Theoretical investigation of the photoinitiated folding of HP-36

SOONMIN JANG,[1] NARASIMHA SREERAMA,[2] VIVIAN H.-C. LIAO,[3] S. HSIU-FENG LU,[4] FENG-YIN LI,[5] SEOKMIN SHIN,[6] ROBERT W. WOODY,[2] AND SHENG HSIEN LIN[4]

[1]Department of Applied Chemistry, Sejong University, Seoul 143-747, Korea
[2]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523, USA
[3]Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 106, Taiwan, Republic of China
[4]Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei 106, Taiwan, Republic of China
[5]Department of Chemistry, National Chung Hsing University, Taichung 402, Taiwan, Republic of China
[6]School of Chemistry, Seoul National University, Seoul 151-747, Korea

(RECEIVED February 9, 2006; FINAL REVISION June 22, 2006; ACCEPTED June 24, 2006)

Abstract

A computational model was developed to examine the phototriggered folding of a *caged* protein, a protein modified with an organic photolabile cross-linker. Molecular dynamics simulations of the modified 36-residue fragment of subdomain B of chicken villin head piece with a photolabile linker were performed, starting from both the *caged* and the *uncaged* structures. Construction of a free-energy landscape, based on principal components as well as on radius of gyration versus root-mean-square deviation, and circular dichroism calculations were employed to characterize folding behavior and structures. The folded structures observed in the molecular dynamics trajectories were found to be similar to that of the wild-type protein, in agreement with the published experimental results. The free-energy landscapes of the modified and wild-type proteins have similar topology, suggesting common thermodynamic/kinetic behavior. The existence of small differences in the free-energy surface of the modified protein from that of the native protein, however, indicates subtle differences in the folding behavior.

Keywords: protein structure/folding; computational analysis of protein structure; phototriggering; uncaged protein; circular dichroism

Deciphering the mechanism by which proteins acquire folded three-dimensional structures remains one of the most important challenges in structural biology. The difficulties in understanding the protein folding problem are partially due to the lack of knowledge concerning the following three major aspects of the folding process. Firstly, detailed structural information about the early stages of folding is not readily available, which is crucial in the understanding of how folding is directed along the productive channels toward the native structure (Myers and Oas 2001). Secondly, it is hard to distinguish the off-pathway intermediates from the on-pathway intermediates during folding, since, with few exceptions, current bulk experiments do not discriminate between the two types of intermediates (Shastry et al. 1998). Finally, most results from folding experiments are based on ensemble-average measurements of a property of a given protein, which can include contributions from a variety of dissimilar conformers.

Refinements in the experimental methods for measuring sub-millisecond folding rates have recently been made and applied to several fast-folding systems (Gruebele 1999; Brockwell et al. 2000; Eaton et al. 2000; Ferguson et al. 2001; Qiu et al. 2002; Snow et al. 2002; Zhu et al. 2003). However, it is difficult to determine with certainty whether

Reprint requests to: Feng-Yin Li, Department of Chemistry, National Chung Hsing University, Taichung 402, Taiwan, Republic of China; e-mail: feng64@nchu.edu.tw; fax: 886-4-22862547; or Robert W. Woody, Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523, USA; e-mail: Robert.Woody@colostate.edu; fax: 1-970-4910494.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.062145106.

a protein is moving through multiple barriers in a sequential pathway to its native structure, or through multiple parallel pathways, each with its own rate-determining step. The difficulty arises because the initial state of a denatured protein is relatively ill-defined and the traditional chemical kinetic analysis cannot be used to characterize the denatured state. For the purpose of studying a specific folding pathway, it may be advantageous to remove the ensemble-averaging effect by examining individual protein molecules. However, it has not been possible so far to specify folding pathways with structural detail, even with recent developments in single-molecule and bulk experiments, because of insufficient sampling (Kiefhaber 1995; Yang and Gruebele 2004; Laurence et al. 2005). To re-examine the problems mentioned above in more detail, a clever scheme for the photochemical initiation of protein folding in a nondenaturing environment has been proposed by Hansen et al. (2000) to improve the time resolution of folding experiments. In this approach, a protein is constrained to a conformationally restrained unfolded state by cross-linking the N terminus to a side chain in the middle of the protein, producing a so-called "caged molecule." This cross-linker is designed to be photolabile so that protein folding can be initiated by cleaving the linker (uncaging process) using flash photolysis. This method provides information on the kinetics of folding of a protein from a starting point that is better defined, both temporally and conformationally, than the usual ensemble resulting from dilution of a denaturant.

We have performed molecular dynamics simulations in order to obtain a better understanding of detailed features at the molecular level for such phototriggered folding experiments. The protein molecule investigated by Hansen et al. (2000) in their phototriggered protein-folding experiments is a 36-residue fragment of the subdomain B of chicken villin head piece (HP-36) (Bretscher and Weber 1979; McKnight et al. 1997; Kubelka et al. 2003; Wang et al. 2003; Chiu et al. 2005), which forms the basis of our computational model. There have been several computer simulations of HP-36, which include both unfolding and folding simulations (Duan and Kollman 1998; Duan et al. 1998; Shen and Freed 2002a; Srinivas and Bagchi 2002; Zagrovic et al. 2002a,b; Fernandez et al. 2003; Lin et al. 2003; van der Spoel and Lindahl 2003; De Mori et al. 2004; Mukherjee and Bagchi 2004; Wen et al. 2004). Our simulations model a specific experimental technique for initiating protein folding and provide a computational approach complementary to the experimental study. We have used the generalized Born model with surface area correction (GB/SA) (Still et al. 1990; Tsui and Case 2000), an implicit solvent model that has been developed to approximate the explicit water in molecular dynamics (MD) simulations in an effort to reduce the computational time.

Hansen et al. (2000) have used circular dichroism (CD) spectroscopy to monitor the photoinitiated folding of HP-36 in their experiments. CD spectroscopy is a valuable experimental tool to estimate and/or monitor the secondary structure of proteins, and analytical methods have been developed for this purpose (Yang et al. 1986; Johnson 1988; Sreerama and Woody 2004). In addition to the analytical applications, protein CD spectra have been calculated using three-dimensional structures of proteins from X-ray diffraction (Woody 1968, 1996; Bode and Applequist 1998; Koslowski et al. 2000; Hirst et al. 2003b; Rogers and Hirst 2004; Sreerama and Woody 2004). The combined MD/CD approach (Blauer et al. 1993; Fleischhauer et al. 1994; Hirst and Brooks 1994; Kiefl et al. 2002; Hirst et al. 2003a), where the MD configurations are used in the calculation of CD spectra, provides a means for following the folding or, at least, analyzing the MD trajectories. The free-energy profile (FEP) is a very useful method to examine the conformational sampling in an ensemble of configurations generated by either MD or Monte Carlo simulations and is widely used (Amadei et al. 1993; Guo et al. 1997; Shea et al. 2000; Garcia and Sanbonmatsu 2001; Zhou et al. 2001; Kamiya et al. 2002). In order to characterize the thermodynamic stability and different folding behavior caused by the photolinker, the FEP principal component (PC) space (Garcia and Sanbonmatsu 2001), was employed.

In this report, we have employed both the combined MD/CD approach and the PC analysis to analyze the structures generated by MD simulations.

