The fast-folding HP35 double mutant has a substantially reduced primary folding free energy barrier

Hongxing Lei,[1,2] Xiaojian Deng,[2] Zhixiang Wang,[2,3] and Yong Duan[2,a)]
[1]*Beijing Institute of Genomics, Chinese Academy of Science, Beijing 100029, People's Republic of China*
[2]*UC Davis Genome Center and Department of Applied Science, University of California at Davis, Davis, California 95616, USA*
[3]*Graduate School, Chinese Academy of Science, Beijing 100029, People's Republic of China*

(Received 18 June 2008; accepted 12 September 2008; published online 16 October 2008)

The LYS24/29NLE double mutant of villin headpiece subdomain (HP35) is the fastest folding protein known so far with a folding time constant of 0.6 μs. In this work, the folding mechanism of the mutant has been investigated by both conventional and replica exchange molecular dynamics (CMD and REMD) simulations with AMBER FF03 force field and a generalized-Born solvation model. Direct comparison to the *ab initio* folding of the wild type HP35 enabled a close examination on the mutational effect on the folding process. The mutant folded to the native state, as demonstrated by the 0.50 Å C_α-root mean square deviation (RMSD) sampled in both CMD and REMD simulations and the high population of the folded conformation compared with the denatured conformations. Consistent with experiments, the significantly reduced primary folding free energy barrier makes the mutant closer to a downhill folder than the wild type HP35 that directly leads to the faster transition and higher melting temperature. However, unlike the proposed downhill folding which envisages a smooth shift between unfolded and folded states without transition barrier, we observed a well-defined folding transition that was consistent with experiments. Further examination of the secondary structures revealed that the two mutated residues have higher intrinsic helical preference that facilitated the formation of both helix III and the intermediate state which contains the folded segment helix II/III. Other factors contributing to the faster folding include the more favorable electrostatic interactions in the transition state with the removal of the charged NH_3^+ groups from LYS. In addition, both transition state ensemble and denatured state ensemble are shifted in the mutant. © *2008 American Institute of Physics*.
[DOI: 10.1063/1.2995987]

INTRODUCTION

The native structures of protein are the outcome of the delicate balance among enthalpic and entropic forces. This balance can be disrupted by the change in the environment such as temperature, solvent, and salt concentration and cause many neurodegenerate diseases.[1] To a lesser extent, mutations may affect the stability and folding process of proteins. The effect of point mutations depends on both the nature and position of the mutations. Generally, conservative mutations especially on surface residues have minor effect on protein structures, but nonconservative mutations on deeply buried residues can have profound effect. Mutational effects can be measured in experiments including Φ-value analyses which enable a systematic assessment on the mutational effects to protein stability and folding kinetics.[2] However, exactly how the mutations disrupt the energy balance is usually not well understood due to the difficulty in evaluating the energy contributions from experimental data.

Owing to the ability to dissect energy contributions and the high spatial and temporal resolutions, computer simulations have been applied to the investigations of mutational effects. Yet, because of the enormous challenge both in reaching the native state directly by computer simulations from fully denatured state and in sampling the transition-state ensemble (TSE), most of the studies have focused on protein stability whereby only the experimental native state and the denatured-state ensembles (DSE) (typically represented as random coils) are needed. Using free energy calculation, Hendsch *et al.*[3] predicted the stabilization effect of replacing buried polar groups with nonpolar residues which was verified in the circular dichroism experiment. Veenstra and Kollman[4] applied free energy calculation to the mutations of T4 lysozyme and achieved good agreement with experiments in both stabilizing and destabilizing mutants. In the study of Zoete and Meuwly,[5] the side chains of amino acids were treated as "pseudoligand" and the molecular mechanics and generalized born/surface area (MM-GBSA) approach was applied to evaluate the contribution to protein stability, demonstrating promising results on monomer insulin. Based on molecular dynamics (MD) simulation, Santos *et al.*[6] proposed that the changed stability of two mutants of β-lactamase was due to the altered van der Waals and hydrogen bonding network, respectively. However, because of the lack of sampling to either the transition state or the folding

a)Author to whom correspondence should be addressed. Electronic mail: duan@ucdavis.edu. FAX: (530)-754-9658. Tel.: (530)-754-5625.

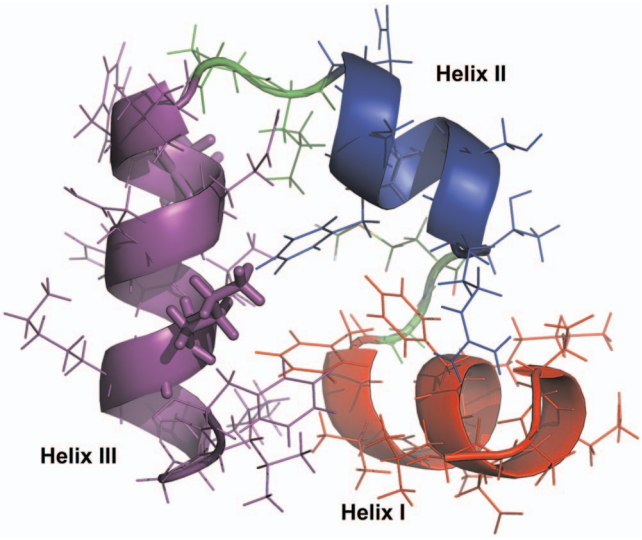

FIG. 1. (Color) The structure of the HP35 double mutant determined by x-ray crystallography (PDB code 2F4K). The backbone is shown as cartoon and the mutated residues (LYS24NLE and LYS29NLE) are shown as sticks. In this view, helices II (in blue) and III (in magenta) are in the plane and helix I (in red) is nearly perpendicular to the plane.

