Conformational Distribution of a 14-Residue Peptide in Solution: A Fluorescence Resonance Energy Transfer Study

Matthew J. Tucker, Rolando Oyola, and Feng Gai*

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received: December 11, 2004; In Final Form: January 18, 2005

We demonstrate here that a nitrile-derivatized phenylalanine residue, *p*-cyanophenylalanine (Phe_{CN}), and tryptophan (Trp) constitute a novel donor−acceptor pair for fluorescence resonance energy transfer (FRET). The Förster distance of this FRET pair was determined to be ∼16 Å and hence is well suited for determining relatively short separation distances. To validate the applicability of this FRET pair in conformational studies, we studied the conformational heterogeneity of a 14-residue amphipathic peptide, Mastoparan X (MPx peptide), in water and 7 M urea solution as well as at different temperatures. Specifically, seven nitrile-derivatized mutants of the MPx peptide, each containing a Phe_{CN} residue that replaces different positions along the peptide sequence (i.e., from position 5 to 11) and serves as a resonance energy donor to the native Trp residue at position 3, were studied spectroscopically. The FRET efficiencies obtained from these peptides allowed us to gain a global picture regarding the conformational distribution of the MPx peptide in different environments. Our results suggest that the MPx molecules exist in water as an ensemble of rather compact conformations, with a radius of gyration of ∼4.2 Å, whereas in 7 M urea the radius of gyration increases to ∼6.5 Å, indicating that the peptide conformations become more extended under this condition. However, we found that temperature had only a negligible effect on the size of the MPx peptide, underlining the difference between the thermally and chemically denatured states of polypeptides. The application of the Gaussian chain or the wormlike chain model allowed us to further obtain the probability distribution function of the separation distance between any two residues along the peptide sequence. We found that the effective bond length of the MPx peptide, obtained by using the Gaussian chain model, is 2.78 Å in water and 4.28 Å in 7 M urea.

Introduction

An increasing body of evidence suggests that the ensemble of denatured state of proteins may contain only a rather limited number of conformations in solution. For example, recent theoretical[1,2] and experimental studies[3−6] have shown that the backbone of a disordered polypeptide can frequently adopt polyproline II structure,[7−11] thus reducing the conformational space accessible to the peptide. In addition, NMR studies[12−17] have shown that many proteins under denaturing conditions can still exhibit residual native and/or nonnative long-range interactions. The existence of such long-range interactions in the denatured state of proteins, either hydrophobic or electrostatic in nature, has strong implication for protein folding. Though nativelike interactions can speed up folding if they facilitate the formation of the folding nucleus,[18] nonnative structural elements may slow folding because they must be disrupted.[19] Moreover, such interactions may result in topologically nativelike denatured proteins, as recently observed in several cases.[20−22] Here, we demonstrate, using a FRET approach[23] and a new intrinsic FRET pair, that even short disordered peptides can show similar behaviors, e.g., forming rather compact conformations in aqueous solution.

Fluorescence resonance energy transfer has been used extensively on a myriad of biological systems to measure the separation distance between a donor and an acceptor, in either static or dynamic fashion. In particular, it has been used to investigate the distribution of the end-to-end distance of random polymers and polypeptides. For example, Hass et al.[24] employed this technique to measure the radial distribution function of peptides consisting of 4−9 residues employing naphthalene as donor and a dansyl group as acceptor. More recent works involving the application of FRET to study conformational distribution of short peptides are those of Grustiananda et al.[25] and Schuler et al.[26] In these and other similar studies, the donor and acceptor (e.g., dye molecules) are typically attached to the termini of the peptides of interest. Therefore, the FRET efficiency only measures the average distance between the termini and thus precludes an accurate determination of the conformational distribution of the entire peptide backbone. A better approach would be to tune the separation distance between the donor and acceptor systematically within the peptide of interest while its conformation is minimally perturbed. Clearly, such an approach requires a donor and an acceptor that can be incorporated easily into the peptide anywhere within the sequence and also have similar properties of natural amino acids. Most commonly used dye molecules do not meet these requirements because they usually have very large sizes and also require special functional groups, such as thiol and amine groups, for labeling.[27] Moreover, even with such groups present, nonspecific labeling can still occur,[27] affecting the interpretation of the FRET results. Here, we demonstrate that *p*-cyanophenylalanine, which may be regarded as a derivative of either phenylalanine (Phe) or tyrosine (Tyr), and tryptophan constitute an ideal FRET pair for studies that involve measuring relatively short separation distances, such as conformational distribution of short disordered peptides in solution.

* To whom correspondence should be addressed. E-mail: gai@sas.upenn.edu.

10.1021/jp044347q CCC: $30.25
Published on Web 02/22/2005

We have shown recently that, owing to its small size, moderate polarity, and environment-sensitive C≡N stretching vibration, Phe_{CN} can be used as a local infrared (IR) probe in biological studies.[28−30] Besides its interesting IR properties, however, Phe_{CN} also exhibits unique fluorescent properties that merit further exploration. For example, our preliminary results indicate that this nonnatural amino acid has a much higher fluorescence quantum yield than its natural counterpart, Phe, and unlike Tyr, it also lacks photoinduced proton transfer.[31] The latter can complicate the interpretation of the photophysics of the chromophore. Most importantly, we find that Phe_{CN} can be selectively excited even when Trp is present and its fluorescence spectrum overlaps significantly with the absorption spectrum of Trp. Therefore, Phe_{CN} and Trp should form an efficient FRET pair that is suitable for a variety of biophysical studies.