Results and Discussion

The structure of the caged molecule was derived from the structure of wild-type HP-36 (PDB code 1VII), determined by NMR (McKnight et al. 1997), with the necessary modifications of Hansen et al. (2000). As shown in Figure 1, the caged-HP-36 molecule was constructed by attaching the linker molecule bromoacetyl-3′-(carboxymethoxy)benzoin (BrAc-CMB) to the N terminus by an amide linkage through the CMB carboxyl group and to the Cys12 side chain in the M12C mutant of HP-36 by a thioether linkage of the Cys with the acetylated hydroxyl group. The ab initio–optimized geometry of the linker molecule was used. Attachment of the linker molecule was followed by energy minimization.

The uncaged HP-36 simulations represent a computational model for the folding of the modified HP-36 molecule of Hansen et al. (2000). Although the MD configurations exhibited a wide range of structural fluctuations, both folding and unfolding events were observed during the simulation. The time series of the structural properties calculated for one of the MD trajectories is shown in Figure 2. Here, we have plotted the radius of gyration (RG) and the backbone RMSD of the MD-generated structures as a function of the simulation time.

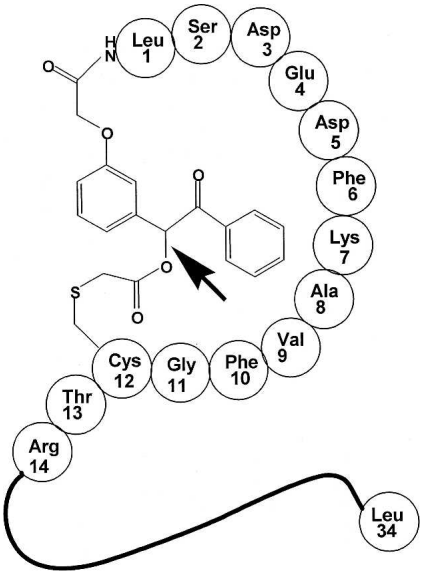

Figure 1. Schematic representation of the caged molecule. The linker molecule, (3′-carboxymethoxy)benzoin, links the N terminus and the side chain of Cys12. The arrow indicates the bond cleaved by photolysis.

Experimentally, uncaged HP-36 folds to a structure similar to that of the wild-type HP-36 (Hansen et al. 2000), and we have used the native structure of wild-type HP-36 determined by NMR as the reference structure in the calculation of backbone RMSD. Folding and unfolding events were characterized based on the similarity of the MD configurations with the folded structure. We have considered the MD-generated structures that have a low RMSD (<4 Å) and low RG as folded. Our definition of the ''folded state'' may be somewhat arbitrary in the absence of a three-dimensional structure for the uncaged protein, but CD spectroscopy suggests that it is similar to that of the wild-type HP-36 (Hansen et al. 2000). Although the variations of RG and RMSD along the MD trajectory, as shown in Figure 2, are quite similar, the lowest RMSD structure may not be the same as the lowest RG structure. However, there is a good correlation between the low RMSD regions and the low RG regions, corresponding to folded structures. The overlapped structure of native HP-36 (red) with a low RMSD (~3.9 Å) uncaged conformation (green) is shown in Figure 3 with the nonstandard linker portion indicated in blue. One can observe that there is a relatively good structural match between these two conformations. The largest differences are in the interhelical strands.

The structural variations observed in uncaged HP-36 simulations, modeling its folding, compare quite well with those for the wild-type HP-36 simulations starting from an extended structure at 400 K. Whereas the excursions of wild type from compact folded structures are short-lived (Fig. 1 in Jang et al. 2003), those for uncaged HP-36 dominate the trajectory, punctuated by short-lived collapses. The RMSD and RG values for a large number of configurations in both simulations are similar (RMSD ~6 Å and RG ~10 Å), the main difference being the larger fluctuations observed in the uncaged simulations and a higher fraction of the unfolded structures. The presence of photoproducts, the polar side chain of the modified cysteine residue and a benzofuran derivative at the N terminus, may have resulted in the larger fluctuations observed in the uncaged MD trajectory.

We have also performed MD simulations of the wild-type HP-36 at 300 K and of the caged molecule at 300 K and 400 K, to compare with that of the uncaged HP-36 results. The overall behavior of wild-type HP-36 at 400 K was similar to that at 300 K, as was the behavior of the caged HP-36 at 400 K in comparison to that at 300 K. We observed slightly larger fluctuations at 400 K. For the wild-type HP-36, the RMSD fluctuated at ~4.2 Å in 300 K and at ~6 Å in 400 K simulations. Analysis of secondary structure content for the caged-molecule simulations at 400 K was performed using Kabsch and Sander's (1983) DSSP program. The average helical content of the caged molecule is ~39%, about 75% of which is located between the C terminus and Cys12. Compared with the helical content of the native HP-36 (53%), the caged molecule at 400 K can be considered

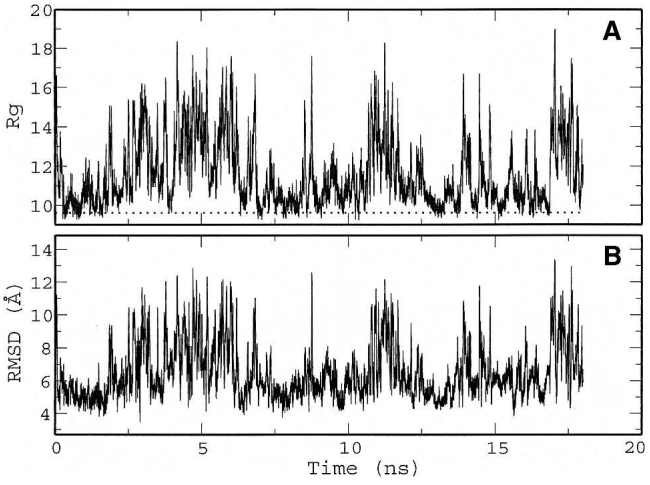

Figure 2. Time series of one of the MD trajectories of uncaged HP-36: radius of gyration (*A*) and backbone RMSD (*B*). The reference structure was the NMR structure of wild-type HP-36 (McKnight et al. 1997; Chiu et al. 2005), and only backbone atoms were used for the RMSD calculation. The dashed line in *A* represents the radius of gyration for the native structure of wild-type HP-36 (RG = 9.4 Å). The dashed line in *B* corresponds to RMSD = 4 Å. The general patterns of variation of the radius of gyration and RMSD with simulation time are quite similar, and reversible folding events (RMSD ~ ≤4 Å) are observed.

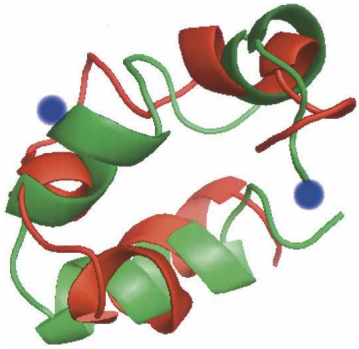

Figure 3. Superimposed images of simulated uncaged (green) and native (red) structure of HP-36. The RMSD is 3.9 Å. The linker parts are shown as blue balls in the uncaged structure for clarity.

partially unfolded, and this agrees with the CD spectrum reported by Hansen et al. (2000). (The 300K trajectory shows similar results.) The CD spectrum reported for the caged molecule (Hansen et al. 2000), while different from that for the native HP-36, also shows evidence for helix content and is far from completely unfolded. However, upon cleavage of the linker, a clear transition from a less helical structure (caged HP-36) to a more helical structure (uncaged HP-36) is observed from CD spectroscopy. Indeed, our MD simulations show similar behavior. Our DSSP analysis with caged and uncaged MD trajectories shows that the helical content is increased to ~47% after cleavage of the linker, corresponding to 89% of native helicity.