"pathways," these studies were limited to protein stability and were unable to provide information on the folding kinetics.

Complementary to the stability study by free energy calculations, unfolding simulations have been exploited to probe TSE and DSE.[7,8] Examples of the work included the studies of Pan and Daggett[9] who applied free energy perturbation to the hydrophobic deletion of chymotrypsin inhibitor 2 using transition-state and denatured-state structures from unfolding simulations and reached agreement with the Φ-value analysis experiment. Cregut and Serrano[10] employed unfolding simulations to study the mutational effects on the B1 domain of protein G, which agreed with experiments. Eleftheriou *et al.*[11] also applied thermal unfolding to the W62G mutant of lysozyme and found that the mutant has a lower stability compared with the wild type. While the results from these studies are encouraging, they did not directly sample the folding pathways and folding processes.

Villin headpiece subdomain (HP35) is a small and fast-folding protein which attracted many experimental and simulation studies. The structure of this 35-residue helical protein has been determined by both nuclear magnetic resonance (NMR) and x-ray crystallography.[12,13] The folding of HP35 has been investigated by experiments using unfolding, truncation, T-jump, and solid-state NMR techniques.[14–17] The folding time of HP35 was determined to be around 4.3 μs from T-jump experiment.[17] HP35 was considered as a two-state folder but this notion has been challenged by a solid-state NMR experiment and by T-jump kinetic experiment in which a two-stage unfolding was observed.[16,17] In an effort to further examine the "speed limit" of protein folding, Kubelka *et al.* mutated the two partially buried LYS residues (LYS24 and 29, Fig. 1) to nonpolar residues NLE (nonstandard amino acid, see Discussion). The mutant can fold with a time constant of 0.6 μs which is very close to the theoretical limit.[18]

Despite the small size and fast folding, it has been difficult to fold HP35 to its native state. In earlier works, the folding of HP35 has been attempted by 1.0 μs simulations with explicit solvent and by folding@home with short trajectories by both implicit and explicit solvent utilizing enormous computing power.[19–21] Attempts have been made by numerous groups using moderate computing power,[22–29] including the first attempt to fold HP35 using simulation with implicit solvent model by Shen and Freed[22] that reached an C_α-RMSD of 3.0 Å, coarse-grained simulation by De Mori *et al.*,[27] Go-model simulation by Faccioli *et al.*,[29] and basin-hopping simulation by Derreumaux[28] and Carr and Wales.[26] Jayachandran *et al.*[21] were able to fold residues 9–32 (corresponding to the helix II/III segment) to 1.7 Å. Their recent simulations on the HP35 double mutant[30] with explicit solvent demonstrated diverse folding pathways and highlighted the complexity of the free energy landscape.

In our recent work, we applied AMBER FF03 force field and a generalized-Born (GB) model to the folding of HP35 using both conventional MD (CMD) with long trajectories (1 μs) and replica exchange MD (REMD).[31,32] The protein folded to the native state as shown by the lowest C_α-RMSD <0.5 Å in both CMD and REMD simulations. Based on the simulations, we proposed a folding pathway in which the protein goes through a folding intermediate state with well-folded helix II/III segment followed by docking of helix I. We also identified the primary folding free energy barrier which was the formation of the intermediate state. In this study, we are motivated to understand the underlying mechanism of the observed mutational effects of LYS24 and LYS29 on NLE. We apply the same strategy of using both CMD and REMD to obtain both the time-dependent kinetic information and temperature-dependent thermodynamic information. The *ab initio* folding of both wild type HP35 and the mutant allowed a close examination of the mutational effect.

RESULTS

In this section, we will first describe folding of the HP35 mutant to the native state. Following that, we will compare the folding of the HP35 mutant with that of the wild type. In the last part, we will discuss the underlying physical effects of the mutations.

Ab initio folding of the HP35 mutant

We conducted ten simulations (1 μs each, all started from the extended chain) with CMD on the HP35 double mutant. In the simulations, the HP35 mutant sampled conformations with C_α-RMSD <3.0 Å in all ten trajectories and C_α-RMSD <2.0 Å in eight trajectories. When C_α-RMSD <2.0 Å was used as cutoff, the folding time varied from 91 to 792 ns. Five of the trajectories reached the folded state within 500 ns (91, 146, 279, 281, and 465 ns, respectively), and three of the trajectories reached the folded state after 500 ns (530, 569, and 792 ns). Furthermore, folding of C_α-RMSD <1.0 Å was observed in four trajectories and one of them is shown in Fig. 2. In this trajectory, HP35 mutant reached the folded state at around 91 ns and stayed in the

