Interconnection of Salt-induced Hydrophobic Compaction and Secondary Structure Formation Depends on Solution Conditions

REVISITING EARLY EVENTS OF PROTEIN FOLDING AT SINGLE MOLECULE RESOLUTION

Shubhasis Haldar and Krishnananda Chattopadhyay

From the Protein Folding and Dynamics Laboratory, Structural Biology and Bioinformatics Division, Indian Institute of Chemical Biology, Council for Scientific and Industrial Research, 4 Raja S.C. Mullick Rd., Kolkata 700032, India

Received for publication, October 23, 2011, and in revised form, January 22, 2012 Published, JBC Papers in Press, February 2, 2012, DOI 10.1074/jbc.M111.315648

Background: Early events of protein folding are difficult to follow. Rapid kinetics measurements with cytochrome c using submillisecond continuous flow mixing devices suggest simultaneous formation of a compact collapsed state and secondary structure. These data seem to indicate that collapse formation is guided by specific short and long range interactions (heteropolymer collapse). A contrasting interpretation has also been proposed, which suggests that the collapse formation is rapid, nonspecific, and a trivial solvent related compaction, which could as well be observed by a homopolymer (homopolymer collapse). We address this controversy using fluorescence correlation spectroscopy (FCS), which enables us to monitor the salt-induced compaction accompanying collapse formation and the associated time constant directly at single molecule resolution. In addition, we follow the formation of secondary structure using far UV CD. The data presented here suggest that both these models (homopolymer and heteropolymer) could be applicable depending on the solution conditions. For example, the formation of secondary structure and compact state is not simultaneous in aqueous buffer. In aqueous buffer, formation of the compact state occurs through a two-state co-operative transition following heteropolymer formalism, whereas secondary structure formation takes place gradually. In contrast, in the presence of urea, a compaction of the protein radius occurs gradually over an extended range of salt concentration following homopolymer formalism. The salt-induced compaction and the formation of secondary structure take place simultaneously in the presence of urea.

What happens in the early stage of protein folding remains an interesting unsolved problem. Rapid kinetics measurements with cytochrome c using submillisecond continuous flow mixing devices suggest simultaneous formation of a compact collapsed state and secondary structure. These data seem to indicate that collapse formation is guided by specific short and long range interactions (heteropolymer collapse). A contrasting interpretation has also been proposed, which suggests that the collapse formation is rapid, nonspecific, and a trivial solvent related compaction, which could as well be observed by a homopolymer (homopolymer collapse). We address this controversy using fluorescence correlation spectroscopy (FCS), which enables us to monitor the salt-induced compaction accompanying collapse formation and the associated time constant directly at single molecule resolution. In addition, we follow the formation of secondary structure using far UV CD. The data presented here suggest that both these models (homopolymer and heteropolymer) could be applicable depending on the solution conditions. For example, the formation of secondary structure and compact state is not simultaneous in aqueous buffer. In aqueous buffer, formation of the compact state occurs through a two-state co-operative transition following heteropolymer formalism, whereas secondary structure formation takes place gradually. In contrast, in the presence of urea, a compaction of the protein radius occurs gradually over an extended range of salt concentration following homopolymer formalism. The salt-induced compaction and the formation of secondary structure take place simultaneously in the presence of urea.

Several experimental and theoretical techniques have been devised to study dynamics of the unfolded state and to monitor early events of protein folding (1–4). Many of these studies involve measurements of rapid kinetics using mixing techniques with sub-millisecond and μs time resolutions (5–8). Cytochrome c has been found to be an attractive model system for these experiments (9).