We also constructed the free-energy landscapes by using RG and RMSD for both the uncaged and wild-type HP36, as shown in Figure 4. The main wells for the two landscapes are quite similar, which indicates that a significant portion of the uncaged population exists as ''native-like'' conformations. However, the uncaged landscape has a well that is rather elongated toward the upper right, which means that the uncaged molecule also populates relatively unfolded conformers compared to the wild-type structure.

The combined MD/CD approach (Blauer et al. 1993; Fleischhauer et al. 1994; Hirst and Brooks 1994; Kiefl et al. 2002; Hirst et al. 2003a), in which the MD-generated structures are used to calculate the corresponding CD spectra (Sreerama and Woody 2004), was used to further analyze the MD configurations of the wild-type and uncaged HP-36. CD spectra calculated using structures at 375-psec intervals from MD simulations of uncaged and wild-type HP-36 are shown in panels A and B, respectively, of Figure 5. The averages of the calculated CD spectra over the MD trajectories are also shown. Figure 6A, and B, compares the CD spectra calculated using the NMR and X-ray structures of wild-type HP-36 and the time average for the MD trajectories with the experimental CD spectra. The experimental CD spectra correspond to the native HP-36 (Fig. 6A; McKnight et al. 1996) and the folded structure of uncaged HP-36, obtained after photolysis of the cyclized form (Fig. 6B; Hansen et al. 2000).

The native structure of wild-type HP-36, as determined by NMR, has three helical segments, two short segments of less than two turns each and a longer C-terminal segment, totaling 53% helical content. CD spectra, both the experimental and those calculated from the NMR and X-ray structures (McKnight et al. 1997; Chiu et al. 2005), are indicative of significant α-helical content. The CD bands in the calculated spectra are blue-shifted by ~5 nm in comparison with the experiment (McKnight et al. 1996; Hansen et al. 2000). In the rotational strength calculations, standard wavelengths of 220 and 190 nm, respectively, have been used for the nπ* and ππ* transitions. Shifts in the wavelengths for these transitions caused by local electrostatic fields have not been taken

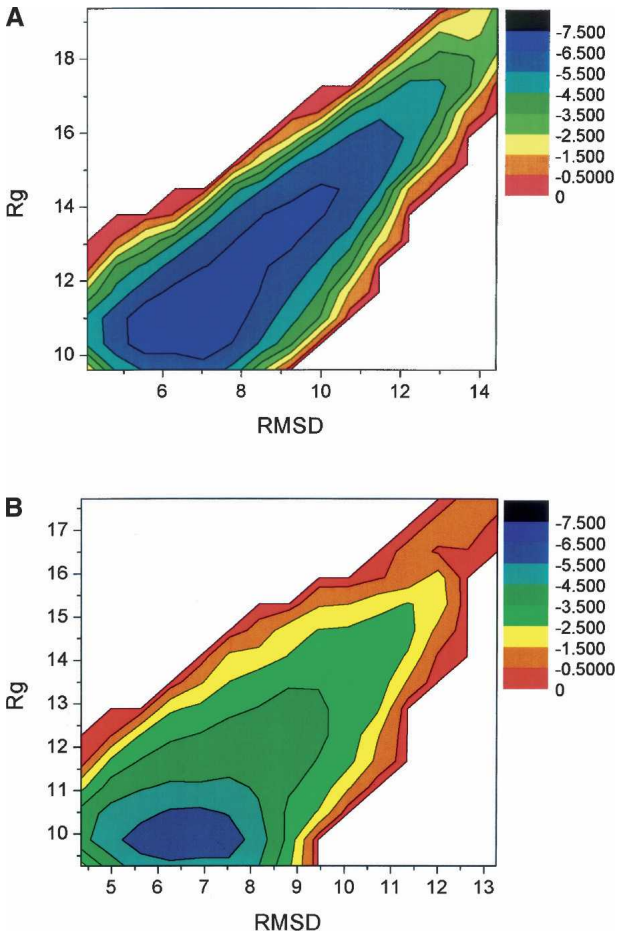

Figure 4. The free energy profiles of the uncaged HP-36 (*A*) and wild-type HP-36 (*B*) at 400 K constructed by using radius of gyration and RMSD. Free energy is in units of kilocalories per mole.

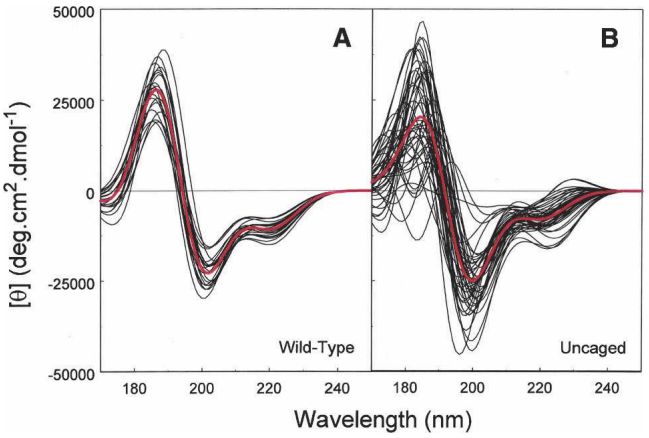

Figure 5. CD spectra calculated from the MD trajectories of wild-type HP-36 (*A*) and uncaged HP-36 (*B*). The thin black curves are the calculated spectra for individual MD configuratons at 375-psec intervals along the trajectories. The thick red curves are the CD spectra averaged over the respective trajectories.

into account. Such shifts are typically of the order of several nanometers. It has been observed previously (Woody 1968) that a red shift of ~6 nm in the position of the $\pi\pi^*$ transition improves agreement of the calculated and experimental band positions for the α-helix, but such an ad hoc shift has not been introduced in the present calculation.

The CD spectra calculated from the MD trajectory of wild-type HP-36 (Fig. 5A) form a family of curves with a characteristic pattern: a short-wavelength positive band followed by two long-wavelength negative bands, similar to that from the NMR structure (Fig. 6) and indicative of significant helical content. The similarity of the calculated CD spectra along the MD trajectory suggests that the simulation is sampling the folded structures, with similar secondary structure compositions. The CD spectra calculated from the MD trajectory of the uncaged molecule (Fig. 5B) show larger variations. While the variations in the CD spectra suggest structural differences, the shape of the majority of the spectra points to the presence of helical structure, but generally less than for the wild-type structure.

The family of CD spectra from the wild-type MD trajectory differs from that calculated from the NMR structure in the amplitudes of the negative bands (Fig. 5A). In the MD-based CD spectra the amplitudes of the 220-nm and the 190-nm bands are smaller and that of the 208-nm band is larger than in the corresponding NMR structure-based spectrum. The uncaged MD-based CD spectra show similar but larger differences from the NMR-based spectrum (Fig. 5B).