To demonstrate the applicability of Phe_{CN} and Trp as a FRET pair in conformational studies, we investigated the conformational distribution of a 14-residue amphipathic peptide, Mastoparan X (MPx peptide with the following sequence: INWKGIAAMAKKLL), which has been shown to lack any detectable secondary structures in aqueous solution, although it becomes helical when bound to membranes or micelles.[32,33] Specifically, seven nitrile-derivatized mutants of the MPx peptide, i.e., MPx−CN*y* ($y = 5$ to 11) peptides with the following sequences

MPx-CN5:IN-Trp-K−Phe_{CN}-IAAMAKKLL

MPx-CN6:IN-Trp-KG-Phe_{CN}-AAMAKKLL

MPx-CN7:IN-Trp-KGI−Phe_{CN}-AMAKKLL

MPx-CN8:IN-Trp-KGIA-Phe_{CN}-MAKKLL

MPx-CN9:IN-Trp-KGIAA-Phe_{CN}-AKKLL

MPx-CN10:IN-Trp-KGIAAM-Phe_{CN}-KKLL

MPx-CN11:IN-Trp-KGIAAMA-Phe_{CN}-KLL

were studied spectroscopically. The goal is to provide a global description regarding the conformational ensemble of the MPx peptide in solution by measuring the separation distance between the donor and acceptor, which is systematically varied in the current case. Our results indicated that the MPx peptide exists in water as an ensemble of rather compact conformations, as assessed by the separation distance between the Trp and Phe_{CN} residues. However, in 7 M urea solution these conformations become more extended. Two theoretical models, i.e., the Gaussian chain model[34,35] and the wormlike chain model,[36,37] were further utilized to quantify the conformational distribution of the MPx peptide in water and 7 M urea solution.

Experimental Section

Materials. 2,2,2-Trifluoroethanol (TFE from Aldrich, 98%), and *p*-cyanophenylalanine (Bachem) were used without further purification. Millipore water was used to prepare aqueous solutions. The MPx−CN*y* peptides are the same as those used in our early infrared studies.[28] The concentration of these peptides was optically determined using the Trp absorbance at 280 nm.

CD Measurements. The far-UV circular dichroism (CD) spectra of MPx−CN5 in 0−80% TFE/water (v/v) mixtures were measured on a 62A DS spectropolarometer (Aviv Associates) with a 1 mm sample holder. The peptide concentration was approximately 40 μM.

Absorption Measurement. All UV−vis spectra were measured on a Lambda 25 UV−vis spectrometer (Perkin-Elmer).

Fluorescence Measurements. The fluorescence spectra were obtained on a Fluorolog 3.10 spectrofluorometer (Jobin Yvon Horiba, NJ) with 2 nm resolution and a 1 cm quartz sample holder. Except the temperature-dependent studies, all other FRET measurements were carried out at 20 °C. Temperature was regulated using a TLC 50 Peltier temperature controller (Quantum Northwest, WA). To minimize self-quenching, the optical density (OD) of each sample at the excitation wavelength was adjusted to be in the range of 0.1−0.2. To achieve a high signal-to-noise (S/N) ratio, an integration time of 2 s/nm was used in these experiments. Although the relatively long integration time could result in sample photobleaching, control experiments showed that the emission intensity changed less than 5% during an hour of measurement, demonstrating that the effect of photobleaching was minimal in these measurements.

Förster Calculation. The Förster distance, R_0, for the Phe_{CN}-Trp FRET pair was determined according to the following equation[31,38]

$$R_0^{\,6} = \left(\frac{9000(\ln 10)\kappa^2 Q_D}{128\pi^5 N\eta^4}\right) J(\lambda) \tag{1}$$

where κ^2 is an orientation factor and has been assumed to be 2/3 in the current study, Q_D is the intrinsic fluorescence quantum yield of the donor in the absence of the acceptor, N is Avogadro's number, and η is the refractive index of the medium. In the present study, $\eta = 1.33$ was used for both water and 20% TFE/water solution, and $\eta = 1.4$ was used for 7 M urea solution. The last term, $J(\lambda)$, is the overlap integral defined by the following equation[31,38]

$$J(\lambda) = \int_0^\infty F_D(\lambda)\, \epsilon_A(\lambda)\, \lambda^4\, \mathrm{d}\lambda \tag{2}$$

where $F_D(\lambda)$ is the peak-normalized emission spectrum of the donor and $\epsilon_A(\lambda)$ is the wavelength-dependent molar absorption coefficient of the acceptor. Subsequently, the separation distance between the donor and acceptor, r, was determined using the following relationship[39]

$$E = \frac{R_0^{\,6}}{R_0^{\,6} + r^6} \tag{3}$$

where E is the FRET efficiency, which was calculated according to the following equation[31,38]

$$E = 1 - \frac{I_{DA}}{I_D} \tag{4}$$

where I_{DA} and I_D are the integrated fluorescence intensities of the donor, with and without the presence of the acceptor, respectively.

Results and Discussion

Photophysics of Phe_{CN}. Compared with those of Phe, the absorption and emission spectra of Phe_{CN} show distinct red shifts. As shown (Figure 1), the absorption spectrum of Phe_{CN} has two maxima in the UV region, centered at approximately 233 and 280 nm, respectively, corresponding to the progression $\pi-\pi^*$ transitions of the benzene ring.[40] In addition, both the molar absorptivity and fluorescence quantum yield of Phe_{CN}

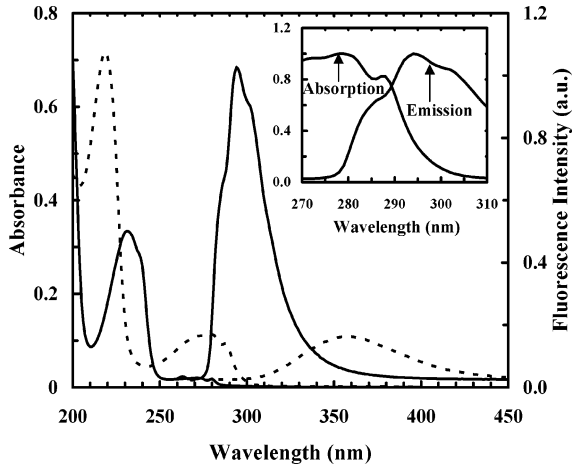

Figure 1. Absorption and emission spectra of Phe_{CN} (solid lines) and Trp (dashed lines) in water. The concentration was ~20 μM for both amino acids. $\lambda_{ex} = 240$ nm was used for fluorescence measurements. Inset: normalized absorption spectrum of Trp and emission spectrum of Phe_{CN}, as indicated.

