Intermolecular Interaction of Myoglobin with Water Molecules along the pH Denaturation Curve

Naoki Baden and Masahide Terazima*

Department of Chemistry, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

Received: January 11, 2006; In Final Form: April 6, 2006

A method of diffusion coefficient (*D*) measurement for proteins based on the pulsed laser-induced transient grating method using a photosensitive cross-linker was applied to the characterization of the pH denaturation process of holo- and apo-myoglobin (Mb) from the viewpoint of protein−water interaction. It was found that the pH denaturation curve monitored by *D* agrees quite well with that determined by the circular dichroism intensity for holo-Mb. This fact indicates that the changes in intermolecular interaction and the α-helix content occur simultaneously during the unfolding process. However, the pH dependence of *D* for apo-Mb was different from that of α-helix content. This different behavior can be explained in terms of the different denaturation steps for the secondary structure and the hydrogen bonding network of the intermediate species around pH 4; i.e., this intermediate is partially unfolded, but the hydrogen bonding network is dominantly an intramolecular one. Taking previously reported properties of this species into account, we conclude that water molecules are trapped in the hydrophobic core of the apo-Mb pH 4 intermediate. This fact suggests that the kinetic intermediate state of the protein folding process is a swollen state without water molecular exchange with the bulk phase.

1. Introduction

To understand the nature of protein stability or protein folding, it will be essentially important to develop an experimental technique to monitor protein conformation in the solution phase under various conditions. Indeed, a variety of techniques have been developed so far for that purpose, such as circular dichroism (CD), nuclear magnetic resonance (NMR), fluorescence detection, absorption measurement, small-angle X-ray scattering (SAXS), and so on.[1] Each method has a characteristic feature to probe protein characteristics. For example, emission intensity or absorption measurements provide us information on rather local protein structures around a chromophore associated with an optical transition. The secondary structure (e.g., α-helices or *β*-sheets) can be monitored by CD or infrared absorption techniques. One can measure the radius of gyration (R_g) by the SAXS method. In comparison with a variety of methods to monitor these geometric conformations, information on the intermolecular interactions between a protein and surrounded water molecules is very scarce, even though this property is very important to characterize and understand the nature of proteins. Recently, it was proposed that the diffusion coefficient (*D*) could be useful for probing this property.[2−5] In this paper, we tried to characterize a protein under different denaturation conditions in terms of the intermolecular interaction with water molecules by using *D*.

According to the Stokes−Einstein relationship, *D* of a spherical molecule with a radius of *R* in continuum medium is given by[6]

$$D = k_B T/a\eta R \quad (1)$$

where k_B, *T*, *η*, and *a* are, respectively, the Boltzmann constant, temperature, viscosity, and a constant representing the boundary condition between the diffusing molecule and the solvent. From this relationship, it seems that *D* can be a representative of only the molecular size (*R*) under a given environment (*T* and *η*). However, this is not correct in general, because the intermolecular interaction also affects the magnitude of *D*.[7] Indeed, it has been reported, for example, that *D* values of many transient radicals are significantly smaller than those of nonradical molecules with a similar shape and size.[8−10] This difference was explained in terms of the enhanced intermolecular interaction between the solute and the solvent. Recently, the time dependence of *D* was reported during the protein refolding process of cytochrome c and interpreted in terms of the changes in the hydrogen bonding network.[2−4] Therefore, this *D* measurement could be potentially useful to probe the intermolecular interaction of a protein with water molecules in solution. However, the relationship between *D* and the conformation of a protein has not been established well yet.

Since *D* is a fundamental and valuable quantity, various experimental methods have been developed so far for *D* measurement, such as spin−echo NMR detection with a field gradient, Taylor dispersion, or dynamic light scattering.[6,11] Each method possesses its merits and disadvantages. For application to *D* measurement of proteins in dilute solution under various conditions, one may encounter some difficulties. For example, dynamic light scattering[12] and pulsed field gradient NMR[13,14] methods need relatively long time measurements to obtain an appropriate signal-to-noise ratio for a sample of low-concentration protein.[6,11] Fluorescence recovery after a photobleaching method[15] requires the synthesis of a protein labeled by a fluorescent molecule. The electrophoresis method requires additional information on the surface charge.[16] The sedimentation[17] measurements have difficulties in application to small proteins. For overcoming these difficulties, we have previously proposed a new convenient method for *D* measurement of macromolecules such as proteins and DNA based on the time-resolved transient grating (TG) method with a photoreactive

* Author to whom correspondence should be addressed. E-mail: mterazima@kuchem.kyoto-u.ac.jp.

10.1021/jp0602171 CCC: $33.50
Published on Web 07/15/2006

cross-linker molecule.[18] The photoreactive labeling reagent that we used was *N*-hydroxysulfosuccinimidyl-4-azidobenzoate (sulfo-HSAB). This method has several merits such as simple preparation, a high sensitivity, and a wide applicability to many macromolecules. It only requires a very short time for measurement (less than 1 s). We have demonstrated the reliability of this method using some proteins.[18] In this paper, we have two aims; we report an experimental procedure and signal analysis for *D* measurement by this new technique and use this method to study *D* of myoglobin (Mb) under various pH conditions for investigating the relationship between the protein conformation and *D*.

Myoglobin is a prototypical globular protein, whose characteristics have been studied for a long time.[19] The stability and the conformational changes during acid denaturation have been studied by CD (the amount of secondary structure)[20] and tryptophan Raman scattering (exposure of the Trp residues to water),[21] both of which show the high cooperativity of the denaturation process with a single transition ($pK_a = 3.5$). In particular, apo-myoglobin (apo-Mb), which lacks the heme group of holo-myoglobin (holo-Mb), has an unique and interesting characteristic in the acid denaturation process; i.e., it shows an equilibrium intermediate around pH 4 instead of two-state unfolding.[20,21] Since this equilibrium intermediate structure has been considered to be similar to that of the kinetic intermediate observed after 5 ms in a pH-jump experiment in a protein folding study,[22] the nature of the intermediate has been investigated extensively and served as a model state for studies of folding mechanism. This pH 4 intermediate has several characteristic properties of a molten globule (MG), such as a compact size,[23] preservation of a considerable amount of native secondary structure,[24] and a disordered tertiary structure.[25] For characterizing the conformation of holo- and apo-Mb in different pH solutions, we have measured the pH dependence of *D* of holo- and apo-Mb by using the TG method. We observed drastic changes in *D* in the transition between folded (pH 5) and unfolded states (pH 2) of both Mb's. While the pH dependence of *D* for holo-Mb was similar to that of the α-helix content obtained by CD measurements, the pH dependences monitored by the CD and *D* measurements are different from each other for apo-Mb. The result is interpreted in terms of the swollen structure of the pH 4 intermediate by water molecules trapped inside the protein.