The far-UV CD spectrum of a protein reflects its secondary structure (Bode and Applequist 1998; Sreerama and Woody 2004). In order to understand the structural changes along the trajectory, we have performed secondary structure analysis of the MD configurations using the DSSP method (Kabsch and Sander 1983). Unfolding and refolding of the helical segments occurs during the simulation, leading to the formation of 3_{10}- and short α-helical segments along the trajectory. In the MD-generated structures, both the number and length of the α-helical segments vary. In comparison, the NMR structure has one long and two short helical segments. The longer helix breaks in the caged-MD simulation due to the attached linker. In the uncaged MD trajectory, 3_{10}-helical segments are formed frequently at the expense of α-helical segments. However, the existence of the 3_{10}-helical segments in the simulation might be an artifact of the force field and/or the implicit solvation model. For example, even with explicit solvation it has been shown that transitions between the α- and π-helix predicted using the CHARMM22 force field are artifactual (Feig et al. 2003). The Ramachandran plot, as shown in Figure 7, indicates a shift of the (ϕ, ψ) angles toward the 3_{10}-helical region. Both experimental (Toniolo et al. 1996) and theoretical (Manning and Woody 1991) studies have established that in comparison with the α-helix, 3_{10}-helices have a CD spectrum that has diminished amplitudes for the 220-nm and the 190-nm bands and increased amplitude for the 208-nm band. The MD-based CD spectra show similar characteristics, and they are consistent with the secondary structural characteristics of the MD-generated structures.

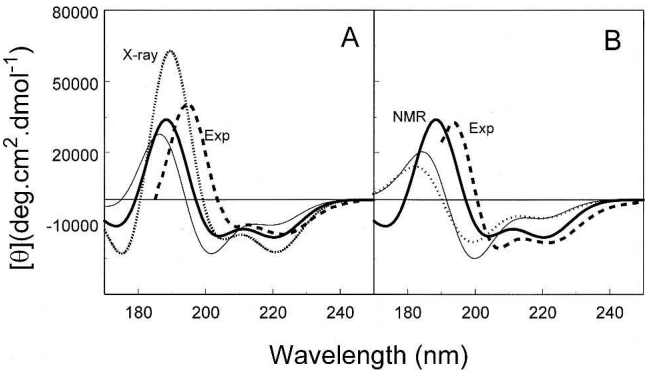

Figure 6. Comparison of CD spectra calculated for NMR and X-ray structures and from the MD trajectories with experiment for wild-type HP-36 (*A*) and uncaged HP-36 (*B*). The thick solid line in each panel is the CD spectrum calculated using the NMR structure of the wild type (McKnight et al. 1997). *A* also shows the spectrum calculated for the X-ray structure of the wild type (Chiu et al. 2005) as a dotted line. The thin solid line in each panel is the time-averaged CD spectrum calculated from the respective MD trajectories. The dotted curve in *B* is the CD spectrum of the uncaged HP-36 averaged over six structures from the trajectory with RMSD < 4 Å. The dashed lines are the experimental CD spectra of wild-type HP-36 (*A*, based on McKnight et al. [1997]) and of uncaged HP-36 after refolding (*B*, based on Hansen et al. [2000]).

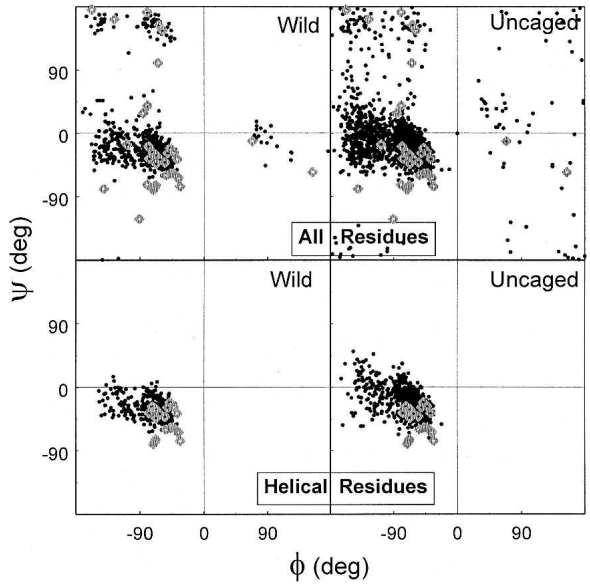

Figure 7. Ramachandran plot of MD configurations selected at equal time intervals of 375 psec along the MD trajectories of the wild-type (*left* panels) and the uncaged (*right* panels) HP-36 (solid circles), compared with that of the NMR structure (McKnight et al. 1997; Chiu et al. 2005) (open diamonds). The φ and ψ values from all residues (*top* panels) and from the residues that form helical structures (*bottom* panels) according to DSSP algorithm (Kabsch and Sander 1983) are plotted. Note that the helical structures from the NMR structure form a tight cluster about the α-helical region of the Ramachandran map, while those from MD simulations are shifted toward the 3_{10} region.

We have also selected structures from the uncaged MD trajectory based on their similarity with the NMR structure, considered folded according to our criterion. The structures that have RMSD < 4 Å from the NMR structure were extracted and the CD spectra were calculated. The calculated CD spectrum averaged over *six* structures from the uncaged HP-36 simulation is compared with the spectrum averaged over the whole trajectory and with that calculated from the NMR structure of the wild type (McKnight et al. 1997) in Figure 6B. The average CD spectrum calculated for the low RMSD structures is similar to that averaged over the entire trajectory in the 222-nm region but agrees somewhat better in amplitude (but not in wavelength) with the experimental CD spectrum and that calculated from the wild-type NMR structure near 208 nm. The differences between the CD spectra calculated from the NMR and X-ray structures and the MD structures are due to the differences in the helical segment characteristics, i.e., the difference between the α- and 3_{10}-helices.

With respect to the PC analysis results, the eigenvalues of the first two PC components (PC1 and PC2) comprise >60% of the total fluctuations in uncaged HP-36 and were used to evaluate the free-energy landscape. The free energy was calculated as $-RT\ \ln(P/P_0)$, where R is the gas constant, T is the temperature, P is the probability of finding the conformation at the given PC values (PC1 and PC2), and P_0 is the normalized probability (Garcia and Sanbonmatsu 2001).

The free-energy landscape calculated from the PC analysis of uncaged HP-36 MD configurations (54,000 conformations, 9 MD trajectories at 400 K) is given in Figure 8. The low RMSD conformations of uncaged HP-36 (Fig. 2) roughly correspond to the low potential energy regions in Figure 4A, corresponding to a folded state. The free-energy landscape of both the uncaged and the wild-type HP-36 from PC analysis have a rather wide stable region at ~ −4 kcal/mol. We have checked the convergence of the free-energy surface by using only the latter half of the trajectories, and we see the consistent feature of two wide wells at the left top and the right bottom of the PC space. A series of free-energy mappings with a different number of trajectories indicates that the error of the free-energy value is <1 kcal/mol. This error range might be not satisfactory for identification of a rigorous folding scenario, since a statistically meaningful number of trajectories (~100) is needed for that purpose (Hunter 2006). However, for the purpose of a combined MD and CD approach, we believe it is not unreasonable.