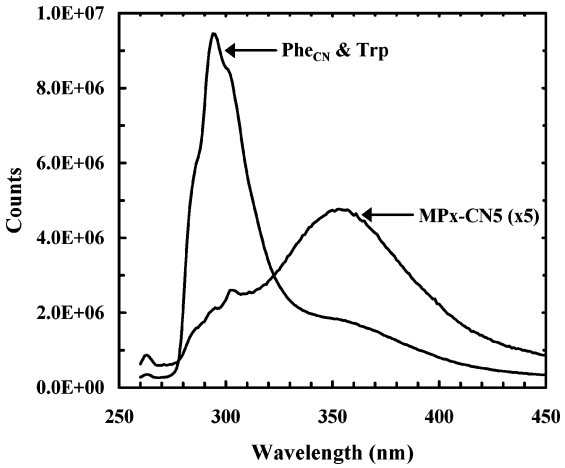

Figure 2. Fluorescence spectra of an equimolar (20 μM) Phe_{CN} and Trp aqueous solution and a 20 μM MPx−CN5 aqueous solution, as indicated. $\lambda_{ex} = 240$ nm.

are also increased significantly compared with those of Phe. For example, the molar absorptivity of Phe_{CN} was determined to be 13,000 ± 500 M^{-1} cm^{-1} at 233 nm and 850 ± 50 M^{-1} cm^{-1} at 280 nm, respectively, compared to the peak values of ~8500 M^{-1} cm^{-1} at 210 nm and ~200 M^{-1} cm^{-1} at 260 nm for Phe.[41] Similarly, the fluorescence quantum yield of Phe_{CN} was determined to be 0.11 ± 0.02, using Trp (quantum yield = 0.13) and Tyr (quantum yield = 0.14) as standards, whereas the quantum yield of Phe is 0.024.[42] This increase in both molar absorptivity and fluorescence quantum yield for Phe_{CN} is most likely due to the common benzene substitution effect with electron donating groups, such as hydroxyl and cyano, which increases the effective charge density of the π system through resonance.[40,43,44]

Although Phe has rarely been used as a fluorescent probe due to its weak fluorescence, the increased molar absorptivity and fluorescence quantum yield of Phe_{CN} make it a better chromophore for conformational studies, provided its emission is sensitive to environment. Moreover, the cyano substitution in Phe shifts the L_a band[40,43,44] of the phenyl group from 210 to 233 nm, where Trp has a much smaller molar absorptivity, making selective excitation of Phe_{CN} possible when other aromatic amino acids are present. The latter is extremely important if one would consider using Phe_{CN} and Trp as a FRET pair with Phe_{CN} serving as the donor (see below). As expected (Figure 2), when excited by 240 nm light, where Trp has a minimum absorbance, the fluorescence intensity of an equimolar Phe_{CN} and Trp aqueous solution is indeed dominated by the Phe_{CN} fluorescence emission.

Phe_{CN}−Trp as a FRET Pair. As shown (inset of Figure 1), the fluorescence emission spectrum of Phe_{CN} overlaps strongly with the absorption spectrum of Trp. Therefore, it is expected that Trp will quench the fluorescence of Phe_{CN} through the mechanism of resonance energy transfer when these two chromophores come within a certain distance of each other. Indeed, the fluorescence intensity of Phe_{CN} in MPx−CN5 peptide, where Phe_{CN} is separated from Trp by only one residue along the polypeptide backbone, becomes almost undetectable (Figure 2), whereas in the meantime the fluorescence intensity of Trp increases significantly as a consequence of resonance energy transfer. Furthermore, a quenching study on the fluorescence of Phe_{CN} with Trp yielded a linear Stern−Volmer plot in the Trp concentration range of 0−50 μM (data not shown). In particular, at 24 μM Trp causes less than 5% change in the Phe_{CN} fluorescence, indicating that the observed Phe_{CN} fluorescence quenching in the case of peptide is mostly due to intramolecular mechanisms, i.e., FRET. Taken together, these results suggest that Phe_{CN} and Trp form an efficient FRET pair. Using eqs 1 and 2, we further determined the Förster distance (R_0) of this FRET pair to be 16.0 ± 0.5 Å. Because many commonly used FRET dye pairs have an R_0 in the range 40−60 Å,[45] the combination of Phe_{CN} and Trp is therefore well suited for biophysical studies that involve the determination of relatively short separation distances, such as folding/unfolding of short loops and β-hairpins. Moreover, one of the biggest advantages of using Trp and Phe_{CN} as a FRET pair is that the former is a natural amino acid and it is straightforward to incorporate it into peptides and proteins. Similarly, the latter, which is a nonnatural amino acid, can also be incorporated easily anywhere into synthetic peptides using the standard methods of chemical synthesis, such as the Fmoc protocol, or into proteins by employing the recently developed biological methods.[46−50] Thus, this FRET pair can potentially be used as a molecular ruler to measure the separation distance between any two positions within a polypeptide. Furthermore, because of its small size, the replacement of a native residue with Phe_{CN} is expected to cause only a minimal perturbation to the native structure.[29] Therefore, this FRET pair is suitable for protein or peptide conformational studies.

MPx−CN*y* in Water. To demonstrate the applicability of using Phe_{CN} and Trp as a FRET pair, we studied the conformational distribution of the MPx peptide under different conditions. It has been shown by two-dimensional ^1H NMR and ^{15}N NMR spectroscopies that the MPx peptide in water does not form any detectable secondary structures.[32,33] Therefore, it provides a good model system for understanding the conformational properties of short disordered peptides in solution, such as the compactness as well as the probability distribution function of the displacement vector between any two residues. Specifically, seven peptides, i.e., MPx−CN*y*, which contains a Trp residue at position 3 and a Phe_{CN} residue at position *y*, with *y* varying from 5 to 11, were studied. Because the sequence separation between the Trp and Phe_{CN} residues in these peptides is systematically varied, the distribution of the separation distances between Trp and Phe_{CN}, which can be determined by measuring the FRET efficiencies, carries information regarding the conformational distribution of the MPx peptide. This, of course, necessitates the assumption that mutations with Phe_{CN}

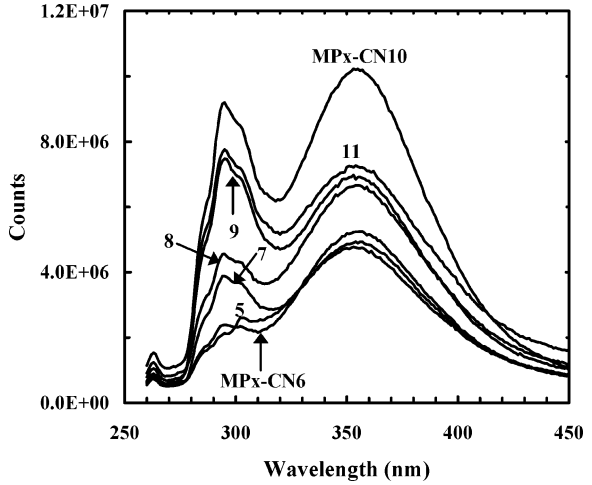

Figure 3. Fluorescence spectra of MPx−CN*y* peptides in water (~20 μM). λ_{ex} = 240 nm. These spectra have been scaled by the corresponding sample OD at the excitation wavelength.

along the sequence induce minimal perturbation to the native conformation.[28] Nevertheless, this approach is drastically different from those that only attach the donor and acceptor to the termini of the polypeptide, because the latter only permits the measurement of the average end-to-end distance.