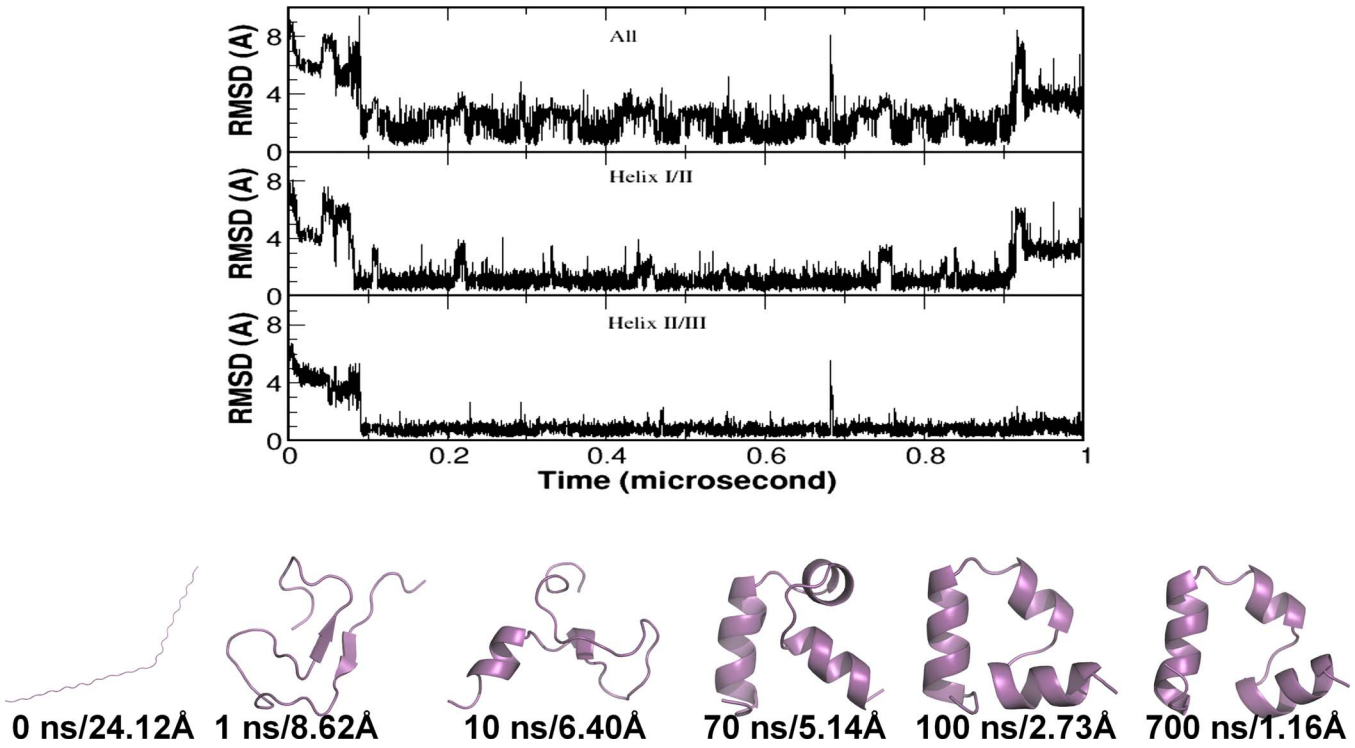

FIG. 2. (Color online) A representative folding trajectory from the CMD simulations of the HP35 mutant. Shown in the three panels are the C_α-RMSD of the whole protein (top, residues 2–34), the helix I/II segment (middle, residues 3–21) and the helix II/III segment (bottom, residues 15–33). Selected snapshots are presented to demonstrate the complete folding process from the extended chain to the fully folded structure.

folded state for about 800 ns. As shown in the figure, the C_α-RMSD fluctuated between 0.4 and 3.0 Å and spent most of the time within 2.0 Å. The observed fluctuation of C_α-RMSD was primarily due to the movement of helix I, as discussed below.

To demonstrate the key events leading to folding to the native state, selected snapshots are shown in Fig. 2. Starting from an extended chain (0 ns), the HP35 mutant quickly collapsed into a coiled structure with a small local β-hairpin structure (1 ns). After 10 ns, helical turns started to form in the helix III region. The formation of the secondary structures was completed at 70 ns, which was followed by the repacking of the structural elements. This eventually evolved to the folded state with a slight variation in the orientation of helix I (two structures with C_α-RMSD of 2.73 and 1.16 Å, respectively, are shown).

In addition to the ten CMD simulations, we also conducted a set of REMD simulations, all starting from the extended chain, using 16 replicas and 240 ns for each replica. Folding was also observed in the REMD simulations where conformations with C_α-RMSD < 1.0 Å were sampled in eight lower temperature replicas ($T < 353$ K) and conformations with C_α-RMSD < 0.5 Å were sampled in five replicas with $T < 310$ K. For better understanding of the conformational sampling, we conducted clustering analysis on the conformations sampled at 296 K and the results are shown in Fig. 3. The top two clusters had very similar population (18.8% and 18.7%) and one of them was a well-folded conformation (1.62 Å C_α-RMSD for the cluster center). In addition, the third cluster (5.8%) was also a folded cluster (3.05 Å C_α-RMSD for the cluster center). Therefore, a combined population of more than 24.5% of folded conformation was sampled in the REMD simulation. In fact, all of the top ten clusters had well-folded helix II/III segment (the intermediate state), indicating the low population of the denatured state. Ideally, the population of the folded conformations should be greater than 50% at 296 K, which is lower than the melting temperature ($T_m = 360$ K from the experiment). However, this property was not faithfully reproduced in this work, indicating room for further improvement in the simulations, including force field and solvation model.