Two long-standing controversial questions related to the nature and time scale of the initial events of protein folding have been addressed by many of these rapid kinetics measurements. First, is it the hydrophobic collapse, which occurs first, followed by the formation of the secondary structure, or do they occur simultaneously? Second, is the initial event of protein folding directed by specific interactions between the participating amino acids, or is it nonspecific? Rapid kinetics measurements show barrier limited formation of a native-like collapsed states in the early time scale (6, 10). These experiments and other data seem to suggest that hydrophobic collapse and secondary structure formation take place simultaneously (11). These studies indicated that collapse is guided by specific short and long range interactions and barrier-limited (heteropolymer collapse) (6, 9–10). A contrasting interpretation was brought up by Sosnick et al. (12, 13), who suggested that the initial collapse is nonspecific. According to this interpretation, the initial collapse is a trivial compaction, which any homopolymer would experience when the denaturant gets diluted by aqueous buffer (homopolymer collapse). The controversy has been found difficult to resolve because of several reasons. First, the number of experimental techniques, which could probe early change in conformation and μsec dynamics in the unfolded state, is limited. Second, pathways of protein folding are heterogeneous with complications arising from parallel pathways, trapped intermediates, and misfolding.

We addressed these questions by making use of fluorescence correlation spectroscopy (FCS)2 and far UV CD to study the salt-induced collapse and secondary structure formation of cytochrome c from Saccharomyces cerevisiae. Bovine serum albumin also was used to verify some of the key experimental

2 The abbreviations used are: FCS, fluorescence correlation spectroscopy; TMR, tetramethyl rodamine 5 maleimide; MEM, maximum entropy method.
results on a different protein system. FCS is a unique experimental technique that can measure conformation of the unfolded state and its fluctuation dynamics in the μs time scale using the same experimental data. Additionally, FCS offers single molecule sensitivity, which could efficiently isolate folding events from any competing undesired processes such as aggregation.

Using FCS, we observed the presence of a fluctuation dynamics between an extended and a collapsed conformer of cytochrome c unfolded at pH 2. The time constant of their interconversion was 50 μs. The equilibrium population of collapsed conformer increased upon the addition of salt as monitored by a decrease in the hydrodynamic radius (r_h). Using far UV CD and FCS, we independently monitored the formation of secondary structure and that of the collapsed state, respectively. In aqueous buffer, the formation of collapsed state followed a cooperative two-state transition occurring at low salt concentration (heteropolymer collapse), although secondary structure forms gradually over a large range of salt concentration. In aqueous buffer, these two events are not simultaneous. A salt-induced compaction of the hydrodynamic radius has been observed in the presence of urea. However, the contraction of the unfolded chain and secondary structure formation occurred simultaneously in urea. In the presence of urea, the contraction followed homopolymer formalism.

**EXPERIMENTAL PROCEDURES**

Iso-1-cytochrome c (C2436), sodium perchlorate (CAS-7601-89-0), and urea (CAS-13-6) were obtained from Sigma. Alexa Fluor 488-maleimide and tetramethyl rhodamine 5 (7601-89-0), and urea (CAS-13-6) were obtained from Molecular Probes (Eugene, OR). All of the other reagents used were of the highest available grade.

Far UV CD spectra were carried out using a Jasco J715 spectropolarimeter. Experiments were performed typically with 9 μM TMR-maleimide-labeled cytochrome c-TMR. Far-UV CD spectra were recorded over a 200–250-nm range, and each spectrum was averaged using 10 repeat scans. Steady state fluorescence experiments were carried out with TMR-labeled protein using a PTI fluorescence spectrometer (Photon Technologies Intl.). The fluorescence emission spectra were recorded between 560–650 nm using an excitation wavelength of 540 nm. Final concentration of the labeled cytochrome c used for the steady-state fluorescence was 1 μM.

MALDI mass spectrometry was carried out using a 4800 MALDI-TOF-TOF mass spectrometer (Applied Biosystems) equipped with a microchannel plate detector. The delay time used was 500 ns. Mass spectra were acquired by averaging 600 to 800 shots. Labeled cytochrome c fragment of 1 μM concentration was used for the mass spectrometry measurements.