The fluorescence emission spectra of these peptides in water show typical features of FRET. As shown (Figure 3), the Phe_{CN} fluorescence, resulting from a 240 nm excitation, is quenched to a certain degree, depending on the position of the Phe_{CN} residue. In the case of MPx−CN5 and MPx−CN6, where Trp is very close to Phe_{CN}, the Phe_{CN} fluorescence is almost completely quenched, whereas in the case of MPx−CN11, where Phe_{CN} is placed further away from the Trp residue, its fluorescence regains substantial intensity. Interestingly, however, the degree of quenching does not totally correlate with the sequence separation between the Phe_{CN} and Trp residues in these peptides. For example, a cursory examination shows that the MPx−CN10 peptide exhibits the strongest Phe_{CN} fluorescence, thus having the longest Phe_{CN}−Trp separation distance. Additionally, the MPx−CN9 and MPx−CN11 peptides have almost identical Phe_{CN} fluorescence, suggestive of quite similar Phe_{CN}−Trp separations in these two cases. Taken together, these results indicate that the MPx peptide does not yield an extended and/or random conformational ensemble in aqueous solution. If the peptide were extended, a longer sequence separation would result in, on average, a longer Phe_{CN}−Trp separation distance. Although these results demonstrated the applicability of Phe_{CN} and Trp to serve as a FRET pair, it should be noted that any mutations will have an effect on the native structure. Though such effects are expected to be minimal[28] for peptides that are intrinsically disordered in solution, we cannot rule out entirely the possibility that the observed irregular intensity distribution of the Phe_{CN} fluorescence is caused by Phe_{CN} substitution alone.

Following common practice, we further calculated the separation distance (r) between the donor (Phe_{CN}) and acceptor (Trp) for each case using eq 3 and the FRET efficiency. The latter was obtained with eq 4 by evaluating the ratio between the integrated areas of the Phe_{CN} fluorescence spectrum obtained with free Phe_{CN} residue and those obtained with the MPx−CN*y* peptide under the same conditions. Although choosing the free amino acid as the reference may cause uncertainties in the determination of the FRET efficiency, the effect is expected to be small because the fluorescence quantum yield of the free Phe_{CN} is within 85% of that of a peptide that contains a single

TABLE 1: FRET Efficiencies (±0.015) and the Corresponding Trp−Phe_{CN} Separation Distances (±0.40 Å)

	7 M urea		H_2O		20% TFE	
peptide	E	r (Å)	E	r (Å)	E	r (Å)
MPx−CN5	0.919	10.45	0.981	8.49	0.997	5.99
MPx−CN6	0.920	10.43	0.980	8.55	0.997	6.14
MPx−CN7	0.833	11.99	0.961	9.60	0.986	7.97
MPx−CN8	0.830	12.03	0.958	9.73	0.985	8.05
MPx−CN9	0.765	12.88	0.921	10.89	0.986	8.00
MPx−CN10	0.695	13.77	0.901	11.36	0.975	8.78
MPx−CN11	0.640	14.20	0.923	10.85	0.957	9.66

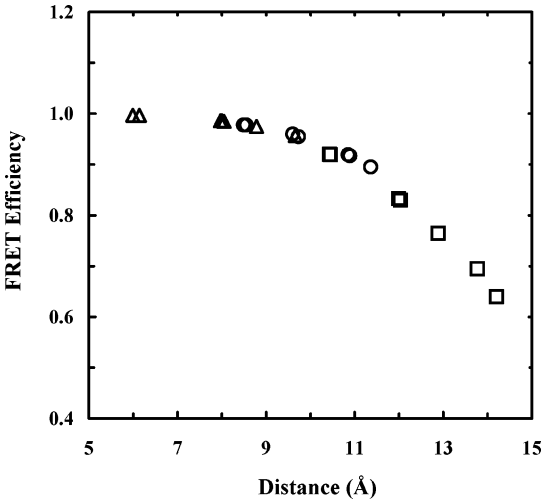

Figure 4. FRET efficiency vs separation distance between Trp and Phe_{CN} in MPx−CN*y* peptides under different conditions, i.e., in water (○), 20% TFE solution (△), and 7 M urea solution (□).

Phe_{CN} and no Trp (data not shown). To accurately determine the integrated area of the Phe_{CN} fluorescence, which overlaps with that of Trp, we fit the fluorescence spectrum of the MPx−CN*y* peptide by using a linear combination of two profiles generated from fitting the fluorescence spectra of Trp and Phe_{CN} plus a linear background. The resulting FRET efficiencies and the corresponding Trp−Phe_{CN} separation distances are summarized in Table 1. As shown (Figure 4), due to a relatively short separation distance between the donor and acceptor, the FRET efficiency is larger than 90% for all MPx−CN*y* peptides, indicating that the conformation of the MPx peptide in water is on average very compact, which is manifested by the fact that the longest Trp−Phe_{CN} separation distance is only about 11 Å, obtained on MPx−CN10 peptide.