Nevertheless, the ability to consistently reach the native

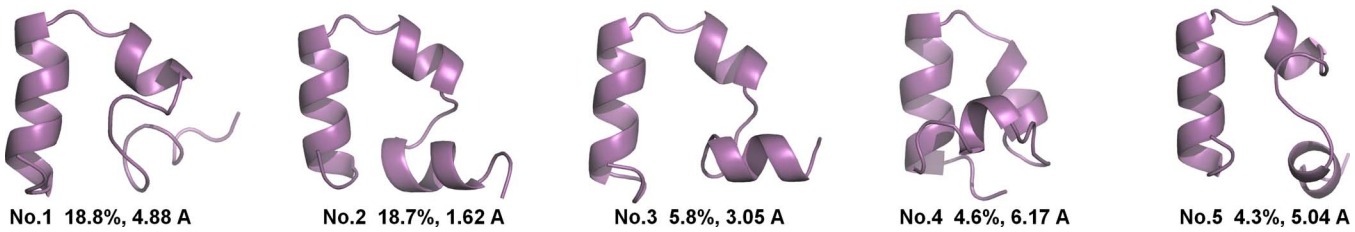

FIG. 3. (Color online) The top 5 clusters (ranked by population) from the REMD simulations of the HP35 mutant (replica $T = 296$ K). Two of them (cluster Nos. 2 and 3) belong to the folded state and the other three belong to the intermediate state.

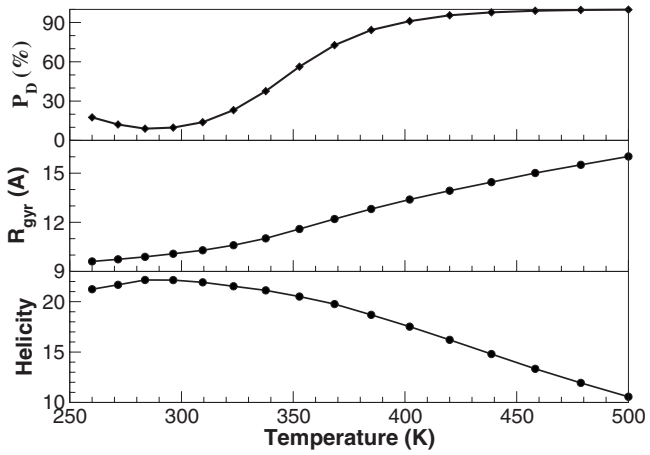

FIG. 4. Temperature dependent folding properties from the REMD simulations of the HP35 mutant. Thermal denaturation is demonstrated by the monotonic increase in the population of the denatured conformations (top), increase in the radius of gyration (middle), and decrease in helicity (bottom).

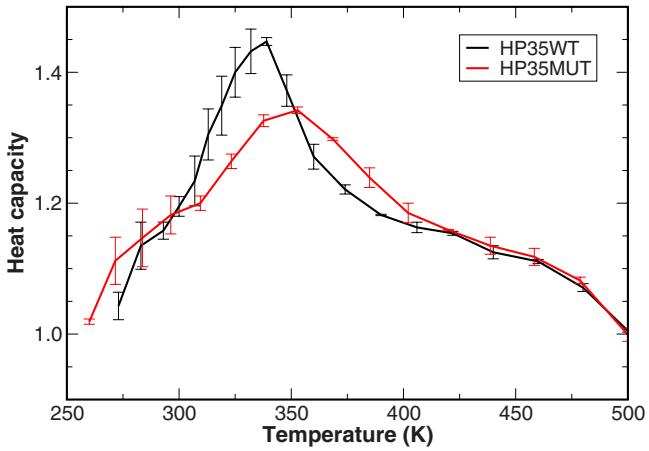

FIG. 5. (Color online) Heat capacity profiles of the HP35 wild type (black line) and mutant (red line) constructed from the REMD simulations. The simulation was divided into two blocks of equal length and heat capacity was calculated on the two blocks separately. The lines correspond to the average value and the error bars correspond to the standard error. The peak of each profile corresponds to the melting temperature T_m. The enhanced thermal stability in the mutant with respect to the wild type is clearly demonstrated by the increased T_m.

state (as defined by C_α-RMSD<2.0 Å or even lower) was an encouraging progress. Further improvement will allow faithful replication of the kinetic and thermodynamic properties. Judging from the fact that most of the so-called "successful" folding simulations done by others failed to reach the native state, our present work is clearly a step forward in which not only the native state was consistently reached but also good agreement was observed between simulation and experiment in terms of certain kinetic and thermodynamic properties.

As temperature increased, the population shifted toward the denatured state gradually. This thermal denaturation of HP35 mutant is shown in Fig. 4. Consistent with the thermal denaturation experiments, there was a sharp increase in the population of the denatured state within the 273–373 K temperature range. In comparison, the increase in the radius of gyration and the decrease in helicity was much slower.

Comparison between the mutant and wild type

In the experimental studies, the mutant has an enhanced thermal stability with the melting temperature of $T_m=360$ K which is higher than the $T_m=342$ K of the wild type, an increase by $\Delta T_m=18$ K. Consistent with the experiments, we observed similar behavior in our simulations. From the heat capacity profiles shown in Fig. 5, the mutant had a $T_m=353$ K and the wild type had a $T_m=339$ K (the peaks of the heat capacity profiles are identified as T_m), an increase by $\Delta T_m=14$ K, in qualitative agreement with experiments.

In addition to the enhanced thermal stability, the mutation also enhanced the folding kinetics. With a folding time of 0.6 μs, the mutant can fold notably faster than the wild type (4.3 μs). Our previous simulations showed that the primary folding transition barrier of the wild type is the formation of helix II/III segment which leads to the folding intermediate. Our present simulations demonstrate that folding of helix II/III segment of the mutant was significantly faster than that of the wild type (Fig. 6). With a cutoff of C_α-RMSD<2.5 Å, the helix II/III of the HP35 mutant reached 80% folded near 150 ns, while that of the wild type only reached 50% folded near 300 ns. Thus, the mutant can cross the folding transition state and reach the folding intermediate state notably faster than the wild type.