Cytochrome c from S. cerevisiae contains one free cysteine (Cys-102), which was labeled with TMR using published procedures (14). Labeling of BSA with Alexa Fluor 488-maleimide was carried out using an identical method. Labeling of cytochrome c with TMR does not result in any significant effect on the structure, conformation, and folding of the protein (14). FCS experiments were carried out with a ConfoCor 3 LSM setup (Carl Zeiss, Evotec, Jena, Germany) using a 40× water immersion objective. The detailed experimental procedure has been published previously (15, 16).

For FCS data analysis, correlation function curves observed with the FCS experiments were analyzed using different models as described below.

**Model 1 (Simple Diffusion Model)—**For a single component system without any conformational event, the diffusion time (τ_D) of a fluorophore (for example, the free dye, TMR or Alexa Fluor 488-maleimide-labeled bovine serum albumin) and the number of particles in the observation volume (N) can be calculated by fitting the correlation function (G(τ)) to Equation 1 (17),

\[
G(\tau) = 1 + \frac{1}{N} \left(1 + \frac{\tau}{\tau_D}\right)^{-0.5}
\]  
(Eq. 1)

where S is the structure parameter, the depth-to-diameter ratio of the Gaussian observation volume.

**Model 2 (Simple Diffusion with Conformation Fluctuation Model)—**Because the correlation functions obtained with cytochrome c-TMR at pH 2 could not be analyzed using model 1 described above, we needed to use model 2 for the successful representation of the data. This model (model 2) assumes a single diffusing species (with the diffusion time of τ_D), which undergoes a chemical reaction or conformational change. The correlation function of the system can be represented by Equation 2 (17),

\[
G(\tau) = 1 + \frac{1}{N(1-F)} \left(1 + \frac{\tau}{\tau_D}\right)^{-0.5}
\]  
(Eq. 2)

The parameter τ_F represents the time constant of a conformational fluctuation between a fluorescent and a non-fluorescent conformer (for example, A ⇔ B, and both of the conformers, A and B, have the same τ_D) provided that the said event occurs faster than the diffusion time (τ_D ≪ τ_F). F is the amplitude of τ_F, which corresponds to the average fraction of the molecules in the non-fluorescent (dark) state (17). For both Equations 1 and 2, a rapid exponential component with amplitude 20–30% and the time constant of 1–3 μs for the triplet state photophysics also is present.

The diffusion coefficient (D) of the molecule can be calculated from τ_D using Equation 3,

\[
\tau_D = \frac{\omega^2}{4D}
\]  
(Eq. 3)

where ω is the beam radius of the observation volume, which can be obtained by measuring τ_D of a fluorophore of known D.

The hydrodynamic radius (r_h) of a labeled molecule can be calculated from D using the Stokes Einstein equation (Equation 4),

\[
D = \frac{kT}{6\pi \eta r_h}
\]  
(Eq. 4)

where k is the Boltzmann constant, T is the temperature, and η corresponds to the viscosity of the solution.
FCS data were further analyzed by maximum entropy method (MEM) (18). This is a model-free method in which the multicomponent correlation function can be represented by $n$, the number of non-interacting fluorescent species. These species can have diffusion time values between 0.001 and 500 ms. MEM minimizes the parameter $\chi^2$ and maximizes the entropic quantity, $S = \sum p_i \ln p_i$ (where $p_i = a_i \sum a_j$) to obtain an optimized fit (15, 18).

**RESULTS**

FCS has been used extensively to measure hydrodynamic radius ($r_H$) of fluorescently labeled molecules at a single molecule resolution (19, 20). Additionally, FCS has been used to monitor conformational dynamics of biological molecules at $\mu$s time scale (3, 4). As a matter of fact, FCS is one of the very few experimental techniques that can measure protein dynamics in the $\mu$s time scale. This is because NMR and stopped flow methods are limited by their time resolution typically at ms; time resolved fluorescence decay experiments are effective only at the ns or ps time range. Computational studies including molecular dynamics simulations are frequent at the ns or ps time scale because of computational constraints one would encounter to perform simulations at $\mu$s time scale.