It is well-known that heat can cause thermal denaturation of protein secondary and tertiary structures. To investigate how temperature affects the conformation of the MPx peptide, we further studied the FRET efficiency from Phe_{CN} to Trp in MPx−CN11 as a function of temperature. Specifically, nine fluorescence emission spectra of both free Phe_{CN} and MPx−CN11 were collected over a temperature range of 0−80 °C with a 10 °C step (data not shown). The measurement of the fluorescence spectrum of the free amino acid at each temperature is utilized to eliminate the intrinsic temperature dependence of the Phe_{CN} fluorescence. As shown (Figure 5), the FRET efficiency obtained with MPx−CN11 exhibits very weak temperature dependence, suggesting that even at 80 °C the MPx peptide still adopts an ensemble of very compact conformations. This is consistent with findings from other studies that the radius of gyration of thermally denatured proteins normally does not deviate significantly from that of their native states.[51,52]

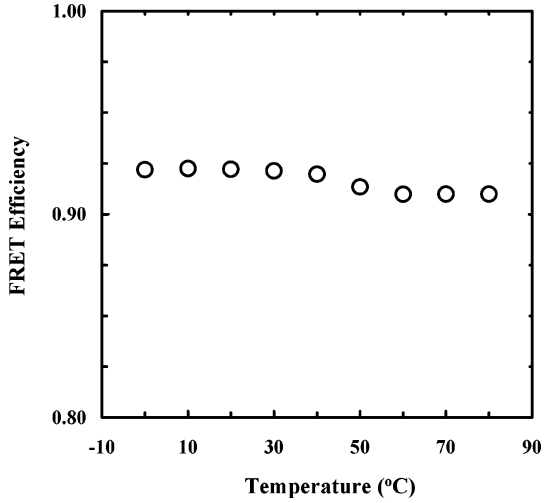

Figure 5. FRET efficiency of MPx−CN11 peptide as a function of temperature.

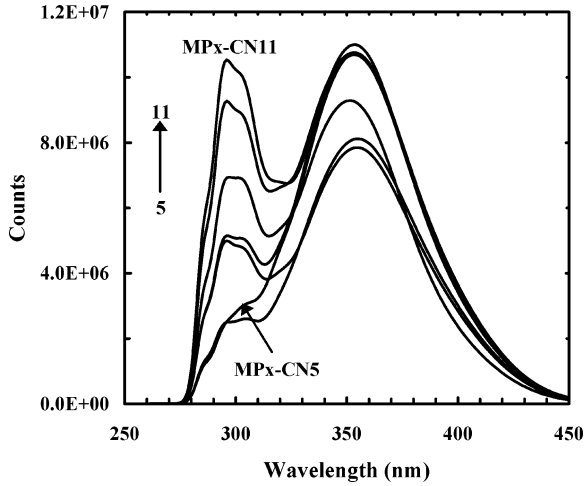

Figure 6. Fluorescence spectra of MPx−CN*y* peptides (∼25 μM) in 7 M urea. λ_{ex} = 240 nm. These spectra have been scaled by the corresponding sample OD at the excitation wavelength.

MPx−CN*y* in 7 M Urea Solution. Chemical denaturants, such as urea and guanidine hydrochloride, are commonly used to unfold proteins. Herein, we further studied the conformational distribution of the MPx peptide in 7 M urea solution using the method discussed above. As shown (Figure 6), the fluorescence emission spectra of the MPx−CN*y* peptides in 7 M urea solution are similar to those observed in water. The major difference is that the fluorescence intensity of the donor steadily increases as a function of the number of residues separating the donor from the acceptor, suggesting that the MPx peptide adopts on average a more extended conformation in 7 M urea solution than in water. Consistent with this picture, the Trp−Phe_{CN} separation distance for each MPx−CN*y* peptide in 7 M urea, calculated from the measured FRET efficiency, is longer than that obtained for the same peptide in water (Table 1 and Figure 4). In particular, the Trp−Phe_{CN} separation distance in MPx−CN11 peptide is lengthened by more than 30%.

MPx−CN*y* in 20% TFE Solution. As a control experiment, we also measured the FRET efficiencies of the MPx−CN*y* peptides in 20% TFE/water (v/v) solution. It is well-known that TFE can promote the formation of secondary structures.[54−56] As expected, the far-UV CD spectrum of MPx−CN5 peptide measured in 20% TFE solution shows typical features of helical structures (Figure 7). Moreover, the Phe_{CN} fluorescence of the

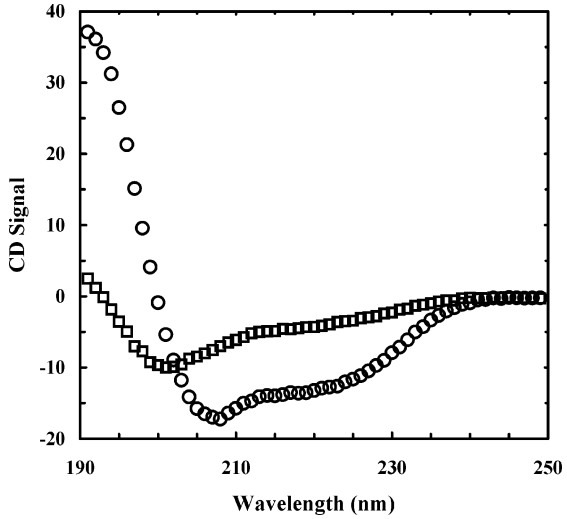

Figure 7. CD spectra of the MPx−CN5 peptide in water (□) and 20% TFE solution (○) at 20 °C. The peptide concentration was about 40 μM.

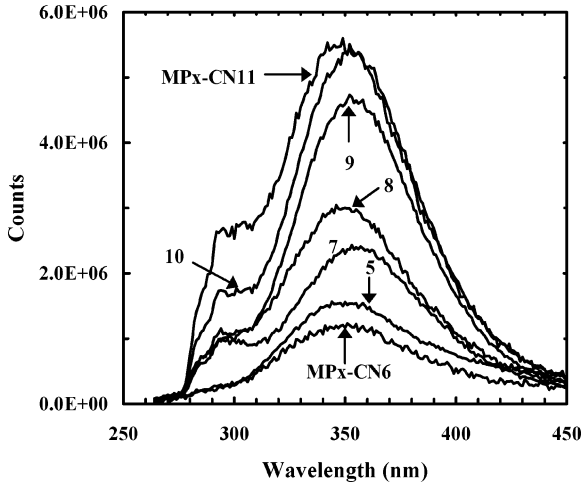

Figure 8. Fluorescence spectra of MPx−CN*y* peptides (∼20 μM) in 20% TFE/water (v/v) solution. λ_{ex} = 240 nm. These spectra have been scaled by the corresponding sample OD at the excitation wavelength.