A key difference is noted in the formation of helix III which facilitates formation of helix II/III segment. Formation of native hydrogen bonds is shown in Fig. 7. It is clear that a key effect of the mutations is the improved formation of helix III while the effect on the other two helices is marginal. This is not surprising because both mutated residues are within helix III. However, the effect went beyond helix III and notably stabilized helix II/III segment.

Reduced primary folding free energy barrier due to the mutations

Protein folding rates are mostly determined by the primary folding free energy barriers. We calculated the free

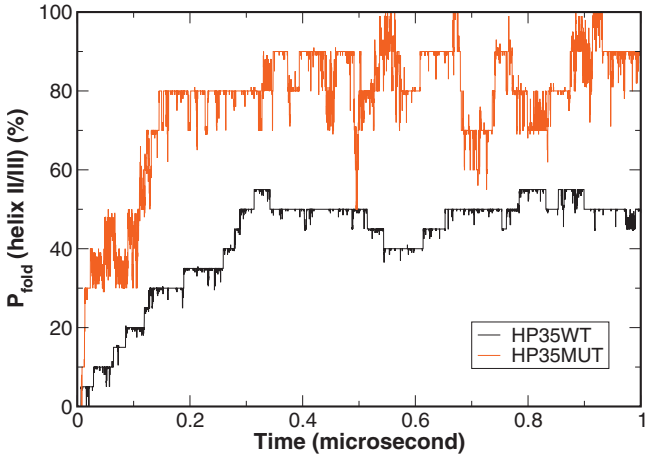

FIG. 6. (Color) Comparison of the helix II/III segment folding for the HP35 wild type (black line) and mutant (red line) from the CMD simulations. The percentage of the folded segment (defined by C_α-RMSD<2.5 Å) was averaged over ten steps. The accelerated transition to the intermediate state in the mutant is clearly shown by the faster formation of the helix II/III segment.

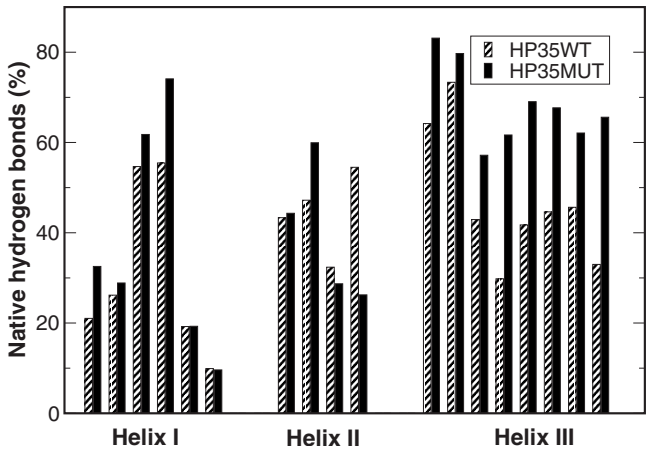

FIG. 7. Native hydrogen bond occupancy from the CMD simulations of the HP35 wild type (gray bar) and mutant (dark bar). Since the first half of the wild type HP35 simulation was excluded due to its slower equilibration process, for consistent comparison, the first half of the simulations (0–0.5 μs) of the mutant was also excluded from the calculation. The accelerated formation of helix III in the mutant is clearly shown by the uniform increase in native hydrogen bond occupancy. Changes in the occupancy of some individual hydrogen bonds were also observed in helices I and II.

energy profiles for the folding of helix II/III segment which was identified as the primary barrier in the wild type folding (Fig. 8). It is rather clear that the free energy barrier (upper panel in Fig. 8) of forming helix II/III segment is much lower for the mutant (~1.0 kcal/mol) than the wild type (~2.8 kcal/mol). To identify the source of this change in free energy barrier, we calculated the potential energy profile for the folding of helix II/III segment. As shown in the lower panel of Fig. 8, the potential energy barrier was ~17 kcal/mol for the wild type and ~9 kcal/mol for the mutant. Therefore, potential energy wise it is much more favorable for the mutant to cross the barrier to form the intermediate state.

Since two charged groups were removed by mutation,

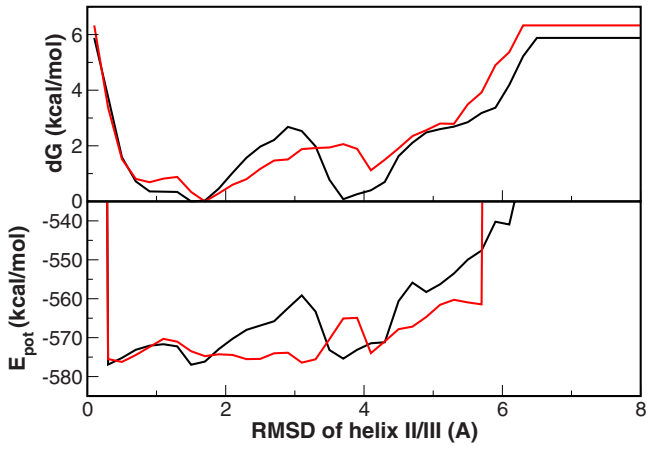

FIG. 8. (Color online) The comparison of energy profiles of the HP35 wild type (black lines) and mutant (red lines) from REMD (replica T=296 K). The folding of the helix II/III segment was used as reaction coordinate to monitor the formation of the intermediate state. Significantly lowered energy barriers (both potential energy and free energy) from the denatured state to the intermediate state in the mutant are clearly shown. For the better comparison of the energy barrier, the potential energy curve of the wild type is shifted upward by 63 kcal/mol.