The wild-type HP-36 free-energy plot (Fig. 5b in Jang et al. 2003) has two well-defined minima connected by a narrow ''bottleneck.'' In contrast, the uncaged HP-36 free-energy plot (Fig. 8) shows two rather poorly defined minima. These observations are consistent with the larger fraction of unfolded structures observed in the uncaged simulations and a tighter clustering of folded structures in wild-type HP-36. Of course, since PC1 and PC2 comprise only ~60% of the total eigenvalues, this effort to

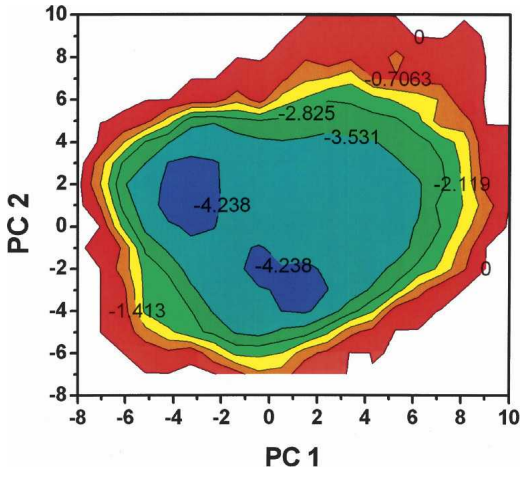

Figure 8. The free-energy profile of the uncaged HP-36 at 400 K calculated from principal component analysis of MD configurations. Free energy is in units of kilocalories per mole.

characterize the free-energy landscape with only two major components may not capture the whole free-energy landscape. This might, in part, explain the subtle free-energy differences.

In spite of some differences, the overall topology of the free energy landscapes of both the wild-type and uncaged HP-36 is similar. This implies that both wild-type and uncaged HP-36 show similar thermodynamic behavior. This is not surprising since it is known that structurally similar proteins, irrespective of their sequence differences, share common features in folding pathways (Ferguson et al. 2001). The differences in the free-energy plots of wild-type and uncaged HP-36 point to differences in the folding mechanism, which might be a result of the chemical modification introduced in the construction of the caged/uncaged HP-36.

The differences between the chemical structures of the uncaged and wild-type HP36 are the Met12 → Cys12 mutation and the attachment of photoproducts at the cysteine side chain and at the N terminus. This modifies Cys12 to carboxymethyl-cysteine. Hansen et al. (2000) suggest, based on CD spectra, that the M12C mutation has little effect on the folding of modified HP-36 because Met12 in wild-type HP-36 is exposed to solvent. According to Frank et al. (2002), the CD spectrum of M12L-HP-36 is also similar to that of wild type, with slightly reduced amplitude. These results suggest that the mutation of Met12 does not significantly alter the native structure. Our combined MD/CD results also suggest similar folded structures for the wild-type and uncaged HP-36, which are in accord with the experimental data. The residue Met12, however, is conserved among the head piece sequences (Vardar et al. 1999). The mutation Met12 → Leu12 reduces the thermal stability of HP-36 and decreases the melting temperature of M12L-HP-36 by ~8°C (Vardar et al. 1999). A reduction in the thermal stability of uncaged HP-36, in comparison with the wild-type HP-36, is expected because of the Met12 → Cys12 mutation used in its construction. The larger structural fluctuations observed in the MD trajectory for the uncaged HP-36 and the observed differences in the free-energy diagrams of the uncaged and wild-type HP36 are consistent with these experimental results.

The combined MD/CD calculations suggest that the uncaged protein shows a folded structure similar to the native structure of the wild-type HP-36 (as illustrated in Fig. 3), and the PC analysis suggests common folding pathways in the uncaged and wild-type HP-36 folding. The implicit solvation model and an elevated temperature of 400 K, used to speed up the conformational sampling/folding, preclude us from estimating the actual folding time in the current simulation. The elevated temperature we have used in MD simulation is somewhat higher than the actual folding temperature, and therefore the population of the unfolded states is higher than the folded state, as is indicated in Figure 2, where the RMSD of the trajectory is generally >4 Å. However, this should not change the general conclusion here, since the overall folding behavior at temperatures not greatly in excess of the folding temperature is not very different from that at and below the folding temperature in "low resolution," unless there are unusual folding paths (Shea and Brooks 2001). (It should be noted that our simulation temperature is substantially lower than that used in the typical unfolding simulation.) As was demonstrated by Jang et al. (2003), this strategy captures the overall picture of the folding event and protein stability. Meanwhile it is also true that, strictly speaking, the detailed free-energy landscape is temperature-dependent and the free-energy surface at physiological temperature is expected to be somewhat rugged (narrower and deeper) than the free-energy shape we presented here. In this sense, the simulation results presented in this article must be considered as reflecting general trends in low resolution rather than an exact picture.

Conclusions

We have introduced a computational model for photoinitiating the folding of a caged protein. Molecular dynamics simulations, combined with CD calculations and PC analysis, were performed to examine the folding dynamics of chemically modified HP-36, starting from its uncaged conformation. The similarity of the calculated CD spectra of the low RMSD configurations with that of the native structure suggests that the "folded" structure of the modified HP-36 is similar to that of the wild-type protein. Although there are subtle differences in the detailed aspects, the overall shape of the free-energy landscapes of chemically modified and wild-type HP-36 are similar, suggesting a common folding pathway. The present study demonstrates that photoinitiated protein folding provides a useful method to investigate folding dynamics starting from a well-characterized initial configuration.

Materials and methods

MD simulations

MD simulations were performed using the AMBER7 package (D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E.I. Cheatham, J. Wang, W.S. Ross, C.L. Simmering, T.A. Darden, K.M. Merz, R.V. Stanton, et al., University of California, San Francisco) with an all-atom AMBER force field, *parm99*. Because the linker portion of the caged molecule is not a standard amino acid residue, its corresponding force field is not available in the *parm99* force field library. We used *gaff* (the generalized AMBER force field) parameters available in the AMBER7

package for the linker molecule. The most important part in constructing a new building block is to develop the charge distribution. We first calculated the charges of the linker segment with Gaussian03 (M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, et al., Gaussian Inc.) at the HF/6–31G* level with geometry optimization in vacuum. The final charges were assigned following the methodology of Eriksson et al. (1999).

In order to facilitate the calculation, the GB/SA implicit-solvation model is used. Although the GB/SA implicit-solvation model is not as accurate as the use of explicit water in MD simulations (Shen and Freed 2002b; Baumketner and Shea 2003; Nymeyer and Garcia 2003), it has been shown that the GB/SA solvation model captures the overall behavior of small peptides in aqueous solution reasonably well (Bursulaya and Brooks 1999). Recent studies have demonstrated that ab initio fast-folding at high temperature using a GB implicit solvent model with an all-atom force field can describe the spontaneous formation of native-like structure for small peptides and proteins (Jang et al. 2002, 2003; Simmerling et al. 2002). The GB parameters of Tsui and Case (2000) were used throughout our simulations. Once the energy-minimized structure was obtained, 2 nsec of MD simulation were performed at 300 K for equilibration. Starting from the initial energy-minimized structure of the caged molecule, we performed several MD simulations of up to 4 nsec at 400 K, and the resulting structure did not show noticeable differences from the 300 K equilibrium run (RMSD < 1.9 Å). This indicates that the initial structure of the caged protein is relatively well defined and the helix structure outside of the loop region is stable up to 400 K.