MPx−CN*y* peptides under this condition is almost completely quenched (Figure 8), indicating that the Trp−Phe_{CN} separation distance in all of these peptides is significantly smaller than R_0 (Table 1 and Figure 4). For example, the longest Trp−Phe_{CN} separation distance, obtained on MPx−CN11, is ∼9.7 Å. These results therefore corroborate the CD results because the Trp−Phe_{CN} separation distance in MPx−CN11 is expected to be ∼10 Å[57] if the latter forms a helical conformation and with the assumption that the Trp and Phe_{CN} residues are separated by two helical turns. In addition, unlike the irregular distribution of the Trp−Phe_{CN} separation distances observed for MPx−CN*y* peptides in water, the Trp−Phe_{CN} separation distances measured in 20% TFE exhibit a distinct pattern as a function of the sequence separation between the Trp and Phe_{CN} residues, indicating that the peptide is well structured. For example, the first two Trp−Phe_{CN} separation distances (i.e., those determined on MPx−CN5 and MPx−CN6) are similar in value followed by the next four separation distances and finally an increase is observed in the last. This trend in separation distance seems to follow a 4-residue variation, as expected for an α-helical structure. Using the NMR structure[53] of the native MPx peptide determined in micelles, we further calculated the separation

TABLE 2: Separation Distances between the C_α of Trp3 and that (C_α') of the Native Residue at Position N, Calculated on the Basis of the NMR Structure of the MPx Peptide[53] and Trp−Phe_{CN} Separation Distances (±0.40 Å) Determined from the FRET Experiments in 20% TFE

position (*n*)	5	6	7	8	9	10	11
$C_\alpha-C_\alpha'$ (Å)	7.34	5.90	5.99	7.86	9.86	10.8	9.42
r_{FRET} (Å)	5.99	6.14	7.97	8.05	8.00	8.78	9.66

distances between the C_α atom of the Trp residue and that of the native residue that has been replaced by Phe_{CN} in the case of MPx−CN*y* peptides and compared these values to the determined FRET distances (Table 2). Indeed, these calculated $C_\alpha-C_\alpha'$ separations are quite comparable with those obtained from the FRET experiments employing the MPx−CN*y* peptides. Any small discrepancies can be attributed to the differences in rotamer distribution as well as the differences in the sizes of the native side chain and the side chain of Phe_{CN}. Taken together, however, these results confirm the applicability of using Trp and Phe_{CN} as a FRET pair for conformational studies.

Analysis Employing the Gaussian Chain and Wormlike Chain Models. For disordered peptides, the FRET measurement reports an ensemble average of the separation distances between the donor and acceptor. Hence, a more adequate description would be a probability distribution function describing the separation distance or displacement vector between the donor and acceptor. Below, we applied the Gaussian chain and the wormlike chain models to gain further insight into the understanding of the conformational distribution of the MPx peptide in water and 7 M urea solution. Although originated from polymer theories, both models have been used frequently to characterize the unfolded states of peptides and proteins.[58]

The Gaussian chain model,[34,35,59] which is commonly used to describe flexible polymer chains, has the following distribution function, $P(r)$, for the displacement vector between two points separated by a distance of r,

$$P(r) = 4\pi r^2 \left(\frac{3}{2\pi d^2}\right)^{3/2} e^{-3r^2/2d^2} \quad (5)$$

where d is the root-mean square distance. For peptides, d may be evaluated according to $d = bl^{1/2} + s$, where l represents the number of peptide bonds separating the two residues of interest, b is the effective bond length, which may be viewed as the average distance between two adjacent α-carbons along the polypeptide backbone, and s is a shifting distance term, introduced to account for the fact that the separation distance measured in a FRET experiment is between two flexible side chains. It has been suggested that the Gaussian chain model is better suited for describing the distribution statistics of residues with a large chain separation[58] and has been used recently to understand residual charge−charge interactions in unfolded proteins[34] as well as results from single-molecule folding studies.[26]

Another commonly used model is the wormlike chain model,[36,37,60] which has been used to model flexible linkers, such as loops connecting regular secondary structures, in proteins.[32] This model takes into account the fact that over a short contour length the tangent vector at each point along the chain may be correlated. Within the framework of the wormlike chain model, a good approximation describing the probability distribution of the displacement vector is given by[53,60]

$$P(r) = \left((4\pi r^2)\left(\frac{3}{4\pi l_p l_c}\right)^{3/2}\right)\left(1 - \frac{5l_p}{4l_c} + \frac{2r^2}{l_c^2} - \frac{33r^4}{80l_p l_c^3} - \frac{79l_p^2}{160l_c^2} - \frac{329r^2 l_p}{120l_c^3} + \frac{6799r^4}{1600l_c^4} - \frac{3441r^6}{2800l_p l_c^5} + \frac{1089r^8}{12800l_p^2 l_c^6}\right) \quad (6)$$

where l_c is the contour length of separation. For peptides, $l_c = lb$, where l is the number of peptide bonds separating the two residues; and similar to that of the Gaussian chain model, b is the effective bond length. In addition, l_p refers to the persistence length over which the tangent vector at each point along the chain is correlated. This characteristic length may also be treated as a measure of the stiffness of the chain.[36] For example, a large persistence length gives rise to a more rigid structure.[61] For long chains, it can be shown easily that for $l_c \gg l_p$ the wormlike chain model yields results consistent with those of the Gaussian chain model, whereas for $l_c \ll l_p$ the system is expected to behave as a rigid rod.[36] A recent study that involves analyzing the loop structures of 1907 proteins from the Protein Data Bank suggests that the wormlike chain model should be used to describe the distribution of end-to-end distances between two residues that have a chain separation of greater than or equal to 5 residues.[36]

For a flexible chain, the measured FRET efficiency, $\langle E \rangle$, is related to the probability distribution function, $P(r)$, through the following equation,

$$\langle E \rangle = \int_0^\infty E(r)\, P(r)\, dr \quad (7)$$

where $E(r)$ is the FRET efficiency corresponding to a particular donor−acceptor separation distance (i.e., r), defined by eq 3. To accurately determine the probability distribution function of the displacement vector for the MPx peptide, we globally fit the FRET efficiencies obtained on all 7 MPx−CN*y* peptides to eq 7, subjecting to the same $P(r)$ corresponding to either the Gaussian chain or wormlike chain model. For the FRET data obtained in water, the best fit using the Gaussian chain model yielded $b = 2.78$ Å and $s = 1.83$ Å, whereas for the FRET data obtained in 7 M urea, the best fit yielded $b = 4.28$ Å and $s = 1.83$ Å.[62] It is worthy of pointing out that if s was set to zero, the best fit yielded an effective bond length of 3.72 Å for water and 5.02 Å for urea, respectively. Taken together, these results allowed us then to quantitatively compare the distributions of the Trp−Phe_{CN} separation distance in different solvents (Figure 9). For short separation distances, such as in the case of the MPx−CN5 peptide, the Gaussian distribution is expected to be quite similar between those obtained in water and 7 M urea, as indicated. However, for longer separation distances, the Gaussian width increases significantly when going from water to 7 M urea as the peptide molecules in the latter case sample more and longer distances.

