we can explain the origin of the macroscopic signals observed in equilibrium and kinetic experiments in microscopic detail. A key feature of our study is the use of a protein energy function optimized for peptide folding,[51] together with a model for water that accurately reproduces its density and viscosity.[55] By careful analysis of the resulting dynamics and explicit modeling of the experimental probe, we obtain results that are in excellent agreement with equilibrium and kinetic data for the W_1H_5 pentapeptide from different sources[1,7] and reveal the hidden complexity in the dynamics of this simple molecule. The fast phase in our model appears to arise from equilibration between different states in the coil ensemble, while the slow phase comes from C-terminal melting of helical states. We also relate our results to kinetic parameters commonly used in "zipper"-type models for helix–coil kinetics, finding reasonable agreement with those obtained experimentally. From our kinetic model we are able to determine directly the time scales of helix nucleation and elongation as 20–70 ns and ~1 ns, respectively, in good accord with previous experimental estimates. In the future it will obviously be important to extend our study to longer helices, in particular the 21-residue helices based on the sequence Ac-AAAAA-$(AAARA)_3$A-NH_2, which have been used in very different experimental approaches to determine helix formation kinetics[7,26] The observed relaxation rates in these systems are generally much slower (100 ns time scale) and will therefore represent a distinct test of our simulation model; using a longer peptide will also mitigate the boundary effects inherent in the present peptide model.

ASSOCIATED CONTENT

S **Supporting Information.** Details on the analysis of the simulations as well as additional discussion. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
rbb24@cam.ac.uk

ACKNOWLEDGMENT

REFERENCES

(1) Mohammed, O.; Jas, G.; Lin, M.; Zewail, A. *Angew. Chem. Int. Ed.* **2009**, *48*, 1521–3773.

(2) Eaton, W. A.; Muñoz, V.; Thompson, P. A.; Henry, E. R.; Hofrichter, J. *Acc. Chem. Res.* **1998**, *31*, 745–753.

(3) Eaton, W. A.; Muñoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R.; Hofrichter, J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327–359.

(4) Ferguson, N.; Fersht, A. R. *Curr. Opin. Struct. Biol.* **2003**, *13*, 75–81.

(5) Muñoz, V.; Thompson, P.; Hofrichter, J.; Eaton, W. *Nature* **1997**, *390*, 196–199.

(6) Dyer, R. B.; Gai, F.; Woodruff, W. H. *Acc. Chem. Res.* **1998**, *31*, 709–716.

(7) Thompson, P. A.; Muñoz, V.; Jas, G. S.; Henry, E. R.; Eaton, W. A.; Hofrichter, J. *J. Phys. Chem. B* **2000**, *104*, 378–389.

(8) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U. S.A.* **2000**, *97*, 7220–7225.

(9) Möglich, A.; Joder, K.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12394–12399.

(10) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76–88.

(11) Muñoz, V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11311–11316.

(12) Alm, E.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11305–11310.

(13) Du, D.; Zhu, Y.; Huang, C.-Y.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15915–15920.

(14) Olsen, K. A.; Fesinmeyer, R. M.; Stewart, J. M.; Andersen, N. H. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15483–15487.

(15) Fierz, B.; Satzger, H.; Root, C.; Gilch, P.; Zinth, W.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2163–2168.

(16) Yeh, I.-C.; Hummer, G. *J. Am. Chem. Soc.* **2002**, *124*, 6563–6568.

(17) Clarke, D. T.; Doig, A. J.; Stapley, B. J.; Jones, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7232–7237.

(18) Williams, S.; Causgrove, T.; Gilmanshin, R.; Fang, K.; Callender, R.; Woodruff, W.; Dyer, R. *Biochemistry* **1996**, *35*, 691–697.

(19) Lednev, I.; Karnoup, A.; Sparrow, M.; Asher, S. *J. Am. Chem. Soc.* **1999**, *121*, 8074–8086.

(20) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. *Biochemistry* **2000**, *104*, 9200–9210.

(21) Huang, C.-Y.; Klemke, J. W.; Getahun, Z.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2001**, *123*, 9235–9238.

(22) Huang, C.-Y.; Getahun, Z.; Zhu, Y.; Klemke, J. W.; DeGrado, W. F.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2788–2793.

(23) Werner, J. H.; Dyer, R. B.; Fesinmeyer, R. M.; Andersen, N. H. *J. Phys. Chem. B* **2002**, *106*, 487–494.

(24) Wang, T.; Zhu, Y.; Getahun, Z.; Du, D.; Huang, C.-Y.; DeGrado, W. F.; Gai, F. *J. Phys. Chem. B* **2004**, *108*, 15301–15310.

(25) Bredenbeck, J.; Helbing, J.; Kumita, J. R.; Woolley, G. A.; Hamm, P. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2379–2384.

(26) Fierz, B.; Reiner, A.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 1057–1062.

(27) Brooks, C. L. *J. Phys. Chem.* **1996**, *100*, 2546–2549.

(28) Ferrara, P.; Apostolakis, J.; Caflisch, A. *J. Phys. Chem. B* **2000**, *104*, 5000–5010.