2. Principle of Measurement

In the TG experiment, a photoinduced reaction is initiated by pulsed excitation laser light beam having sinusoidally modulated intensity, which is produced by the interference of two excitation light waves.[26] The sinusoidal modulations of the concentrations of a reactant and a product lead to the modulation of the refractive index (δn), which is called the species grating. The released thermal energy by the nonradiative transition from the excited states creates the temperature modulation, which leads to the refractive index change. This grating is called the thermal grating. These modulations can be monitored by the diffraction efficiency of a probe beam (TG signal). The intensity of the TG signal (I_{TG}) is given by[27]

$$I_{TG}(t) = \alpha\{\delta n(t)\}^2 \quad (2)$$

where α is a constant. The signal intensity becomes weaker as the spatial modulations of the refractive index become uniform, which is accomplished by thermal diffusion or translational mass diffusion.

Solving the diffusion equation, one may find that the decay rate constant of the thermal grating signal should be $D_{th}q^2$ (D_{th}, thermal diffusivity of the solution; q, grating wavenumber).[28] Similarly, the species grating signal decays with a rate constant of Dq^2, where D is the diffusion coefficient of the chemical species. When concentrations of the chemical species and D of the reactant (D_r) and the product (D_p) are constants in time, the amplitude of the modulation of the refractive index change (δn) can be expressed by a sum of exponential functions[28]

$$\delta n(t) = \delta n_{th} \exp(-D_{th}q^2t) - \sum_r \delta n_r \exp(-D_r q^2 t) + \sum_p \delta n_p \exp(-D_p q^2 t) \quad (3)$$

where δn_{th} (<0) is the refractive index change due to the temperature change and δn_r (>0) and δn_p (>0) are, respectively, the refractive index changes due to the presence of the reactant and the product. The sign of δn_r is negative, because the phase of the spatial concentration modulation is 180° shifted from that of the product. Therefore, from the temporal profile of the TG signal, the diffusion coefficients of chemical species involved in the reaction can be measured.

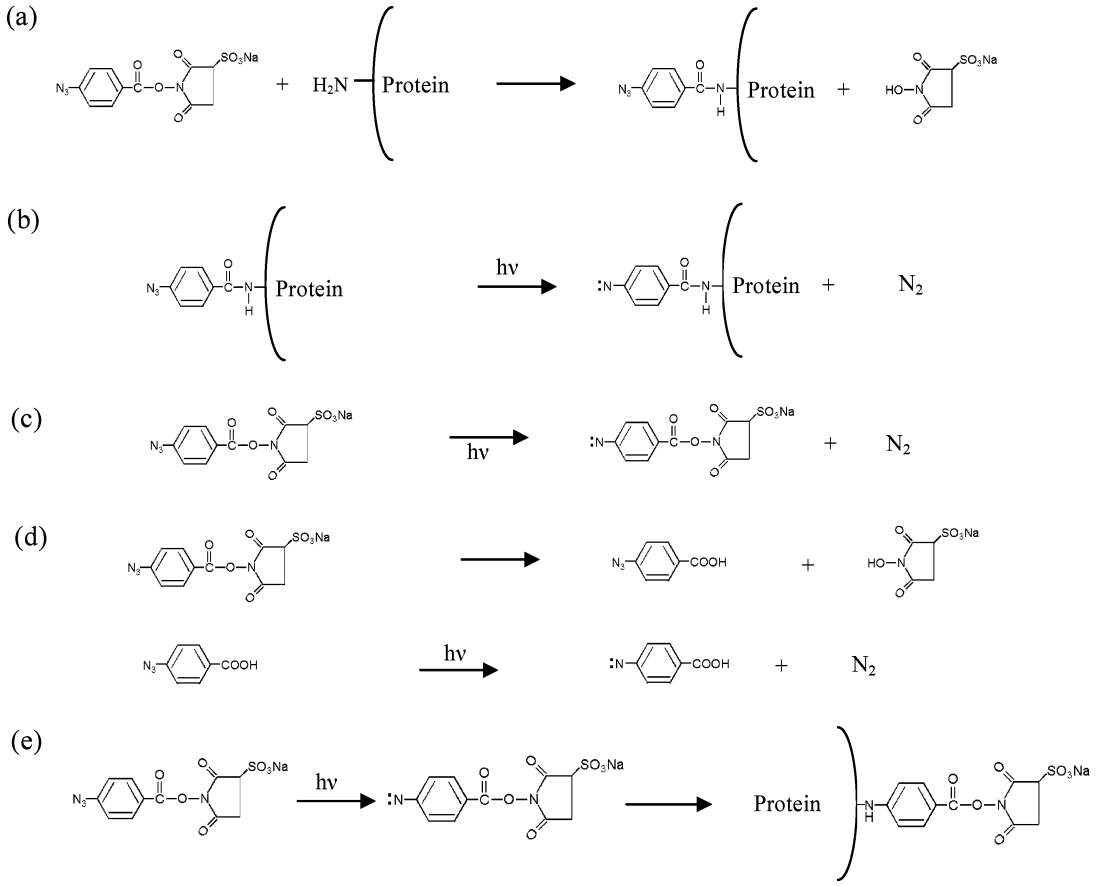
For photolabeling the proteins by the spatially modulated light, sulfo-HSAB was used. This sulfo-HSAB has been developed and used as a photoactive cross-linking reagent.[29] When this reagent is mixed with protein solution, the *N*-hydroxysuccinimide (NHS) end reacts with the proteins (Scheme 1a). This reaction should take place in darkness to protect the aryl azide from photolysis. The photoactive end (azide) releases a nitrogen molecule upon photoexcitation (Scheme 1b). This photochemical reaction induces the refractive index modulation to produce the TG signal, from which *D* can be measured.

3. Experimental Section

The experimental setup of the TG measurement was similar to those reported previously.[9,10,18] An excimer laser (XeCl operation, a pulse width of 20 ns) was used for excitation, and a photodiode laser (780 nm) as a probe laser. The TG signal induced by the pulsed laser irradiation was detected by a photomultiplier tube and fed into a digital oscilloscope for recording. A sample solution of ca. 0.1 mL was contained in a quartz cell with a path length of 2 mm. The excitation beams were slightly focused at the sample position to have a spot size of ca. 1 mm in diameter. Hence, the volume of the excited region is much smaller than that of the unilluminated region. Sometimes, the depletion of the photoreactive labeling reagent at the laser-irradiated region was substantial, in particular, when the concentration of the sample was low. In such a case, the sample solution was stirred after each measurement to restore the initial concentration in the photoexcited region. The obtained TG signals were averaged around 10−20 times to improve the signal-to-noise ratio. All measurements were carried out at room temperature.