Although these results certainly indicated that concentrated chemical denaturant solution, such as 7 M urea, can force the MPx peptide to adopt a more extended conformation, the effective bond lengths obtained from the Gaussian chain model for both cases are somewhat shorter than values reported in other studies for unfolded proteins.[58] A reasonable explanation for this discrepancy is that the current study employed a relatively short peptide, with only 14 residues. Previous studies have shown that the effective bond length depends on the number of residues. For example, for a 16-residue peptide, b was determined to be 5.6 Å, whereas for a 104-residue protein, b was determined to be about 8.8 Å.[58]

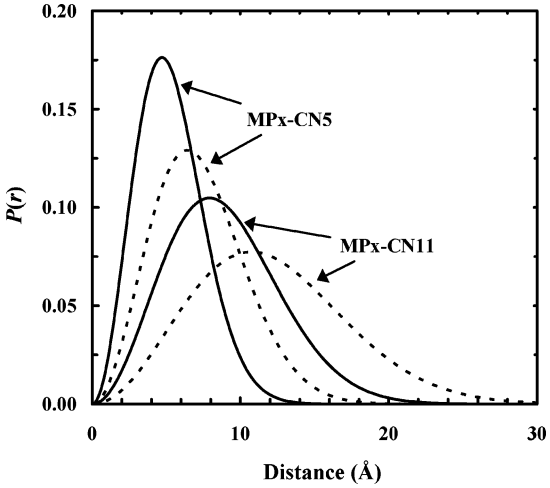

Figure 9. Probability distribution of the separation distance between Trp and Phe_{CN} in MPx−CN5 and MPx−CN11 in water (solid lines) and 7 M urea (dashed lines), as indicated.

Interestingly, the wormlike chain model failed to provide a meaningful description about the conformational distribution of the MPx peptide in water because the best fit to the measured FRET efficiencies yielded a very large and unreasonable value of b and also a small value of l_p. Because the wormlike chain model is best suited for describing chains with strong correlation between tangential vectors along the chain,[61] a plausible reason for the breakdown of the wormlike chain model here stems from the fact that the conformations of the MPx peptide in water are rather compact, and most likely contain static and/or transient residual secondary structures as well as other compact or even folded conformations, thus the correlation between adjacent vectors does not persist.

In the case of 7 M urea, the best fit to the FRET efficiencies employing the wormlike chain model yielded $b = 4.1$ Å and $l_p = 3.2$ Å, respectively. Similar to the results obtained with the Gaussian chain model, these values also suggest that in 7 M urea the MPx peptide adopts more extended conformations. Moreover, the effective bond length obtained from the wormlike chain model is similar to that yielded by the Gaussian chain model, suggesting that both models are capable of describing the conformational distribution of polypeptides in concentrated chemical denaturant solutions. In addition, the value of the persistent length obtained here (i.e., 3.2 Å) is comparable to those suggested by mechanical protein unfolding studies, which range from 3.3 to 4.2 Å (for a review, see ref 58), and also the value of 3.04 Å suggested for flexible linkers connecting folded domains in proteins.[36]

Results from FRET experiments indicate that the conformation of the MPx peptide in aqueous solution is compact, whereas in 7 M urea solution it is significantly elongated. Besides the probability distribution function of the displacement vector, another commonly used indicator of the compactness of a protein or peptide conformation is the radius of gyration, R_g, which can be calculated according to the following equation,[58]

$$R_g = b\left(\frac{N}{6}\right)^{1/2} \tag{8}$$

where b is the effective bond length and N is the total number of residues in the peptide. Using the values determined from the Gaussian chain model, we obtained a radius of gyration of 4.2 Å for the MPx peptide in water and 6.5 Å for the MPx peptide in 7 M urea, respectively. These results further illustrate the difference between the conformational ensembles of the MPx peptide in H_2O and 7 M urea.

Conclusion

Phe_{CN} (*p*-cyanophenylalanine) has an increased fluorescent quantum yield and a higher molar absorptivity than Phe, thus making it a better fluorescent probe. In addition, Phe_{CN} can be selectively excited even when Trp is present (e.g., using an excitation wavelength of 240 nm) and the resulting fluorescence spectrum overlaps significantly with the absorption spectrum of Trp. Hence, Phe_{CN} and Trp can be used together to form an efficient FRET pair, whose Förster distance was determined to be about 16 Å. Compared with other commonly used FRET-coupled dye pairs, the Trp−Phe_{CN} FRET pair offers certain advantages for biophysical studies. For example, incorporating these amino acids into a peptide sequence requires only standard synthesis protocols and their perturbation to the native conformation is expected to be minimal. Using a 14-residue peptide, we tested the applicability of the Trp−Phe_{CN} FRET pair for conformational studies. Our results indicated that the supposedly disordered MPx peptide molecules exist in water as an ensemble of rather compact conformations, whereas a high concentration of urea can effectively lengthen these conformations. Globally fitting the measured FRET efficiencies to the Gassuain chain model resulted in the following effective bond lengths: $b = 2.78$ Å for water and $b = 4.28$ Å for 7 M urea. Although this study alone cannot provide more detailed information regarding the conformations formed by the MPx peptide in these solutions, our results are in accord with recent observations that the conformation of short peptides in aqueous solution is not random and/or extended.

Acknowledgment. We gratefully acknowledge financial support from the NSF (CHE-0094077) and the NIH (GM-065978, RR-13456). We also wish to thank Drs. Z. Getahun and W. F. DeGrado for providing the peptide samples.