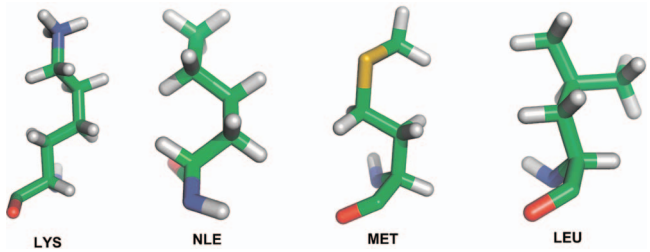

FIG. 9. (Color) Structures of amino acids relevant to the LYS→NLE mutation, including LYS, NLE, MET, and LEU. This is to examine the source of the mutational effects.

there are at least two possible scenarios for the decrease in the potential energy barrier: the mutant may either have less charge repulsion at the transition state or may have less favorable electrostatic interactions in the DSE in comparison to the wild type. Interestingly, the relative energy differences between the denatured state and intermediate state were similar for both the wild type and the mutant. However, the energy of the transition state decreased significantly in the mutant. What is more interesting is that the change also takes place in the positions of the denatured and transition states and both shifted toward higher RMSD for the mutant. This further underscores the notion that folding kinetics can be greatly influenced by a few mutational perturbations when these mutations are targeted at the key sites responsible for the stabilization of the TSE as have been revealed by the Φ-value analysis experiments.

DISCUSSION

In a recent simulation work on the same HP35 mutant, Ensign *et al.*[30] investigated the folding rate of the mutant using the same force field (AMBER FF03) but with explicit solvent. Starting from various unfolded structures generated by unfolding simulation at 373 K, the authors observed heterogeneous folding behavior with various folding rates. The ones with nativelike structures in the helix I and three core phenylalanine residues displayed the closest folding rate to the experimentally determined folding rate. Based on this observation, the authors suggested that those nativelike structures may be close to the unfolded structures in the experiment. Here in our work, we started all simulations from the fully unfolded structures (the straight chain) that enabled us to observe the complete folding process without the potential bias due to selection of starting coordinates. Furthermore, since we obtained the *ab initio* folding of both the wild type and the mutant using the identical simulation parameter set, we were able to compare the difference and investigate the mutational effect. Evidently, the explicit representation of solvent water is ideally suitable to explore the influence of solvent at the molecular level. Therefore, the study of Pande *et al.* and our current work are complimentary to each other with different emphases.

The mutation of LYS to NLE neutralizes the charged side chain by replacing the NH_3^+ group with an H (Fig. 9). The neutralized side chains are subject to a reduced penalty when they are buried in the mutant.[18] In addition, the mutation may also change the local secondary structure prefer-

ence as illustrated by the helical propensity. Since NLE is a nonstandard amino acid, we will infer its helical propensity from similar nonpolar amino acids. Chemical composition wise, NLE is similar to LEU. Structure wise, NLE is similar to MET. We can consider NLE as a linear version of LEU or an $S \rightarrow CH_2$ mutant of MET. Both LEU and MET have higher helical propensity than LYS.[33] Therefore we expect the helical propensity of NLE higher than that of LYS. Indeed, REMD simulations on the Ace-Gly-Lys-Gly-Nme and Ace-Gly-Nle-Gly-Nme showed that NLE peptide has a slightly higher helicity than the LYS peptide by ~3%. This should partially account for the better formation of helix III in the mutant. When the neighboring effects are considered, slightly higher helicity in helix III was predicted by AGADIR (Ref. 34) when LYS is replaced by either LEU or MET (data not shown). This better formed helix III can facilitate the folding of the intermediate state and significantly enhance the folding rate of HP35.

Protein folding rate is largely determined by the primary free energy barriers. The change in the denatured state and transition state can both affect the folding rate in a significant manner. These effects are difficult to predict from the native structure alone. Rather, a detailed knowledge in the folding pathway is necessary to achieve reliable prediction of the mutational effect. From the free energy profile of the formation of the intermediate state (Fig. 8), it is quite clear that both TSE and DSE shifted to larger RMSD upon mutation. Further analyses of the TSE show that the mutation caused substantial changes, although it is difficult to attribute the free energy change to specific residue-residue interactions. In fact, the charge to neutral mutations ($NH_3^+ \rightarrow H$) in the two lysine residues caused substantial structural rearrangement in both the DSE and TSE.

Downhill folding has been a subject of great interest in the field. In the downhill folding process, the protein shifts between the denatured and folded states without notable thermodynamic barriers.[35–37] For small barriers (barrier $< 3RT$), the folding can be an incipient downhill process, as shown by lambda repressor and WW domain.[32] In the cases of downhill or incipient downhill folding, the process exhibits the characteristics of nonexponential and noncooperative behavior. In our study, we observed a significantly reduced folding free energy barrier. After the double mutation, the folding free energy barrier was reduced to ~1.0 kcal/mole, notably smaller than the wild type and comparable to RT at room temperature ($RT = \sim 0.6$ kcal/mole). Thus, the double mutant may be an example of incipient downhill folding. Interestingly, we observed a cooperative folding process as shown by the heat capacity (Fig. 5), folding free energy profiles (Fig. 8), and other measures (Fig. 4). These observations are consistent with experimental results.[18] Thus, the double mutant, although can fold extremely rapidly with a time constant of 0.6 μs, appears to differ from the "downhill" folder which supposes to fold nonexponentially and noncooperatively as envisaged by the funnel theory.