The sample solution was prepared by the following procedure. First, a concentrated aqueous solution of sulfo-HSAB (100 times higher than that for measurement) was prepared. Then this solution was mixed with a protein aqueous solution with a volume ratio of 1:100 in darkness shortly before the measurement. The sample solution was filtered by a membrane with a pore size of 0.45 μm. The pH value was adjusted by concentrated HCl aqueous solution. The cross-linker reagent, sulfo-HSAB, was purchased from Pierce Co. Horse Mb and apo-Mb were obtained from Nacalai Tesque, Inc. and Sigma, respectively.

SCHEME 1: Reaction Schemes of the (a) Acylation Reaction of Sulfo-HSAB with a Protein in Darkness, (b) Photolysis of the Aryl Azide after Sulfo-HSAB Labeling on a Protein, (c) Photoreaction of Sulfo-HSAB, (d) Hydrolysis and Subsequent Photolysis of Sulfo-HSAB, and (e) Photoinduced Addition of Free Sulfo-HSAB to a Protein



4. Results

4.1. TG Signal of Sulfo-HSAB. The fundamental feature of the TG signal observed after photoexcitation of sulfo-HSAB without a protein was reported previously.[18] Before analyzing the TG signal of the protein solution, we should first describe the features and the analysis of the sulfo-HSAB signal, because this signal will be the basis of the signal analysis of protein solutions later. The observed TG signal of sulfo-HSAB in water is shown in Figure 1. Under a high grating wavenumber ($q^2 > 1.2 \times 10^{12}$ m^{-2}) condition, the signal rises within the excitation pulse width, decays to the baseline, and shows a rise−decay curve. However, in a low grating wavenumber region ($q^2 < 1.0 \times 10^{12}$ m^{-2}), one additional rise−decay signal was observed after the thermal diffusion signal. We found that the signal after complete decay of the thermal grating signal can be reproduced well by a sum of four exponential functions. One rate constant among these components does not depend on q^2, whereas the others depend on q^2. The q^2 dependence indicates that these dynamics represent molecular diffusion processes. Hence, the observed TG signal intensity (I_{TG}) can be expressed by

$$I_{TG}(t) = \alpha\{\delta n_k \exp(-kt) + \delta n_1 \exp(-D_1 q^2 t) - \delta n_2 \exp(-D_2 q^2 t) + \delta n_3 \exp(-D_3 q^2 t)\}^2 \quad (4)$$

where k is the q^2-independent rate constant and D_i ($i = 1, 2, 3$) are the diffusion coefficients of chemical species involved in the reaction ($D_1 > D_2 > D_3$). The preexponential factor of δn_k is negative, and the others are positive. From the analysis of the TG signals at different wavenumbers, the time constant of the q^2-independent component k^{-1} was determined to be about 76 μs, which is close to the reported lifetime of a nitrene-derived intermediate, 100 μs.[30] The diffusion coefficients of D_1, D_2, and D_3 are 6.3 ($\pm$1.2) $\times 10^{-9}$, 8.0 ($\pm$1.0) $\times 10^{-10}$, and 3.0 ($\pm$0.5) $\times 10^{-10}$ m^2 s^{-1}, respectively. Considering the magnitude of D and the reaction scheme, we may attribute the diffusing molecule of D_1 to a reaction product with a relatively small size, such as the nitrogen molecule. Since the reported D value of nitrogen molecule in water (2.3×10^{-9} m^2 s^{-1})[31] is smaller than D_1, some other smaller molecules such as OH^-, which could be generated by the reaction of nitrene with a water molecule,[32] may contribute to this component simultaneously, and the D_1 value from the profile may not be accurate. The second component $D_2 = 8.0 \times 10^{-10}$ m^2 s^{-1} is too large to be assigned to either sulfo-HSAB (parent) or the nitrene product (Scheme 1c), because this D is quite larger than that of a molecule with a smaller size and with similar functional groups such as *p*-coumaric acid (4.4×10^{-10} m^2 s^{-1}).[33] We tentatively attributed this component to the diffusion of molecules induced by the photoreaction of the hydrolyed sulfo-HSAB, 4-azidobenzoic acid (Scheme 1d).[34] Although we have not exactly identified the diffusing species for these D_1 and D_2 components, this assignment is not important for the main purpose in this paper, measurement of the protein diffusion. The last diffusion component (D_3) may be attributed to that of sulfo-HSAB and the nitrene product. We consider that, since the D values of these molecules are expected to be similar, only one decay component was observed in the TG signal. The positive sign of this component, δn_3, indicates that the contribution of the

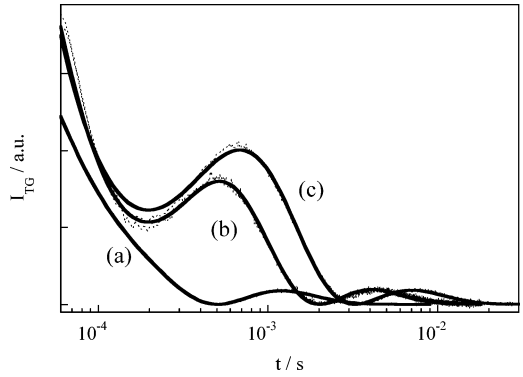


Figure 1. 1. TG signals (dotted lines) of the sulfo-HSAB solution (200 μM) (without any protein) at various grating wavenumbers (q^2): (a) 2.6×10^{12}, (b) 8.8×10^{11}, and (c) 4.6×10^{11} m^{-1}. The fitting curves based on eq 4 are also shown (solid lines).