(29) Hiltpold, A.; Ferrara, P.; Gsponer, J.; Caflisch, A. *J. Phys. Chem. B* **2000**, *104*, 10080–10086.

(30) Hummer, G.; García, A. E.; Garde, S. *Phys. Rev. Lett.* **2000**, *85*, 2637–2640.

(31) Jun, B.; Weaver, D. L. *J. Chem. Phys.* **2000**, *112*, 4394–4401.

(32) Hummer, G.; Garcia, A. E.; Garde, S. *Proteins* **2001**, *42*, 77–84.

(33) Buchete, N.-V.; Straub, J. E. *J. Phys. Chem. B* **2001**, *105*, 6684–6697.

(34) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2782–2787.

(35) Doshi, U. R.; Muñoz, V. *J. Phys. Chem. B* **2004**, *108*, 8497–8506.

(36) Gnanakaran, S.; Garcia, A. E. *Proteins* **2005**, *59*, 773–782.

(37) Paschek, D.; Gnanakaran, S.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6765–6770.

(38) Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005**, *88*, 2472–2493.

(39) Sorin, E. J.; Rhee, Y. M.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2006**, *356*, 248–256.

(40) Dzubiella, J. *J. Am. Chem. Soc.* **2008**, *130*, 14000–14007.

(41) Dzubiella, J. *J. Phys. Chem. B* **2009**, *113*, 16689–16694.

(42) Hegefeld, W. A.; Chen, S.-E.; DeLeon, K. Y.; Kuczera, K.; Jas, G. S. *J. Phys. Chem. A* **2010**, *114*, 12391–12402.

(43) Mukherjee, S.; Chowdhury, P.; Bunagan, M. R.; Gai, F. *J. Phys. Chem. B* **2008**, *112*, 9146–9150.

(44) Ma, H.; Wan, C.; Zewail, A. H. *J. Am. Chem. Soc.* **2006**, *128*, 6338–6340.

(45) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Am. Chem. Soc.* **2004**, *126*, 698–699.

(46) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.

(47) Best, R. B.; Buchete, N.-V.; Hummer, G. *Biophys. J.* **2008**, *95*, L7–L9.

(48) Best, R. B.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

(49) Best, R. B.; Mittal, J. *J. Phys. Chem. B* **2010**, *114*, 8790–8798.

(50) Mittal, J.; Best, R. B. *Biophys. J.* **2010**, *99*, L26–L28.

(51) Best, R. B.; Mittal, J. *J. Phys. Chem. B* **2010**, *114*, 14916–14923.

(52) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950–1958.

(53) Freddolino, P. L.; Harrison, C. B.; Liu, Y.; Schulten, K. *Nature Phys.* **2010**, *751*, 751–758.

(54) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

(55) Abascal, J. L. F.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.

(56) Buchete, N.-V.; Hummer, G. *J. Phys. Chem. B* **2008**, *112*, 6057–6069.

(57) Buchete, N.-V.; Hummer, G. *Phys. Rev. E* **2008**, *77*, 030902.

(58) Swope, W. C.; Pitera, J. W.; Suits, F. *J. Phys. Chem. B* **2004**, *108*, 6571–6581.

(59) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. *J. Chem. Phys.* **2007**, *126*, 155101.

(60) Noe, F.; Horenko, I.; Schutte, C.; Smith, J. C. *J. Chem. Phys.* **2007**, *126*, 155102.

(61) Voelz, V. A.; Singh, V. R.; Wedemeyer, W. J.; Lapidus, L. J.; Pande, V. S. *J. Am. Chem. Soc.* **2010**, *132*, 4702–4709.

(62) Bowman, G. R.; Beauchamp, K. A.; Boxer, G.; Pande, V. S. *J. Chem. Phys.* **2009**, *131*, 124101.

(63) Noe, F.; Schutte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19011–19016.

(64) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. *J. Am. Chem. Soc.* **2010**, *132*, 1526–1528.

(65) Bowman, G. R.; Voelz, V. A.; Pande, V. S. *J. Am. Chem. Soc.* **2011**, *133*, 664–667.

(66) Noe, F.; Doose, S.; Daidone, I.; Lollmann, M.; Sauer, M.; Chodera, J. D.; Smith, J. C. *Proc. Natl. Acad. Sci. U.S.A.*

(67) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(68) Gunsteren, W. F. V.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173–185.

(69) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963–1974.

(70) Sreerama, N.; Woody, R. W.; Methods Enzym., D. *Numerical Computer Methods* **2004**, *383*, 318–351.

(71) Bullheller, B. M.; Rodger, A.; Hirst, J. D. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2020–2035.

(72) Northrup, S. H.; Hynes, J. T. *J. Chem. Phys.* **1978**, *69*, 5246–5260.

(73) Sriraman, S.; Kevrekidis, I. G.; Hummer, G. *J. Phys. Chem. B* **2005**, *109*, 6479–6484.

(74) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526–535.

(75) Schwarz, G. *J. Mol. Biol.* **1965**, *11*, 64–77.

(76) Steiner, R. F.; Kirby, E. P. *J. Phys. Chem.* **1969**, *73*, 4130–4135.