In contrast to the prevailing side-chain/backbone paradigm, Rose *et al.* proposed that protein folding starts from the building of a set of conformations driven by backbone hydrogen bonding which forms local helices and β-sheets.[38] In a recent work by St-Pierre *et al.*,[39] the complex folding scenario was attributed to the formation of a set of conformers before reaching the intermediate state. In the current work, although transient formation of β-sheets was observed in several trajectories such as displayed in Fig. 2, the formation of helical conformations was apparent even before reaching the intermediate state (data not shown). The frustration mostly came from two sources: the mispacking of folded helical segments before the formation of the intermediate state and the slow conformational search in helix I after the formation of the intermediate state.

METHODS

The simulations were conducted with the AMBER simulation package.[40] The all-atom point-charge force field FF03 was chosen to represent the protein.[41] The combined GB (Ref. 42) and surface area model was chosen to mimic solvation effect (surface tension=0.005 kcal/mol Å^2). The partial charges for the nonstandard amino acid norleucine were derived using an approach similar to that in FF03. The geometries of the two representative conformations (alpha and beta) were optimized at HF/6-31G* and the electrostatic potentials for charge fitting were then calculated at B3LYP/cc-pVTZ‖HF/6-31G* in a medium (ether) mimicking the protein interior environment. The polarizable continuum model was used to consider the solvent effects.

Starting from the extended polypeptide chain of the HP35 mutant, short minimization (1000 steps) and equilibration (20 ps with random seed) were applied to the system. The equilibrated structures were the starting point for the simulation trajectories. There were 16 replicas in the REMD simulations[43] and the target temperatures were 260.0, 271.6, 283.7, 296.3, 309.5, 323.3, 337.7, 352.8, 368.5, 384.9, 402.1, 420.0, 438.7, 458.3, 478.7, and 500.0 K. Temperature exchanges were attempted every 2000 steps. Temperature was set to 300 K in CMD simulations and was controlled by applying Berendsen's thermostat[44] with a coupling time constant of 2.0 ps. The choice of thermostat has been tested in a separate study on the B domain of protein A,[45] which demonstrated similar ensemble properties and even folding behavior for Berendsen's thermostat and Andersen's thermostat. Ionic strength was set to 0.2 M. The cutoffs for both general nonbonded interaction and GB pairwise summation were set to 12 Å. The time step was 2 fs in CMD and 1 fs in REMD. SHAKE was applied for hydrogen-connected bond constraint.[46] Slow-varying terms were evaluated every four steps. The coordinates were saved every 10 ps in CMD and 2 ps in REMD. The simulations were run on an advanced micro devices (AMD) cluster (4 CPU on each node), and it took ~40 days to complete each 1 μs CMD simulation and ~30 days to complete the 240 ns REMD simulation.

Clustering was conducted on the REMD trajectories at each temperature. The snapshots were clustered using a hierarchical clustering method. Two snapshots are considered as neighbors when their pairwise C_α-RMSD is below 2.5 Å. One residue from both termini was excluded in the clustering and other RMSD calculations due to the high flexibility. The snapshot with the most neighbors was identified as the center

of the cluster. The process was iterated to identify other clusters from the remaining snapshots. Heat capacity was calculated using $C=(\langle E^2\rangle-\langle E\rangle^2)/RT^2$, where E is the potential energy, R is the gas constant, and T is the temperature. A standard criterion for hydrogen bond was used in the analyses where the cutoff for donor-acceptor distance was set to 3.5 Å and the donor-hydrogen-acceptor angle cutoff was set to 120°. Helicity was evaluated using a simple main chain dihedral cutoff: $\Phi=-57\pm40°$ and $\Psi=-47\pm40°$. We used smaller cutoff of ±30° in the previous studies.[47] The larger cutoff used here is more tolerable to the dynamic nature of the structures. The choice of either value did not have significant effect on the estimation of helicity (data not shown).

CONCLUSIONS

To investigate the effect of the LYS24/29NLE mutation on the folding of HP35, we conducted extensive *ab initio* folding simulations, including a set of ten CMD simulations (1 μs each) and a set of REMD simulations with 16 replicas for 240 ns. Consistent with the simulations on the wild type HP35, folding was obtained for the mutant in both CMD and REMD simulations where C_α RMSD of 0.5 Å was reached, and the population of the folded conformation was relatively high (>24.5%). Consistent with the experimental studies on the mutant and wild type, the mutant demonstrated higher thermal stability and faster folding compared with the wild type. The higher intrinsic helicity of NLE compared to LYS likely contributed to the faster formation of helix III. More importantly, the primary free energy barrier was significantly lower in the mutant which leads to the faster folding of the intermediate state with folded helix II/III segment. Furthermore, unexpected changes in both the DSE and TSE were observed. In short, the simple LYS24/29NLE mutation has profound effect on the folding of HP35.

ACKNOWLEDGMENTS

This work was supported by research grants from NIH (Grant Nos. GM64458 and GM67168 to Y.D.). Usage of PYMOL, GRACE, VMD, MATLAB, and RASMOL graphics packages are gratefully acknowledged.