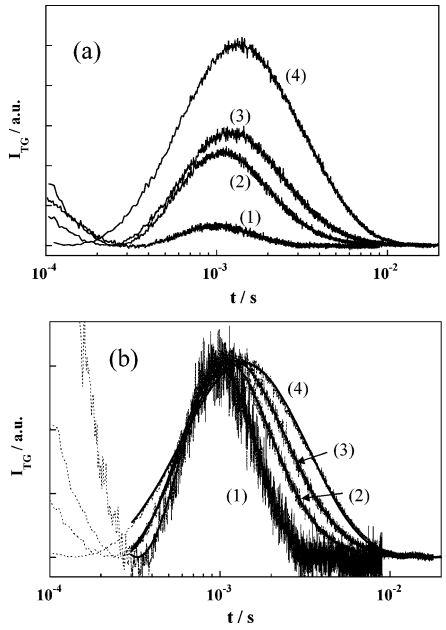


Figure 2. (a) TG signals of sulfo-HSAB with holo-Mb at [HSAB] = 500 μM and different holo-Mb concentrations: (1) 1, (2) 10, (3) 100, and (4) 500 μM. (b) The intensities of the diffusion signal are normalized.

nitrene product, which has positive δn_p, is larger than that of the reactant, sulfo-HSAB, which has negative δn_r.

4.2. TG Signal of Holo-Mb. When holo-Mb is added to the sulfo-HSAB solution, the TG signal intensity increases, and the time constant changes significantly (Figure 2). This change should be explained by the photoreaction of sulfo-HSAB with the protein. When sulfo-HSAB is mixed with the protein, the acylation reaction takes place with its primary amines of the protein (Scheme 1a), and upon the photoirradiation, the azide group dissociates to produce N_2 and the nitrene product (Scheme 1b). From the temporal profile of the signal, we expect that D of the diffusing species can be determined. However, since we found that the temporal shape depends on the mixing ratio of the concentration of the protein to that of sulfo-HSAB, we first examined the appropriate concentrations for a proper D measurement of the protein.

We measured the TG signals at various holo-Mb concentrations ([hMb]) with a constant sulfo-HSAB concentration ([HSAB]) of 500 μM in phosphate buffer (10 mM). The observed TG signals at different protein concentrations are shown in Figure 2. The increase of the signal intensity upon adding the protein to the solution depends on [hMb]. For clearly

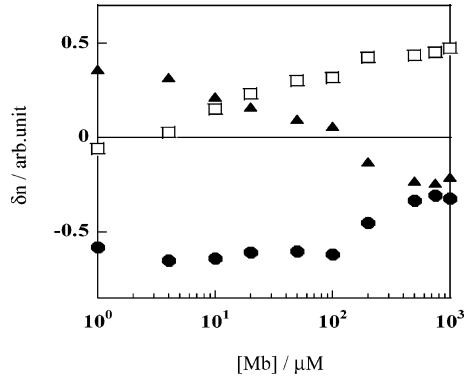


Figure 3. Relative amplitudes (δn) of the diffusion components of D_2 (δn_2, ●), D_3 (δn_3, ▲), and D_{hMb} (δn_{hMb}, □) determined from the fits of the TG signals at [HSAB] = 500 μM and various holo-Mb concentrations (1−1000 μM) (e.g., Figure 2) by eq 5. The amplitudes are normalized by $|\delta n_2| + |\delta n_3| + |\delta n_{hMb}| = 1$.

showing the profile, the signal intensities are normalized at the peak maximum intensity in Figure 2b. It is clear from the figure that the slow decay component becomes stronger with increasing protein concentration. Since the molecular size of holo-Mb is much larger than those of the other species in the solution, the slowest diffusion component should be attributed to protein diffusion. The observed concentration dependence suggests that the decay component of the signal should be analyzed by two components, the D_3q^2 component in eq 4, which was observed without the protein, and the protein signal. This contribution has to be estimated for the analysis.

To avoid ambiguity of the fitting, we fitted the rise−decay curve by a sum of the three exponential terms

$$I_{TG} = \alpha\{-\delta n_2 \exp(-D_2q^2t) + \delta n_3 \exp(-D_3q^2t) + \delta n_{hMb} \exp(-D_{hMb}q^2t)\}^2 \quad (5)$$

where the first and second terms are those in eq 4 and D_{hMb} is the diffusion coefficient of holo-Mb. Since the other components in eq 4 ($\delta n_k \exp(-kt)$ and $\delta n_1 \exp(-D_1q^2t)$) appear in a much earlier time range than that of the protein diffusion, we can safely neglect these terms for the determination of D of the protein. Furthermore, D_2 and D_3 are fixed to the values obtained from the sulfo-HSAB solution ($D_2 = 8.0 \times 10^{-10}$ m^2 s^{-1}, $D_3 = 3.0 \times 10^{-10}$ m^2 s^{-1}). Even through the use of this restriction for the fitting, determined D_{hMb}, in particular, at low concentrations of Mb was ambiguous, because of the small amplitudes of δn_{hMb}. Hence, we first fitted the signal to determine D_{hMb} at rather high concentrations, and this value (1.0×10^{-10} m^2 s^{-1}) was used for the fitting in the whole concentration range. The signal can be reproduced very well by this equation with these restrictions, and the determined relative refractive index changes at various protein concentrations are plotted in Figure 3.

This dependence shows that, at low [hMb], the contribution of holo-Mb is small and most of the TG signal comes from sulfo-HSAB which is not bound to the protein (free sulfo-HSAB). The magnitude of the protein contribution increases with increasing [hMb], and the relative contributions of δn_2 and δn_3 decrease. In a higher concentration range of Mb, the sign of δn_2 changes from positive to negative. At present, we do not know the exact mechanism of this sign change, but we may explain this phenomenon tentatively as follows. The negative sign of δn_3 at a higher concentration suggests that free sulfo-HSAB was still present at the high [Mb] (probably because not all molecules in the reagent can associate with the protein) so that it contributed to the negative δn_3. At the same time, the

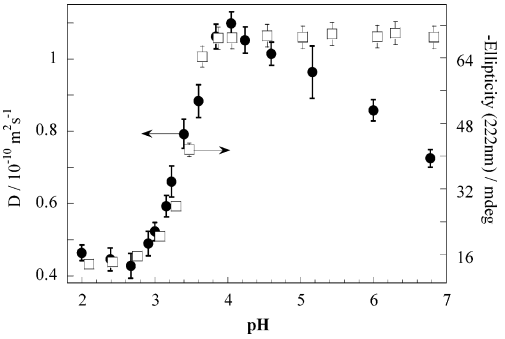


Figure 4. pH dependences of the diffusion coefficient (●) and ellipticity at 222 nm (□) of holo-Mb.