The ''phototriggered protein'' was constructed by separating the linker from the equilibrated structure of the caged molecule, according to the reaction suggested in the experiment (Hansen et al. 2000). Separation of the linker results in the uncaged HP-36 structure, containing photoproducts resulting from the photolysis of CMB: a phenylbenzofuran derivative attached to the N terminus and a carboxymethyl-cysteine side chain. After the uncaged protein was constructed, we carried out another energy minimization followed by nine MD simulations, each with a different initial velocity distribution, for up to 18 nsec at 400 K. The higher temperature was chosen to reduce the length of MD simulations. The time step was set to 1.5 fsec, and the Berendsen thermostat (Berendsen et al. 1984) was used for temperature control with a coupling constant of 1.0 psec. The bond distances between heavy and hydrogen atoms were fixed using the SHAKE algorithm (Ryckaert et al. 1977). MD configurations were saved at 0.375-psec intervals for further analysis.

CD calculations

CD spectra were calculated for structures at 375-psec intervals from the MD trajectories, for the NMR structure (McKnight et al. 1997), and for the X-ray structure (PDB code 1WY3) (Chiu et al. 2005) of wild-type HP-36. The origin-independent matrix method (Bayley et al. 1969; Goux and Hooker 1980) was employed to compute the rotational strengths and transition energies for a given structure, which are necessary for calculating the CD spectrum. In the CD calculations, the protein molecule is treated as a collection of independent chromophores (peptide groups and aromatic side chains) with specific electronic transitions (three on the peptide group, four on tyrosine and phenylalanine side chains, and six on the tryptophan side chain).[6] The transition parameters were those used by Woody and Sreerama (1999). These parameters combine empirical and theoretical quantities. The transition energies (wavelengths) and electric dipole transition moment magnitudes are based upon experimental data for simple amides and aromatic side-chain chromophores, with the exception of three higher energy transitions in indole, for which theoretical values were used. Transition charge densities were modeled by a distributed dipole or monopole approximation. For the peptide $n\pi^*$ and $\pi\pi^*$ transitions, the monopoles were located as previously described (Woody 1968). For aromatic side chains, the monopole charges for the peptide $\pi\pi^*$ transitions were assigned to reproduce the experimental transition moment directions for N-acetylglycine (Clark 1995). Those for the peptide $n\pi^*$ transition were calculated from INDO/S (Ridley and Zerner 1973) wave functions for Ń-methylacetamide. The monopole charges for the aromatic side-chain $\pi\pi^*$ transitions were generated from π-MO wave functions calculated using the parameters of Nishimoto and Forster (1966). The magnetic dipole transition moment for the peptide $\pi\pi^*$ transitions were taken to be zero, relative to the carbonyl carbon as an origin. The magnetic dipole transition moments of the peptide $n\pi^*$ and the side-chain $\pi\pi^*$ transitions were obtained from the previously mentioned MO calculations. These parameters reproduce the major features of a wide range of globular proteins with an accuracy comparable to that obtained with ab initio parameters (Woody and Sreerama 1999; Hirst et al. 2003b) and work best for proteins with substantial α-helix and little β-sheet, the category to which the villin head piece belongs.

Interaction between all chromophores in a given protein, evaluated in the framework of the matrix method, gives the rotational strengths and energies of transitions of the composite protein molecule. The CD spectrum is calculated by assuming Gaussian bands for individual transitions. The methodology for calculating CD spectra of proteins has been described by Sreerama and Woody (2004).

PC (principal component) analysis

PC analysis, also called covariance analysis or essential dynamics, is a standard mathematical tool used to detect correlations between the dominant features in a given data set with many dimensions. PC analysis defines a new coordinate system for the data set, with the special property that the covariance is zero for any two coordinates. In this sense, these new coordinates can be called uncorrelated. These coordinates are ordered according to the variance of the data in that coordinate. This allows for a reduction of dimensionality of the space by neglecting the coordinates with small variance, thus concentrating on the coordinates with the largest spread of fluctuations. In other words, major fluctuational motions characteristic of a given

[6]In a recent study, Goldmann et al. (2001) have questioned the validity of the independent chromophore model for polypeptides because their calculations showed extensive mixing of the amide $\pi\pi^*$ transitions with interpeptide charge-transfer (CT) configurations. However, the model used by Goldmann et al. predicts that the manifolds formed by the $\pi\pi^*$ configurations and the CT configurations overlap in energy. More reliable ab initio calculations (Serrano-Andrés and Fülscher 1998, 2001) show that there is an energy gap of ~1 eV separating the two types of excited state, and this is supported by experiment (Koslowski and Woody 2002). The larger energy gap strongly suppresses the $\pi\pi^*$-CT mixing and supports the validity of the independent chromophore approximation.

ensemble of conformations begin to emerge through PC analysis. This makes PC values relatively robust parameters that can be used to generate free-energy landscapes. Free-energy profiles can also be generated based on other physical quantities, such as RG, RMSD, number of hydrogen bonds, and number of native contacts (Guo et al. 1997; Shea et al. 2000).

To further analyze the MD configurations and to examine the overall characteristics of thermodynamic behavior, we have calculated the free-energy landscape using PC analysis from uncaged and caged trajectories separately. PC analysis is based on the diagonalization of the covariance matrix built from the positional fluctuations of the backbone atoms in MD trajectories and provides a means for sampling conformations. It was performed using 54,000 configurations (one every 3 psec) from nine uncaged MD trajectories at 400 K with different initial velocity distributions, each 18 nsec long. The PC components provide useful parameters to describe the protein dynamics, in many cases better than the conventional parameters such as RG, RMSD, accessible surface area, etc. (Kamiya et al. 2002).

Acknowledgments

We appreciate helpful discussion with Y. Pak. S.J. is acknowledged for support from Korea Sci. & Eng. Foundation (R08-2003-000-10777-0). S.F.L. is supported by a post-doc fellowship from Academia Sinica, Taiwan. Financial support from an NIH grant (EB002803) to 1R.W.W. is gratefully acknowledged.

References

Amadei, A., Linssen, A.B.M., and Berendsen, H.J.C. 1993. Essential dynamics of proteins. *Proteins* **17:** 412–425.

Baumketner, A. and Shea, J.E. 2003. Kinetics of the coil-to-helix transition on a rough energy landscape. *Phys. Rev. E* **68:** 051901-1–051901-10.

Bayley, P.M., Nielsen, E.B., and Schellman, J.A. 1969. Rotatory properties of molecules containing two peptide groups: Theory. *J. Phys. Chem.* **73:** 228–243.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81:** 3684–3690.

Blauer, G., Sreerama, N., and Woody, R.W. 1993. Optical activity of hemoproteins in the Soret region. Circular dichroism of the heme undecapeptide of cytochrome *c* in aqueous solution. *Biochemistry* **32:** 6674–6679.

Bode, K.A. and Applequist, J. 1998. Globular protein ultraviolet circular dichroic spectra. Calculation from crystal structures via the dipole interaction model. *J. Am. Chem. Soc.* **120:** 10938–10946.