In a typical FCS experiment, intensity fluctuations of a labeled molecule inside a small observation volume are monitored with time. Intensity fluctuations may arise due to the diffusion of the molecule (with a diffusion time of $\tau_D$) in or out of the observation volume (Equation 1) or due to the kinetics of a conformational change (with the time constant of $\tau_R$) whose rate is faster than the molecular diffusion ($\tau_R \ll \tau_D$, Equation 2). Analysis of the diffusion component ($\tau_D$) using Equations 3 and 4 provides hydrodynamic radius ($r_{H1}$). Analysis of the kinetic component provides information about the time constant ($\tau_R$) of the conformational fluctuations in the $\mu$s time scale.

We have used FCS and far UV CD to study salt-induced collapse and secondary structure formation of TMR-labeled cytochrome c (cytochrome c-TMR) at pH 2. Labeling with TMR does not affect the conformation and folding of cytochrome c (14, 21). At low pH (for example, at pH 2), the protein becomes positively charged. To avoid the repulsive interactions between the positively charged residues at acidic pH, it unfolds and becomes extended. Unfolding of cytochrome c-TMR at pH 2 leads to a large decrease in the secondary structure as observed by far UV CD experiments (supplemental Fig. S1). The correlation functions obtained by the FCS experiments with cytochrome c-TMR at pH 2 could not be fit to Equation 1 resulting in non-random residual distributions (Fig. 1a). The data instead have been fit to Equation 2 (Fig. 1b), which contains an exponential time constant ($\tau_R$ of 50 $\mu$s) in addition to a diffusion component (of diffusion time of $\tau_D$). Using the $\tau_D$ value obtained from the diffusion component, we have calculated the hydrodynamic radius ($r_{H1}$) of the protein at pH 2 to be 31 Å (using Equations 3 and 4, “Experimental Procedures”). The parameter $\tau_R$ denotes the time constant of a conformational fluctuation occurring between a non-fluorescent and a fluorescent conformer (also see “Experimental Procedures” for explanations of Equations 1 and 2). The parameter $F$ denotes the amplitude of $\tau_R$ describing the relative population of the compact non-fluorescent state in equilibrium. Fitting of the FCS data using two exponential components (1 diffusion + 2 exponential components) does not improve the quality of the fit.
Fluorescence Correlation Spectroscopy on Cytochrome c Folding

The presence of an equilibrium containing three states (for example, one described in Equation 6) can be ruled out. This is because, the use of Equation 2 (which represents a two-state equilibrium between \( U \) and \( I_C \), Equation 5 above) is sufficient to successfully fit the correlation functions data of cytochrome c-TMR at pH 2. The use of an equation containing two exponential components (with time constants of \( \tau_{E1} \) and \( \tau_{E2} \)) in addition to a diffusion component (of diffusion time of \( \tau_{D} \)), which would define a three states equilibrium (Equation 6), does not improve the quality of the fit.

\[
U \text{ (extended)} \rightleftharpoons I_C \text{ (collapsed)}
\]  
\[
I_C' \text{ (collapsed)} \rightleftharpoons U \text{ (extended)} \rightleftharpoons I_C'' \text{ (collapsed)}
\]  

It has been shown that the presence of heme quenches TMR fluorescence of the protein (22). The compact collapsed conformer (\( I_C \)) is non-fluorescent because heme (the quencher) is close to the attached fluorophore (TMR) ensuring efficient quenching of TMR fluorescence. Alternatively, the extended conformer (\( U \)) is strongly fluorescent because heme (the quencher) and TMR (the fluorophore) are distant from each other. The parameter \( F \), which corresponds to the population of the non-fluorescent conformer (see “Experimental Procedures”), would represent the population of \( I_C \) in the equilibrium between \( U \) and \( I_C \). The time constant \( \tau_{E} \) of the interconversion between \( U \) and \( I_C \) has been found to be 50 µs. The population of \( I_C \) increases with sodium perchlorate concentration as observed by a decrease in \( r_{14} \) (Fig. 2a) and an increase in \( F \) (Fig. 2b). The decrease in \( r_{14} \) with sodium perchlorate occurs in a single co-operative step that could be fit successfully using a two-state transition model, and the midpoint is 23 mM. The increase in \( F \) with the sodium perchlorate concentration occurs simultaneously (Fig. 2b), and a fit using the two state transition hypotheses leads to a midpoint of 20 mM.