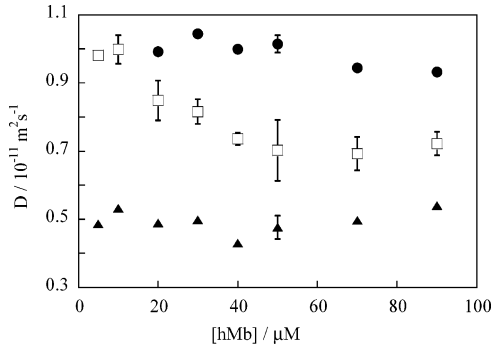


Figure 5. [hMb] dependence of D_{hMb} (5−90 μM) at various pH values; pH 7 (□), pH 5 (●), and pH 2 (▲).

contribution of the nitrene product giving positive δn decreased with increasing [Mb]. The decrease of the nitrene contribution may be explained by a recombination reaction of the nitrene product with Mb at a high [Mb] (Scheme 1e), because the combined product diffuses with the same rate as that of Mb not sulfo-HSAB (δn_3). Hence, the sign of δn_3 changed with increasing [Mb].

We found that when [hMb] is more than half of [HSAB] the decay component is mostly composed of the holo-Mb contribution and the analysis becomes much easier. Therefore, hereafter, we used the same concentration of the protein with [HSAB] for the *D* measurement of the proteins so that the dominant contribution to the signal decay comes from the protein diffusion.

4.3. pH Dependence of *D* of Holo-Mb. Using the method described in the previous section, we measured D_{hMb} (50 μM in aqueous solution) at various pH values and plotted in Figure 4. We found that, with decreasing pH, *D* slightly increases in the range of 7 > pH > 5, shows a plateau in the range of 5 > pH > 4, decreases suddenly in the range of 4 > pH > 3, and becomes almost constant again in the range of 3 > pH > 2.

First, we investigate the origin of the slight increase in the pH range of 7 > pH > 5. Since this pH range is close to the isoelectric point of holo-Mb (pH ≈ 7), we examined the protein−protein interaction in this pH range by measuring the concentration dependence of holo-Mb. We found that *D* gradually increases with decreasing [hMb] (Figure 5). Around 10 μM, *D* becomes 1.0 (±0.1) × 10^{-10} m^2 s^{-1}, which is close to *D* in the range of 5 > pH > 4 and the reported *D* of the native holo-Mb (0.9−1.1 × 10^{-10} m^2 s^{-1}).[12,13,35,36,39] Therefore, we concluded that the smaller *D* at the 50 μM concentration at pH ≈ 7 compared to that at pH 4 is due to the attractive interaction between holo-Mb molecules.[37] No [hMb] dependence was observed in *D* at lower pH values, such as pH 5 (Figure 5). At this low pH, the surface of holo-Mb should be positively charged, and the electric repulsion between the proteins weakens the attractive intermolecular interaction. The observed increase in *D* from pH 7 to pH 5 may represent the gradual decrease of the attractive interaction by increasing the surface charge. After lowering from pH 5, we did not observe concentration dependence in *D* as shown in Figure 5 for pH 2. Considering the signal-to-noise ratio of the signal and this concentration dependence, we used [hMb] = 50 μM for most cases hereafter.

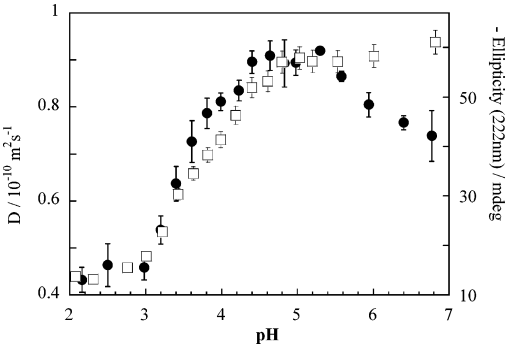


Figure 6. pH dependences of D_{aMb} (●) and ellipticity at 222 nm (□) of apo-Mb.

The pH range of 3 < pH < 4, in which significant *D* change is observed, is close to the pH range of acid denaturation of holo-Mb.[20,21] For comparison, the CD spectrum of holo-Mb was measured at various pH values, and the ellipticity at 222 nm, which is a good indicator of the α-helix content, is plotted in Figure 4. Both *D* and ellipticity depend on pH similarly.

At pH 2, the secondary structure of holo-Mb is mostly destroyed, and *D* decreases to 0.45 × 10^{-10} m^2 s^{-1}, which is quite smaller than that of the native one. This decrease is not due to any dimer or oligomer formation, because we did not observe any concentration dependence of *D* (Figure 5). The significant decrease of *D* upon unfolding was previously observed also for cytochrome c.[3]

4.4. pH Dependence of *D* of Apo-Mb. The essential feature of the pH dependence of *D* of apo-Mb (D_{aMb}) is similar to that of D_{hMb}. With a decrease in pH, *D* increases initially, decreases in the range of 5 > pH > 3, and becomes a constant in the range of 3 > pH > 2. As explained above for the case of holo-Mb, the initial increase of *D* is attributed to the interprotein interactions of apo-Mb. The D_{aMb} value at pH 5 should be that for the native apo-Mb. The *D* value of apo-Mb at pH 2 should also be that for an unfolded apo-Mb not aggregation formation, because *D* does not depend on the concentration in the experimental concentration range. The denaturation curve monitored by the ellipticity at 222 nm is shown in Figure 6, which agrees quite well with those reported previously.[20,21,38] This profile indicates that the denaturation of apo-Mb is not a single step but an intermediate species is involved around pH 4. An interesting point in this profile is that this pH dependence of *D* is different from that of the secondary structure monitored by the ellipticity.

5. Discussion

5.1. *D* of Native and Unfolded States. As mentioned in the Introduction, *D* is determined by the size and shape of a solute and the intermolecular interaction between the solute and the solvent. At first, we tried to explain the observed pH dependence of *D* from the viewpoint of the radius change of the proteins. Sometimes, the radius of the molecule is calculated from the *D* value based on the Stokes−Einstein relationship, and the radius is called the hydrodynamic radius (R_h). The smaller *D* means larger R_h. There are also some attempts to make a relation

between R_g and D. One of the recent examples is reported by He et al.[39] They have proposed an empirical equation that can reproduce D (in $m^2\ s^{-1}$) of a variety of native proteins from a radius of gyration (R_g in Å) and the molecular weight (M in g mol^{-1}) in a solution of temperature T (in K) and viscosity η (in Pa s)

$$D = 6.85 \times 10^{-15}\, T/(\eta\sqrt{M^{1/3}R_g}) \tag{6}$$

Using the average value of the reported R_g for each state,[12,40−42] we calculated D values of the native holo- and apo-Mb and found that the calculated D values ($D_{hMb} = 0.94 \times 10^{-10}$ $m^2\ s^{-1}$, $D_{aMb} = 0.90 \times 10^{-10}\ m^2\ s^{-1}$) agree with those obtained by the TG measurements ($D_{hMb} = 1.04 \times 10^{-10}\ m^2\ s^{-1}$, $D_{aMb} = 0.90 \times 10^{-10}\ m^2\ s^{-1}$) within the reliability of this equation (15%). This result indicates that the absence of heme affects only the size of Mb. In other words, the interaction between Mb and water is not affected by the presence (or depletion) of the heme.