References and Notes

(1) Han, W. G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587.
(2) Mezei, M.; Fleming, P. J.; Srinivasan, R.; Rose, G. D. *Proteins* **2004**, *55*, 502.
(3) Rucker, A. L.; Creamer, T. P. *Protein Sci.* **2002**, *11*, 980.
(4) Shi, Z. S.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190.
(5) Eker, F.; Cao, X.; Nafie, L.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2002**, *124*, 1433.
(6) Wang, F.; Polavarapu, P. L. *Biopolymers* **2003**, *70*, 614.
(7) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. *Science* **2002**, *295*, 1719.
(8) Tiffany, M. L.; Krimm, S. *Biopolymers* **1968**, *6*, 1379.
(9) Drake, A. F.; Siligardi, G.; Gibbons, W. A. *Biophys. Chem.* **1988**, *31*, 143.
(10) Dukor, R. K.; Keiderling, T. A. *Biopolymers* **1991**, *31*, 1747.
(11) Woody, R. W. *Adv. Biophys. Chem.* **1992**, *2*, 37.
(12) Wilson, G.; Hecht, L.; Barron, L. D. *Biochemistry* **1996**, *35*, 12518.
(13) Mok, Y. K.; Kay, C. M.; Kay, L. E.; Forman-Kay, J. *J. Mol. Biol.* **1999**, *289*, 619.
(14) Zhang, O.; Forman-Kay, J. D. *Biochemistry* **1997**, *36*, 3959.
(15) Shortle, D. R. *Curr. Opin. Struct. Biol.* **1996**, *6*, 24.
(16) Neri, D.; Billeter, M.; Wider, G.; Wuthrich, K. *Science* **1992**, *257*, 1559.
(17) Wong, K. B.; Clarke, J.; Bond, C. J.; Neira, J. L.; Freund, S. M. V.; Fersht, A. R.; Daggett, V. *J. Mol. Biol.* **2000**, *296*, 1257.
(18) Tang, Y.; Rigotti, D. J.; Fairman, R.; Raleigh, D. P. *Biochemistry* **2004**, *43*, 3264.
(19) Cho, J. H.; Sato, S.; Raleigh, D. P. *J. Mol. Biol.* **2004**, *338*, 827.
(20) Ohnishi, S.; Lee, A. L.; Edgell, M. H.; Shortle, D. *Biochemistry* **2004**, *43*, 4064.
(21) Shortle, D.; Ackerman, M. S. *Science* **2001**, *293*, 487.

(22) Lietzow, M. A.; Jamin, M.; Dyson, H. J.; Wright, P. E. *J. Mol. Biol.* **2002**, *322*, 655.
(23) Stryer, L. *Annu. Rev. Biochem.* **1978**, *47*, 819.
(24) Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1807.
(25) Gustiananda, M.; Liggins, J. R.; Cummins, P. L.; Gready, J. E. *Biophys. J.* **2004**, *86*, 2467.
(26) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743.
(27) Karolin, J.; Fa, M.; Wilczynska, M.; Ny, T.; Johansson, L. B. A. *Biophys. J.* **1998**, *74*, 11.
(28) Tucker, M. J.; Getahun, Z.; Nanda, V.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2004**, *126*, 5078.
(29) Getahun, Z.; Huang, C. Y.; Wang, T.; De Leon, B.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2003**, *125*, 405.
(30) Huang, C. Y.; Wang, T.; Gai, F. *Chem. Phys. Lett.* **2003**, *371*, 731.
(31) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy;* Second ed.; Academic Press: New York, 1999.
(32) Wakamatsu, K.; Okada, A.; Miyazawa, T.; Ohya, M.; Higashijima, T. *Biochemistry* **1992**, *31*, 5654.
(33) Hori, Y.; Demura, M.; Iwadate, M.; Ulrich, A. S.; Niidome, T.; Aoyagi, H.; Asakura, T. *Eur. J. Biochem.* **2001**, *268*, 302.
(34) Zhou, H. X. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3569.
(35) Zhou, H. X.; Dill, K. A. *Biochemistry* **2001**, *40*, 11289.
(36) Zhou, H. X. *J. Phys. Chem. B.* **2001**, *105*, 6763.
(37) Kratky, O.; Porod, G. *Rec. Trav. Chim.* **1949**, *68*, 1106.
(38) Steinberg, I. Z. *Annu. Rev. Biochem.* **1971**, *40*, 83.
(39) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 719.
(40) Lewis, F. D.; Holman, B. *J. Phys. Chem.* **1980**, *84*, 2326.
(41) Wetlaufer, D. B. *Adv. Protein Chem.* **1962**, *17*, 303.
(42) Demchenko, A. P. *Ultravilolet Spectroscopy of Proteins*; Springer-Verlag: New York, 1986.
(43) Borst, D. R.; Korter, T. M.; Pratt, D. W. *Chem. Phys. Lett.* **2001**, *350*, 485.
(44) Youn, Y. Y.; Kwon, C. H.; Choe, J. C.; Kim, M. S. *J. Chem. Phys.* **2002**, *117*, 2538.
(45) Wu, P.; Brand, L. *Anal. Biochem.* **1994**, *218*, 1.
(46) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705.
(47) Severinov, K.; Muir, T. W. *J. Biol. Chem.* **1998**, *273*, 16205.
(48) Evans, T. C., Jr.; Benner, J.; Xu, M.-Q. *Protein Sci.* **1998**, *7*, 2256.
(49) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498.
(50) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 19.
(51) Dill, K. A.; Shortle, D. *Annu. Rev. Biochem.* **1991**, *60*, 795.
(52) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153.
(53) Kusunoki, H.; Wakamatsu, K.; Sato, K.; Miyazawa, T.; Kohno, T. *Biochemistry* **1998**, *37*, 4782.
(54) Goodman, M.; Listowsky, I. *J. Am. Chem. Soc.* **1962**, *84*, 3770.
(55) Buck, M. *Q. Rev. Biophys.* **1998**, *31*, 297.
(56) Andersen, N. H.; Dyer, R. B.; Fesimeyer, R. M.; Gai, F.; Liu, Z.; Neidigh, J. W.; Tong, H. *J. Am. Chem. Soc.* **1999**, *121*, 9879.
(57) Ceighton, T. E. *Proteins: Structures and Molecular Properties*; W. H. Freeman and Company: New York, 1984.
(58) Zhou, H. X. *Biochemistry* **2004**, *43*, 2141.
(59) Zhou, H. X. *Biophys. J.* **2002**, *83*, 2981.
(60) Gobush, W.; Stockmayer, W. H.; Yamakawa, H.; Magee, W. S. *J. Chem. Phys.* **1972**, *57*, 2839.
(61) Wilhelm, J.; Frey, E. *Phys. Rev. Lett.* **1996**, *77*, 2581.
(62) In these fittings, the *s* was treated as a global parameter.