The presence of an equilibrium between an extended and a compact conformer in the unfolded state of cytochrome c has been shown before by fluorescence spectroscopy and small angle x-ray scattering experiments (14, 23, 24). The formation of the collapsed unfolded states for multiple other proteins has also been demonstrated (25–29). Shastry and Roder (6)
observed a 60-µs time kinetics of the formation of the compact state of cytochrome c using a continuous flow capillary mixing technique. Pascher et al. (29) observed a 40-µs time kinetics for the early events of cytochrome c folding. These results are comparable with the 50-µs time constant observed in the present study.

Several control experiments have been carried out to establish that the parameter \( \tau_R \) measures a true physical event and is not a fitting artifact. First, experiments have been carried out with TMR (the free dye) using the same setup and experimental conditions. Supplemental Fig. S2 shows the correlation functions obtained by the FCS experiments with TMR at pH 2. The data are fit successfully to Equation 1 and the use of \( \tau_R \) is not needed. Second, FCS experiments have been carried out with cytochrome c-TMR using different laser power and pinhole diameters. The possibility of an artifactual extra component at large pinhole diameter has been reported previously (30). We see no systematic variation in the value of \( \tau_R \) and \( F \) with the variation of laser power and pinhole diameters studied at pH 2 in the absence or presence of sodium perchlorate (supplemental Figs. S3 and S4). This experiment suggests that \( \tau_R \) represents a true conformational event and not an erroneous component arising from the imperfect Gaussian approximation of the confocal volume element.

In the third control experiment, we used cyanogen bromide-treated labeled cytochrome c-TMR as the control. Cyanogen bromide treatment with the protein yields a fragment with molecular weight of 3139 Da (as judged by mass spectrometry, supplemental Fig. S5). Although the fragment was still labeled with TMR, it did not contain the heme group as judged by absorption spectroscopy (supplemental Fig. S6). The absence of heme in the cyanogen bromide fragment resulted in a complete absence of fluorescence quenching observed in the full-length protein (supplemental Fig. S7 shows the fluorescence spectra of native and cyanogen bromide-treated fragment of cytochrome c with the same TMR concentration). Because the fragment lacks the non-fluorescent state and the parameter \( F \) in Equation 2 denotes the relative amplitude of the non-fluorescent state, the correlation function of the cyanogen bromide-treated fragment is not expected to contain the \( \tau_R \) component. Alternatively, the presence of any \( \tau_R \) component in the correlation function of the cyanogen bromide fragment would represent an artifact and not a physical motion. The correlation function data of the cyanogen bromide-treated protein can be fit successfully to Equation 1 and the use of \( \tau_R \) is not needed (supplemental Fig. S8).

The addition of sodium perchlorate at pH 2 results in the formation of a molten globule-like intermediate (defined in the present work as \( I_C \)). The formation of \( I_C \) is accompanied by the formation of large secondary structure (supplemental Fig. S1). The salt-induced formation of secondary structure has been monitored independently by the increase in the ellipticity at 222 nm using far UV CD. Fig. 3 shows a comparison between the formation of \( I_C \) (monitored by the decrease in \( r_{IC} \), and an increase in \( F \)) and that of \( I_S \) (monitored by the increase in the ellipticity at 222 nm). Although the formation of \( I_C \) is a two-state cooperative transition, the formation of \( I_S \) occurs through a shallow transition and completes at higher salt concentration.