However, if we simply apply this eq 6 to the unfolded holo- and apo-Mb, then the calculated D values ($D = 0.73 \times 10^{-10}$ $m^2\ s^{-1}$, for both holo- and apo-Mb) do not reproduce the observed values at all ($D = 0.45 \times 10^{-10}\ m^2\ s^{-1}$). This is not surprising, because the above equation is optimized for the native state of the protein. Indeed, Wilkins et al. have suggested an empirical equation that can predict a hydrodynamic radius (R_h) of an unfolded protein from its number of residues [13]

$$R_h = 2.21N^{0.57} \tag{7}$$

Through the use of this equation, R_h of unfolded Mb ($N = 153$) is calculated to be 3.89 nm, which gives a D value of $0.55 \times 10^{-10}\ m^2\ s^{-1}$ in aqueous solution at 298 K. This value is in agreement with the D value of acid-induced unfolded Mb (pH 2). The different equations for different states mean that the factor which governs the diffusion is different depending on the conformational state. Furthermore, the pH dependencies of D and R_g are different from each other as shown later. These facts indicate that D and R_g provide different information on the protein conformation.

Contrary to the clearly defined R_g from the geometrical data, there are many factors affecting R_h (or, equivalently, D). The shape of the molecule is one of the factors governing D. Perrin theoretically derived the friction for molecules that have significantly different shapes from spherical. For example, the ratio of D of a prolate molecule (D_{prl}), which has a semimajor axis a and a minor axis b, to D of a spherical molecule (D_0), which has the same volume, is expressed as[16]

$$\frac{D_{prl}}{D_0} = \frac{\ln[P + (P^2 - 1)^{1/2}]}{P^{-1/3}(P^2 - 1)^{1/2}} \tag{8}$$

as a function of $P = a/b$. Also that of an oblate molecule (D_{obl}) is given by

$$\frac{D_{obl}}{D_0} = \frac{P^{2/3}\arctan[(P^2 - 1)^{1/2}]}{(P^2 - 1)^{1/2}} \tag{9}$$

To explain the experimentally observed ratio (D(pH = 5)/D(pH = 2) = 2.3 for holo-Mb, 2.0 for apo-Mb) by molecular shape difference alone, the random coil state should have an extremely elongated shape ($a/b > 20$). However, such an extremely elongated shape is unreasonable, because the unfolded apo-Mb at pH 2 has enough flexibility to access various conformations (high conformational entropy), and its averaged shape over the possible conformational ensemble should be near a random coil.[23,43] Therefore, we can conclude that the molecular shape cannot be the main factor to determine the difference in D between the two states (folded and unfolded Mb).

However, the interaction between a protein and the water molecules affects the friction. When the protein is completely denatured and consequently the solvent-exposed surface area of the protein increases, it should cause an increase in the number of binding sites for water molecules on a protein. Hence the solvent molecules will interact with the protein surface, and consequently the protein molecule feels more frictional drag in the solvent, and its diffusion is more restrained. By using microwave dielectric spectroscopy, Kamei et al.[44] have observed an appreciable increase in the hydration number of water in the acid-induced unfolded state of apo-Mb. The denaturation curve monitored by D may be consistent with this change.

5.2. Conformation of Mb along the pH Denaturation Curve. The similarity of the denaturation curves monitored by the ellipticity (Figure 4) and D for the holo-Mb case may reflect the fact that the intermolecular interaction between the protein and the water molecules is enhanced by the destruction of the secondary structure. A similar feature was observed previously for cytochrome c [3] and poly-L-glutamic acid (PLG).[44] In the both cases, the denaturation curves monitored by the ellipticity agree well with that of D. The change in the molecular shape was excluded from the cause of the D change for both cases. Furthermore, the change in partial molar volume of cytochrome c was negligible during the protein unfolding.[3] We interpreted the D change during the unfolding of the α-helices in terms of smaller frictions of the α-helices with water than those of random-coiled macromolecules; that is, D can be a good indicator reflecting the conformation of macromolecules in the solution phase.

However, the denaturation curve monitored by the ellipticity is different from that of D for apo-Mb. This non-coincidence of the two transition curves indicates that even when the α-helices are partially destroyed the friction for diffusion does not change to the same extent.

The pH denaturation profile of apo-Mb monitored by the ellipticity cannot be explained by the two-state model, but it indicates the presence of an equilibrium intermediate species. This intermediate has been attracting many researchers, and the properties were investigated by using techniques such as NMR,[22,46] SAXS,[12,40] fluorescence,[43] and Raman measurements[21] of Trp and Tyr. These investigations showed that this intermediate has a compact hydrophobic core, which is composed of the AGH helices. From several characteristic properties, such as compact size, preservation of considerable amount of native secondary structure, and disordered tertiary structure, this intermediate state has been characterized as a MG conformation. Interestingly, it was proposed that this equilibrium intermediate possesses a similar structure to that of a kinetic intermediate observed in a pH-jump experiment.[22]

The larger R_g of this intermediate (23.0 Å)[12,40,42] compared with R_g of the native form (19.3 Å)[12,40−42] indicates that a part of the side chains or the backbone structure is swollen by water at the intermediate. We calculated D of this intermediate conformation by the empirical equation proposed by He et al.[39] with the reported R_g. Interestingly, the calculated D value ($0.82 \times 10^{-10}\ m^2\ s^{-1}$) is very close to the experimentally observed value. This agreement suggests that water molecules causing the swollen structure are "trapped" inside the protein on the basis of the following consideration. If water molecules can