Bretscher, A. and Weber, K. 1979. Villin: The major microfilament-associated protein of the intestinal microvillus. *Proc. Natl. Acad. Sci.* **76:** 2321–2325.

Brockwell, D.J., Smith, D.A., and Radford, S.E. 2000. Protein folding mechanisms: New methods and emerging ideas. *Curr. Opin. Struct. Biol.* **10:** 16–25.

Bursulaya, B.D. and Brooks, C.L. 1999. Folding free energy surface of a three-stranded β-sheet protein. *J. Am. Chem. Soc.* **121:** 9947–9951.

Chiu, T.K., Kubelka, J., Herbst-Irmer, R., Eaton, W.A., Hofrichter, J., and Davies, D.R. 2005. High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl. Acad. Sci.* **102:** 7517–7522.

Clark, L.B. 1995. Polarization assignments in the vacuum UV spectra of the primary amide, carboxyl, and peptide groups. *J. Am. Chem. Soc.* **117:** 2375–2382.

De Mori, G.M.S., Micheletti, C., and Colombo, G. 2004. All-atom folding simulations of the villin headpiece from stochastically selected coarse-grained structures. *J. Phys. Chem. B* **108:** 12267–12270.

Duan, Y. and Kollman, P.A. 1998. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **282:** 740–744.

Duan, Y., Wang, L., and Kollman, P.A. 1998. The early stage of folding of villin headpiece subdomain observed in a 200-nanosecond fully solvated molecular dynamics simulation. *Proc. Natl. Acad. Sci.* **95:** 9897–9902.

Eaton, W.A., Munoz, V., Hagen, S.J., Jas, G.S., Lapidus, L.J., Henry, E.R., and Hofrichter, J. 2000. Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **29:** 327–359.

Eriksson, M.A.L., Morgantini, P.Y., and Kollman, P.A. 1999. Binding of organic cations to a cyclophane host as studied with molecular dynamics simulations and free energy calculations. *J. Phys. Chem. B* **103:** 4474–4480.

Feig, M., MacKerell, A.D., and Brooks, C.L. 2003. Force field influence on the observation of π-helical protein structures in molecular dynamics simulations. *J. Phys. Chem. B* **107:** 2831–2836.

Ferguson, N., Johnson, C.M., Macias, M., Oschkinat, H., and Fersht, A. 2001. Ultrafast folding of WW domains without structured aromatic clusters in the denatured state. *Proc. Natl. Acad. Sci.* **98:** 13002–13007.

Fernandez, A., Shen, M.Y., Colubri, A., Sosnick, T.R., Berry, R.S., and Freed, K.F. 2003. Large-scale context in protein folding: Villin headpiece. *Biochemistry* **42:** 664–671.

Fleischhauer, J., Grotzinger, J., Kramer, B., Kruger, P., Wollmer, A., Woody, R.W., and Zobel, E. 1994. Calculation of the circular-dichroism spectrum of cyclo- (L-Tyr-L-Tyr) based on a molecular-dynamics simulation. *Biophys. Chem.* **49:** 141–152.

Frank, B.S., Vardar, D., Buckley, D.A., and McKnight, C.J. 2002. The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain. *Protein Sci.* **11:** 680–687.

Garcia, A.E. and Sanbonmatsu, K.Y. 2001. Exploring the energy landscape of a β-hairpin in explicit solvent. *Proteins* **42:** 345–354.

Goldmann, E., Asher, S.A., and Mukamel, S. 2001. Electronic excitations of polyalanine: Test of the independent chromophore approximation. *Phys. Chem. Chem. Phys.* **3:** 2893–2903.

Goux, W.J. and Hooker, T.M. 1980. Chiroptical properties of proteins. I. Near-ultraviolet circular dichroism of ribonuclease S. *J. Am. Chem. Soc.* **102:** 7080–7087.

Gruebele, M. 1999. The fast protein folding problem. *Annu. Rev. Phys. Chem.* **50:** 485–516.

Guo, Z.Y., Brooks, C.L., and Boczko, E.M. 1997. Exploring the folding free energy surface of a three-helix bundle protein. *Proc. Natl. Acad. Sci.* **94:** 10161–10166.

Hansen, K.C., Rock, R.S., Larsen, R.W., and Chan, S.I. 2000. A method for photoinitating protein folding in a nondenaturing environment. *J. Am. Chem. Soc.* **122:** 11567–11568.

Hirst, J.D. and Brooks, C.L. 1994. Helicity, circular-dichroism and molecular-dynamics of proteins. *J. Mol. Biol.* **243:** 173–178.

Hirst, J.D., Bhattacharjee, S., and Onufriev, A.V. 2003a. Theoretical studies of time-resolved spectroscopy of protein folding. *Faraday Discuss.* **122:** 253–267.

Hirst, J.D., Colella, K., and Gilbert, A.T.B. 2003b. Electronic circular dichroism of proteins from first-principles calculations. *J. Phys. Chem. B* **107:** 11813–11819.

Hunter, P. 2006. Into the fold. *EMBO Rep.* **7:** 249–252.

Jang, S., Shin, S., and Pak, Y. 2002. Molecular dynamics study of peptides in implicit water: Ab initio folding of β-hairpin, β-sheet, and ββα-motif. *J. Am. Chem. Soc.* **124:** 4976–4977.

Jang, S.M., Kim, E., Shin, S., and Pak, Y. 2003. Ab initio folding of helix bundle proteins using molecular dynamics simulations. *J. Am. Chem. Soc.* **125:** 14841–14846.

Johnson Jr., W.C. 1988. Secondary structure of proteins through circular dichroism spectroscopy. *Annu. Rev. Biophys. Biophys. Chem.* **17:** 145–166.

Kabsch, W. and Sander, C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22:** 2577–2637.

Kamiya, N., Higo, J., and Nakamura, H. 2002. Conformational transition states of a β-hairpin peptide between the ordered and disordered conformations in explicit water. *Protein Sci.* **11:** 2297–2307.

Kiefhaber, T. 1995. Kinetic traps in lysozyme folding. *Proc. Natl. Acad. Sci.* **92:** 9029–9033.

Kiefl, C., Sreerama, N., Haddad, R., Sun, L.S., Jentzen, W., Lu, Y., Qiu, Y., Shelnutt, J.A., and Woody, R.W. 2002. Heme distortions in sperm-whale carbonmonoxy myoglobin: Correlations between rotational strengths and heme distortions in MD-generated structures. *J. Am. Chem. Soc.* **124:** 3385–3394.

Koslowski, A. and Woody, R.W. 2002. Recent developments in the electronic spectroscopy of amides and α-helical polypeptides. *Biophys. Chem.* **101:** 535–551.

Koslowski, A., Sreerama, N., and Woody, R.W. 2000. Theoretical approach to electronic optical activity. In *Circular dichroism: Principles and applications* (eds. N. Berova et al.), pp. 55–96. John Wiley & Sons, Inc, New York.

Kubelka, J., Eaton, W.A., and Hofrichter, J. 2003. Experimental tests of villin subdomain folding simulations. *J. Mol. Biol.* **329:** 625–630.

Laurence, T.A., Kong, X., Jager, M., and Weiss, S. 2005. Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins. *Proc. Natl. Acad. Sci.* **102:** 17348–17353.