[1] G. Calloni, S. Zoffoli, M. Stefani, C. M. Dobson, and F. Chiti, J. Biol. Chem. **280**, 10607 (2005).
[2] M. Petrovich, A. L. Jonsson, N. Ferguson, V. Daggett, and A. R. Fersht, J. Mol. Biol. **360**, 865 (2006).
[3] Z. S. Hendsch, T. Jonsson, R. T. Sauer, and B. Tidor, Biochemistry **35**, 7621 (1996).
[4] D. L. Veenstra and P. A. Kollman, Protein Eng. **10**, 789 (1997).
[5] V. Zoete and M. Meuwly, J. Comput. Chem. **27**, 1843 (2006).
[6] J. Santos, V. A. Risso, M. P. Sica, and M. R. Ermacora, Biophys. J. **93**, 1707 (2007).
[7] A. R. Fersht and V. Daggett, Cell **108**, 573 (2002).
[8] C. L. Brooks III, Curr. Opin. Struct. Biol. **8**, 222 (1998).
[9] Y. Pan and V. Daggett, Biochemistry **40**, 2723 (2001).
[10] D. Cregut and L. Serrano, Protein Sci. **8**, 271 (1999).
[11] M. Eleftheriou, R. S. Germain, A. K. Royyuru, and R. Zhou, J. Am. Chem. Soc. **128**, 13388 (2006).
[12] C. J. McKnight, P. T. Matsudaira, and P. S. Kim, Nat. Struct. Biol. **4**, 180 (1997).
[13] T. K. Chiu, J. Kubelka, R. Herbst-Irmer, W. A. Eaton, J. Hofrichter, and D. R. Davies, Proc. Natl. Acad. Sci. U.S.A. **102**, 7517 (2005).
[14] B. S. Frank, D. Vardar, D. A. Buckley, and C. J. McKnight, Protein Sci. **11**, 680 (2002).
[15] Y. Tang, D. J. Rigotti, R. Fairman, and D. P. Raleigh, Biochemistry **43**(11), 3264 (2004).
[16] R. H. Havlin and R. Tycko, Proc. Natl. Acad. Sci. U.S.A. **102**, 3284 (2005).
[17] J. Kubelka, W. A. Eaton, and J. Hofrichter, J. Mol. Biol. **329**, 625 (2003).
[18] J. Kubelka, T. K. Chiu, D. R. Davies, W. A. Eaton, and J. Hofrichter, J. Mol. Biol. **359**, 546 (2006).
[19] Y. Duan and P. A. Kollman, Science **282**, 740 (1998).
[20] B. Zagrovic, C. D. Snow, M. R. Shirts, and V. S. Pande, J. Mol. Biol. **323**, 927 (2002).
[21] G. Jayachandran, V. Vishal, A. E. Garcia, and V. S. Pande, J. Struct. Biol. **157**, 491 (2007).
[22] M. Y. Shen and K. F. Freed, Proteins **49**, 439 (2002).
[23] D. R. Ripoll, J. A. Vila, and H. A. Scheraga, J. Mol. Biol. **339**, 915 (2004).
[24] S. Jang, E. Kim, S. Shin, and Y. Pak, J. Am. Chem. Soc. **125**, 14841 (2003).
[25] T. Herges and W. Wenzel, Structure (London) **13**, 661.
[26] J. M. Carr and D. J. Wales, J. Chem. Phys. **123**, 234901 (2005).
[27] G. M. De Mori, G. Colombo, and C. Micheletti, Proteins **58**, 459 (2005).
[28] P. Derreumaux, Phys. Rev. Lett. **85**, 206 (2000).
[29] P. Faccioli, M. Sega, F. Pederiva, and H. Orland, Phys. Rev. Lett. **97**, 108101 (2006).
[30] D. L. Ensign, P. M. Kasson, and V. S. Pande, J. Mol. Biol. **374**, 806 (2007).
[31] H. Lei and Y. Duan, J. Mol. Biol. **370**, 196 (2007).
[32] H. Lei, C. Wu, H. Liu, and Y. Duan, Proc. Natl. Acad. Sci. U.S.A. **104**, 4925 (2007).
[33] V. Munoz and L. Serrano, Proteins **20**, 301 (1994).
[34] V. Munoz and L. Serrano, Nat. Struct. Biol. **1**, 399 (1994).
[35] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins **21**, 167 (1995).
[36] S. S. Cho, P. Weinkam, and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **105**, 118 (2008).
[37] M. Sadqi, D. Fushman, and V. Munoz, Nature (London) **442**, 317 (2006).
[38] G. D. Rose, P. J. Fleming, J. R. Banavar, and A. Maritan, Proc. Natl. Acad. Sci. U.S.A. **103**, 16623 (2006).
[39] J. F. St-Pierre, N. Mousseau, and P. Derreumaux, J. Chem. Phys. **128**, 045101 (2008).
[40] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26**, 1668 (2005).
[41] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. M. Wang, and P. Kollman, J. Comput. Chem. **24**, 1999 (2003).
[42] A. Onufriev, D. Bashford, and D. A. Case, Proteins **55**, 383 (2004).
[43] U. H. E. Hansmann and Y. Okamoto, Phys. Rev. E **56**, 2228 (1997).
[44] H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[45] H. Lei, C. Wu, Z. X. Wang, Y. Zhou, and Y. Duan, J. Chem. Phys. **128**, 235105 (2008).
[46] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[47] W. Zhang, H. X. Lei, S. Chowdhury, and Y. Duan, J. Phys. Chem. B **108**, 7479 (2004).