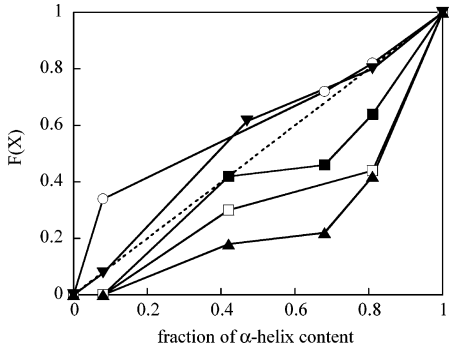


Figure 7. Plot of various folding parameters of apo-Mb calculated by eq 11 versus the α-helix content measured by the ellipticity at 222 nm: $F(D)$ from the present study (▼), $F(R_g)$ (○), $F(m(\text{urea}))$ (■), $F(m(\text{GdnHCl}))$ (□), and $F(\Delta C_p)$ (▲) from ref 47.

penetrate freely through the space of the loosened structure without any additional interaction, then D of such a protein is expected to be similar to the native state. However, generally, since the unfolding of the α-helices induces intermolecular interaction between the protein and water molecules, the unfolding of the α-helices leads to a significant decrease in D as shown above. The fact that the observed D value can be reproduced by the increase of R_g suggests that the diffusion is determined only by the size; that is, the water molecules inside the protein are rather isolated, and the intermolecular interaction between the protein and the water as a cause of the friction is similar to that of the native form.

Extent of folding may be measured using various physical quantities. For example, R_g or α-helix content from the CD measurement has been used as good indicators for the folding. The m-value, which is defined by the denaturant-concentration dependence of the free energy change upon unfolding (ΔG_U)

$$\Delta G_U = \Delta G_U(H_2O) - m[\text{denaturant}] \qquad (10)$$

($\Delta G_U(H_2O)$, the free energy change in the absence of denaturant) has been frequently used as an indicator of a change in the solvent-accessible surface area of the protein during the unfolding transition. Similarly, the isobaric heat capacity change (ΔC_p) is also considered to be a suitable indicator representing the hydrophobic interaction. If we use a physical property, X, as an indication of folding, then the extent of folding may be defined as

$$F(X) = \frac{X - X(\text{U})}{X(\text{native holo-Mb}) - X(\text{U})} \qquad (11)$$

where X(U) and X(native holo-Mb) are X of a denaturant-induced unfolded state and native holo-Mb.

The conformational transition of apo-Mb has been studied by X-ray scattering, CD, and differential scanning calorimetry (DSC) by Nishii et al.[47] The extent of folding using R_g ($F(R_g)$), ΔC_p ($F(\Delta C_p)$), and the m-value ($F(m)$) were plotted versus the α-helix content, which was estimated from the ellipticity at 222 nm obtained by CD measurement (Figure 7). It was found that the extent of folding monitored by these properties was not parallel with it. Both ΔC_p and the m-value changed sharply in the final stage of the folding (MG → native holo-Mb). Since these properties are related to the change in solvent-accessible surface (SAS) upon unfolding,[48] the changes in ΔC_p and the m-value suggest that substantial hydrophobic interactions are not formed until the transformation from MG to the native form. In other words, the water molecules inside the protein were squeezed out at this stage.

For comparison, the extent of folding from D ($F(D)$) measured in the present study is plotted in Figure 7. Interestingly, the profile of $F(D)$ does not correlate with the α-helix content but deviates to the opposite side of those of $F(\Delta C_p)$ and $F(m)$. This trend indicates that D changes mainly during the transition from the unfolded state to MG, whereas the change is rather minor during the MG → native transition. On the basis of the intermolecular interaction picture of the D change, one may conclude that the hydrogen bonding network changes from intermolecular to intramolecular at the unfolded → MG step; that is, the intramolecular hydrogen bonding is dominant in MG. However, from the profiles of $F(\Delta C_p)$ and $F(m)$, it was suggested that MG is hydrated. These two features can be consistently explained if rather small numbers of water molecules are trapped inside the protein interior in the MG state.

Such trapped water molecules have been considered to be important in forming a compact denatured MG state of α-lactoalbumin,[49] which is stabilized by nonspecific hydrophobic interaction. This hydrophobic interaction between hydrated hydrophobic groups called solvent-separated hydrophobic interaction may be important at an early stage of protein folding. The present observation showed that a similar solvent-separated hydrophobic interaction exists also for the pH 4 intermediate of apo-Mb. Furthermore, a recent study on the protein folding process of cytochrome c showed that the hydrogen bonding network changes from intermolecular to intramolecular before the dominant α-helix formation.[3] The picture of the isolated water clusters inside the hydrophobic core is also consistent with this intermediate state of cytochrome c folding.

6. Conclusion

We have measured the diffusion coefficient (D) of myoglobins (holo- and apo-Mb) at various pH values. The pH dependence of D agrees well with that measured by the ellipticity for holo-Mb, whereas the difference is more evident in the case of apo-Mb. This result shows that holo-Mb unfolds by acid more cooperatively than apo-Mb does. This difference may be due to the structural stabilization caused by the F helix of holo-Mb that is absent in apo-Mb. In the range of 7 > pH > 5.5, we observed the increase of D with decreasing pH for both holo- and apo-Mb. This pH dependence is explained by the interprotein interaction at the isoelectric point of myoglobin. This interaction is weakened by the stronger electrostatic repulsion with decreasing pH. The acid denaturation of apo-Mb may be summarized as follows. In the native → MG process, R_g slightly increases by extending some side chains and keeping a hydrophobic core of AGH helices. At this stage, water molecules enter into the protein interior to be swollen. These water molecules contribute to the increase in solvent-accessible surface as suggested by the ΔC_p, m-value, and Raman cross section measurements. However, since these molecules are rather isolated from the bulk phase, D does not change so much from the native D value. In the further MG → unfolded denaturation stage, the backbone conformation is extended to increase R_g, and the intermolecular interaction with the solvent water molecules is enhanced to decrease D.

References and Notes

(1) Pain, R. H. *Mechanism of Protein Folding*, 2nd ed.; Oxford University Press: New York, 2000.
(2) Nada, T.; Terazima, M. *Biophys. J.* **2003**, *85*, 1876.
(3) Nishida, S.; Nada, T.; Terazima, M. *Biophys. J.* **2004**, *87*, 2663.
(4) Nishida, S.; Nada, T.; Terazima, M. *Biophys. J.* **2005**, *89*, 2004.
(5) Eitoku, T.; Nakasone, Y.; Matsuoka, D.; Tokutomi, S.; Terazima, M. *J. Am. Chem. Soc.* **2005**, *127*, 13238.