Lin, C.Y., Hu, C.K., and Hansmann, U.H.E. 2003. Parallel tempering simulations of HP-36. *Proteins* **52:** 436–445.

Manning, M.C. and Woody, R.W. 1991. Theoretical CD studies of polypeptide helices: Examination of important electronic and geometric factors. *Biopolymers* **31:** 569–586.

McKnight, C.J., Doering, D.S., Matsudaira, P.T., and Kim, P.S. 1996. A thermostable 35-residue subdomain within villin headpiece. *J. Mol. Biol.* **260:** 126–134.

McKnight, C.J., Matsudaira, P.T., and Kim, P.S. 1997. NMR structure of the 35-residue villin headpiece subdomain. *Nat. Struct. Biol.* **4:** 180–184.

Mukherjee, A. and Bagchi, B. 2004. Contact pair dynamics during folding of two small proteins: Chicken villin head piece and the Alzheimer protein β-amyloid. *J. Chem. Phys.* **120:** 1602–1612.

Myers, J.K. and Oas, T.G. 2001. Preorganized secondary structure as an important determinant of fast protein folding. *Nat. Struct. Biol.* **8:** 552–558.

Nishimoto, K. and Forster, L.S. 1966. SCFMO calculations of heteratomic systems with the variable β approximation. I. Hetero-atomic molecules containing nitrogen or oxygen atoms. *Theor. Chim. Acta* **4:** 155–165.

Nymeyer, H. and Garcia, A.E. 2003. Simulation of the folding equilibrium of α-helical peptides: A comparison of the generalized Born approximation with explicit solvent. *Proc. Natl. Acad. Sci.* **100:** 13934–13939.

Qiu, L.L., Pabit, S.A., Roitberg, A.E., and Hagen, S.J. 2002. Smaller and faster: The 20-residue Trp-cage protein folds in 4 μs. *J. Am. Chem. Soc.* **124:** 12952–12953.

Ridley, J. and Zerner, M. 1973. Intermediate neglect of differential overlap technique for spectroscopy: Pyrrole and azines. *Theor. Chim. Acta* **32:** 111–134.

Rogers, D.M. and Hirst, J.D. 2004. Calculations of protein circular dichroism from first principles. *Chirality* **16:** 234–243.

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **23:** 327–341.

Serrano-Andrés, L. and Fülscher, M.P. 1998. Theoretical study of the electronic spectroscopy of amides. III. Charge-transfer transitions in polypeptides. *J. Am. Chem. Soc.* **120:** 10912–10920.

———. 2001. Charge-transfer transitions in neutral and ionic polypeptides. *J. Phys. Chem. B* **105:** 9323–9330.

Shastry, M.C.R., Sauder, J.M., and Roder, H. 1998. Kinetic and structural analysis of submillisecond folding events in cytochrome *c*. *Acc. Chem. Res.* **31:** 717–725.

Shea, J.E. and Brooks, C.L. 2001. From folding theories to folding proteins: A review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* **52:** 499–535.

Shea, J.E., Onuchic, J.N., and Brooks, C.L. 2000. Energetic frustration and the nature of the transition state in protein folding. *J. Chem. Phys.* **113:** 7663–7671.

Shen, M.Y. and Freed, K.F. 2002a. All-atom fast protein folding simulations: The villin headpiece. *Proteins* **49:** 439–445.

———. 2002b. Long-time dynamics of met-enkephalin: Comparison of explicit and implicit solvent models. *Biophys. J.* **82:** 1791–1808.

Simmerling, C., Strockbine, B., and Roitberg, A.E. 2002. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* **124:** 11258–11259.

Snow, C.D., Nguyen, N., Pande, V.S., and Gruebele, M. 2002. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* **420:** 102–106.

Sreerama, N. and Woody, R.W. 2004. Computation and analysis of protein circular dichroism spectra. *Methods Enzymol.* **383:** 318–346.

Srinivas, G. and Bagchi, B. 2002. Foldability and the funnel of HP-36 protein sequence: Use of hydropathy scale in protein folding. *J. Chem. Phys.* **116:** 8579–8587.

Still, W.C., Tempczyk, A., Hawley, R.C., and Hendrickson, T. 1990. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112:** 6127–6129.

Toniolo, C., Polese, A., Formaggio, F., Crisma, M., and Kamphuis, J. 1996. Circular dichroism spectrum of a peptide 3_{10}-helix. *J. Am. Chem. Soc.* **118:** 2744–2745.

Tsui, V. and Case, D.A. 2000. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **56:** 275–291.

van der Spoel, D. and Lindahl, E. 2003. Brute-force molecular dynamics simulations of villin headpiece: Comparison with NMR parameters. *J. Phys. Chem. B* **107:** 11178–11187.

Vardar, D., Buckley, D.A., Frank, B.S., and McKnight, C.J. 1999. NMR structure of an F-actin-binding ''headpiece'' motif from villin. *J. Mol. Biol.* **294:** 1299–1310.

Wang, M.H., Tang, Y.F., Sato, S.S., Vugmeyster, L., McKnight, C.J., and Raleigh, D.P. 2003. Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale. *J. Am. Chem. Soc.* **125:** 6032–6033.

Wen, E.Z., Hsieh, M.J., Kollman, P.A., and Luo, R. 2004. Enhanced ab initio protein folding simulations in Poisson–Boltzmann molecular dynamics with self-guiding forces. *J. Mol. Graph. Model.* **22:** 415–424.

Woody, R.W. 1968. Improved calculation of the nπ* rotational strength in polypeptides. *J. Chem. Phys.* **49:** 4797–4806.

———. 1996. Theory of circular dichroism of proteins. In *Circular dichroism and the conformational analysis of biomolecules* (ed. G.D. Fasman), pp. 25–67. Plenum, New York.

Woody, R.W. and Sreerama, N. 1999. Comment on ''Improving protein circular dichroism calculations in the far-ultraviolet through reparametrizing the amide chromophore.'' *J. Chem. Phys.* **111:** 2844–2845.

Yang, W.A. and Gruebele, M. 2004. Detection-dependent kinetics as a probe of folding landscape microstructure. *J. Am. Chem. Soc.* **126:** 7758–7759.

Yang, J.T., Wu, C.-S.C., and Martinez, H.M. 1986. Calculation of protein conformation from circular dichroism. *Methods Enzymol.* **130:** 208–269.

Zagrovic, B., Snow, C.D., Khaliq, S., Shirts, M.R., and Pande, V.S. 2002a. Native-like mean structure in the unfolded ensemble of small proteins. *J. Mol. Biol.* **323:** 153–164.

Zagrovic, B., Snow, C.D., Shirts, M.R., and Pande, V.S. 2002b. Simulation of folding of a small α-helical protein in atomistic detail using worldwide-distributed computing. *J. Mol. Biol.* **323:** 927–937.

Zhou, R.H., Berne, B.J., and Germain, R. 2001. The free energy landscape for β-hairpin folding in explicit water. *Proc. Natl. Acad. Sci.* **98:** 14931–14936.

Zhu, Y., Alonso, D.O.V., Maki, K., Huang, C.Y., Lahr, S.J., Daggett, V., Roder, H., DeGrado, W.F., and Gai, F. 2003. Ultrafast folding of α_3: A de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci.* **100:** 15486–15491.