Moreover, the formation of \( I_S \) is significantly less cooperative. A comparison between these two processes (the formation of \( I_C \) and \( I_S \)) suggests that \( I_C \) contains \( \sim 60\% \) of the secondary structure, whereas the remaining secondary structure forms gradually as the \( I_S \) state forms (Fig. 3). Although the kinetics of the formation of \( I_S \) is not measured directly, we could unambiguously show that the formation of \( I_C \) and \( I_S \) are not simultaneous and they correspond to two separate thermodynamic events. As mentioned before, \( I_S \) is similar to the molten globule-like intermediate states studied extensively for multiple proteins including cytochrome c (31). The formation of the native state \( (N) \) occurs in subsequent slow steps requiring further refolding (such as using a buffer with a pH of 7.4). It may be important to point out that extreme care has been exercised to rule out any aggregation related complications in these measurements with the protein at low pH and in the presence of salt. This is particularly important for the bulk CD experiments, which require relatively high concentrations. For every experiment, the reversibility of the CD measurements has been established.

Although urea has been widely used as a protein denaturant, the mechanism of its action has been debated extensively (32–36). Although an “indirect” mechanism has been proposed, it is now generally believed that urea unfolding occurs through “direct” binding to protein backbone. Interestingly enough, there are controversies regarding the nature of the direct binding of urea: either it is hydrophobic and directed toward the apolar residues (36), or it is performed through hydrogen bonding and directed toward the polar residues of the protein (37). It has been shown using hydrogen deuterium exchange that urea unfolds protein by forming direct hydrogen bonds to the peptide groups (38).

In a separate experiment, we use FCS and far UV CD to monitor the formation of \( I_C \) and \( I_S \) induced by sodium perchlorate in the presence of increasing concentration of urea at pH 2. The interaction of urea with the protein would oppose the formation of compact collapsed state, and thus, these experiments should monitor the competing effects of urea and sodium per-
chlorate on the formation of IC and IS. Fig. 4 shows the variation of \( r_{H} \) and the ellipticity at 222 nm with sodium perchlorate concentration in the absence and presence of different concentrations of urea at pH 2. At low urea concentration (0 and 1 M urea), the formation of IC and IS is not simultaneous. The formation of IC occurs through a sharp two-state transition with a large decrease in \( r_{H} \) and increase in \( F \), whereas IS forms gradually through a shallow transition (see above).

In the presence of 3 M urea and beyond at pH 2, \( F \) has been found to be very low (\(< 1\%\)) or it may even be absent. In addition, no change in \( F \) (which is inherently less) is observed in the presence of added sodium perchlorate (Figure 5). The absence of any significant amplitude of \( F \) in the absence or presence of sodium perchlorate concentrations suggests that there is no rapid equilibrium between the extended (U) and any compact (IC-like) state. Thus, the interconversion between U and IC must be slow with time constant higher than the diffusion time (\(< 9270 \) D). This conclusion comes directly from Equation 2 (see “Experimental Procedures”), which assumes \( R_{d} < 9270 \) for \( R_{d} \) and \( F \) to be observed in the experimental correlation functions. This inference is supported also by Fig. 4, c and d, which show non-two state formation of IC, which occurs simultaneously with the formation of IS. The values of ellipticity at 222 nm and \( r_{H} \) in the absence and presence of high concentrations of sodium perchlorate for all urea concentrations are shown in Table 1. These values were used as the start and end points of all the experiments described above. Supplemental Fig. S9 shows the far UV CD spectra of the protein in the absence and presence of sodium perchlorate in the presence of different concentrations of urea.