(6) Cussler, E. L. *Diffusion; Mass Transfer in Fluid Systems*, 2nd ed.; Cambridge University Press: New York, 1997.
(7) Srinivas, G.; Bhattacharyya, S.; Bagchi, B. *J. Chem. Phys.* **1999**, *110*, 4477.
(8) Terazima, M. *Acc. Chem. Res.* **2000**, *33*, 687.
(9) Terazima, M.; Hirota, N. *J. Chem. Phys.* **1993**, *98*, 6257.
(10) Terazima, M,; Okamoto, K.; Hirota, N. *J. Chem. Phys.* **1995**, *102*, 2506.
(11) Tyrrell, H. J. V.; Harris, K. R. *Diffusion in Liquids: A Theoretical and Experimental Study*; Butterworths: London, 1984.
(12) Gast, K.; Damaschun, H.; Misselwitz, R.; Müller-Frohne, M.; Zirwer, D.; Damaschun, G. *Eur. Biophys. J.* **1994**, *23*, 297.
(13) Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. *Biochemistry* **1999**, *38*, 16424.
(14) Nesmelova, I. V.; Skirda, V. D.; Fedotov, V. D. *Biopolymers* **2002**, *63*, 132.
(15) Meyvis, T. K. L.; De Smedt, S. C.; Van Oostveldt, P.; Demeester, J. *Pharm. Res.* **1999**, *16*, 1153.
(16) van Holde, K. E.; Johnson, W. C.; Ho, P. S. *Principles of Physical Biochemistry*, int. ed.; Prentice Hall: Upper Saddle River, NJ, 1998; p 213.
(17) Kawahara, K. *Biochemistry* **1969**, *8*, 2551.
(18) Baden, N.; Terazima, M. *Chem. Phys. Lett.* **2004**, *393*, 539.
(19) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, 4th ed.; W. H. Freeman and Company: New York, 2005; p 132.
(20) Goto, Y.; Fink, A. L. *Methods Enzymol.* **1994**, *232*, 3.
(21) Chi, Z.; Asher, S. A. *Biochemistry* **1999**, *38*, 8196.
(22) Jennings, P. A.; Wright, P. E. *Science* **1993**, *262*, 892.
(23) Griko, Y. V.; Privalov, P. L.; Venyaminov, S. Y.; Kutyshenko, V. P. *J. Mol. Biol.* **1988**, *202*, 127.
(24) Eliezer, D.; Chung, J.; Dyson, J.; Wright, P. E. *Biochemistry* **2000**, *39*, 2894.
(25) Hughson, F. M.; Barrick, D.; Baldwin, R. L. *Biochemistry* **1991**, *30*, 4113.
(26) Eichler, H. J.; Gunter, P.; Pohl, D. W. *Laser Induced Dynamic Gratings*; Springer-Verlag: Berlin, 1986.
(27) Kogelnik, H. *Bell. Syst. Tech. J.* **1969**, *48*, 2909.
(28) Terazima. M. *Adv. Photochem.* **1998**, *24*, 255.
(29) Yeung, C. W. T.; Moule, M. L.; Yip, C. C. *Biochemistry* **1980**, *19*, 2196.
(30) Buchmueller, K. L.; Hill, B. T.; Platz, M. S.; Weeks, K. M. *J. Am. Chem. Soc.* **2003**, *125*, 10850.
(31) *Landolt-Börnstein*; Springer-Verlag: Berlin, 1968; Band II, Teil 5, S. 611.
(32) Kwok, W. M.; Chan, P. Y.; Phillips, D. L. *J. Phys. Chem. B* **2004**, *108*, 19068.
(33) Takeshita, K.; Hirota, N.; Terazima, M. *J. Photochem. Photobiol., A* **2000**, *134*, 103.
(34) Lomant, A. J.; Fairbanks, G. *J. Mol. Biol.* **1976**, *104*, 243.
(35) Bismuto, E.; Gratton, E.; Lamb, D. C. *Biophys. J.* **2001**, *81*, 3510.
(36) Riveros-Moreno, V.; Wittenberg, J. B. *J. Biol. Chem.* **1972**, *247*, 895.
(37) The concentration dependence at pH 7 shows that *D* becomes almost constant over the [hMb] > 50 μM range. From the Stokes−Einstein relationship, we expect that *D* decreases by a factor of $2^{1/3} = 1.26$ by dimer formation. The constant *D* value in the [hMb] > 50 μM range is close to the *D* of the dimer. Hence, we may suggest that the intermolecular interaction of holo-Mb in this concentration range is not strong enough to make a larger aggregate, but it may form a dimer. Such a formation of a dimer has been observed previously for myoglobin (Van Den Oord, A. H. A.; Wesdorp, J. J.; Van Dam, A. F.; Verheij, J. A. *Eur. J. Biochem.* **1969**, *10*, 140.) and also for cytochrome c.[3] When the concentration of Mb and sulfo-HSAB was increased further (>350 μM), we found that the final decay rate of the TG signal became slower depending on the time after the mixing of the solutions. We consider that the cross-linking between Mb's occurs by sulfo-HSAB under this condition. Within the concentration range that we used in this study, we did not observe this time-dependent change of the decay.
(38) Goto, Y.; Calciano, L. J.; Fink, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 573.
(39) He, L.; Niemeyer, B. *Biotechnol. Prog.* **2003**, *19*, 544.
(40) Kataoka, M.; Nishii, I.; Fujisawa, T.; Ueki, T.; Tokunaga, F.; Goto, Y. *J. Mol. Biol.* **1995**, *249*, 215.
(41) Uzawa, T.; Akiyama, S.; Kimura, T.; Takahashi, S.; Ishimori, K.; Morishima, I.; Fujisawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1171.
(42) Nishii, I.; Kataoka, M.; Tokunaga, F.; Goto, Y. *Biochemistry* **1994**, *33*, 4903.
(43) Rischel, C.; Thyberg, P.; Rigler, R.; Poulsen, F. M. *J. Mol. Biol.* **1996**, *257*, 877.
(44) Kamei, T.; Oobatake, M.; Suzuki, M. *Biophys. J.* **2002**, *82*, 418.
(45) Inoue, K.; Baden, N.; Terazima, M. *J. Phys. Chem. B* **2005**, *109*, 22623.
(46) Hughson, F. M.; Wright, P. E.; Baldwin, R. L. *Science* **1993**, *249*, 1544.
(47) Nishii, I.; Kataoka, M.; Goto, Y. *J. Mol. Biol.* **1995**, *250*, 223.
(48) Alonso, D. O. V.; Dill, K. A. *Biochemistry* **1991**, *30*, 5974.
(49) Arai, M.; Kuwajima, K. *Adv. Protein Chem.* **2000**, *53*, 209.