Because the above experiments have been carried out with only one protein (cytochrome c-TMR), we wanted to repeat some of these measurements with at least another protein system to find out whether any general trend could be established. FCS experiments with Alexa Fluor 488-maleimide-labeled bovine serum albumin (BSA-Alexa Fluor 488-maleimide) suggest a sharp collapse formation at pH 2 in the presence of

The data deviate from this model in the presence of urea and the drawn lines are only for presentation purposes (b–d).
increasing concentrations of sodium perchlorate (supplementary Fig. S10). Identical to what has been observed with cytochrome c-TMR, collapse formation follow a shallow transition in the presence of urea at pH 2 (Fig. 4). Although the results obtained with BSA-Alexa Fluor 488-maleimide are identical to that of cytochrome c-TMR, we are currently investigating several other protein systems containing different secondary structural elements (it should be noted that both cytochrome c-TMR and BSA-Alexa Fluor 488-maleimide are predominately α-helical) to determine whether the nature of secondary structure may have any influence on the interconnection between the hydrophobic collapse and secondary structure formation.

DISCUSSION

The nature of the collapse or whether it is a homopolymer (nonspecific) or heteropolymer (specific) has been the subjects of long-standing debates. A collapse process pertaining to a homopolymer is expected to be continuous and extended (39). Computational and experimental studies suggest that homopolymer collapse comprise multiple kinetic stages leading to extended process (39–41). The observed formation of \( L_c \) in the absence of urea (Fig. 4a) could be fit successfully to a two-state transition. In contrast, the presence of high concentration of urea results in relatively extended transitions for \( I_c \), which deviates from the two-state behavior (Fig. 4d). To further clarify, we have shown the MEM distribution of \( r_{D_2} \) values in the presence of different sodium perchlorate concentration at pH 2 in the absence (Fig. 6a) and presence of 3 M urea (Fig. 6b). In the absence of urea, the MEM profiles remain unchanged at a relatively large value of \( r_{D_2} \) before the collapse and shift sharply to the lower value as the transition occurs (Fig. 6a). In the presence of 3 M urea, this is not the case, and the shift in the MEM profiles occurs in continuous manner (Fig. 6b). The FCS data presented in this paper thus suggest that the collapse in the absence of urea may not follow nonspecific homopolymer formalism. In the presence of urea, on the other hand, the data seem to suggest a homopolymer-like nonspecific compaction. The presence of “specific” and “nonspecific” components in the early folding of Barstar has been identified using multisite fluorescence resonance energy transfer measurements (42).

Using the considerations outlined in the previous paragraph, the specific and nonspecific contributions to the secondary structure formation have been calculated in the presence of different urea concentrations and shown in Fig. 7. Fig. 7 shows that the nonspecific component to the secondary structure of \( I_s \) increases with the urea concentration in a sigmoidal fashion.

![FIGURE 6](image)

**FIGURE 6.** The variation of the maximum entropy profiles of cytochrome c-TMR with sodium perchlorate concentration in the absence (a) and presence (b) of 3 M urea at pH 2. In the absence of urea, the maximum entropy profiles shift sharply toward lower diffusion time as \( I_c \) forms. This behavior of maximum entropy profiles is not consistent with the homopolymer collapse hypothesis. In the presence of urea, however, \( I_c \) forms slowly and gradually over an extended range of sodium perchlorate concentration, as expected for a homopolymer collapse.

![FIGURE 7](image)

**FIGURE 7.** The variation in the percentage of nonspecific (red) and specific contributions (black) to the secondary structure with urea concentration. A percentage of nonspecific component increases and that of specific component decreases with urea concentration. It is important to note that these percentage calculations are based on total secondary structure change between the \( U \) and \( I_s \) and the native state (N) is not considered.

For example, the nonspecific component of the secondary structure is ∼40% in the absence of urea, whereas it increases to 100% in the presence of 4 M urea. The specific component, on the other hand, decreases identically from 60% (in the absence of urea) to 0% (in the presence of 4 M urea). In the presence of

---

**TABLE 1**

| Urea Concentration (M) | Ellipticity at 222 nm | Ellipticity at 222 nm | Ellipticity at 222 nm | Ellipticity at 222 nm | Ellipticity at 222 nm |
|------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                         | \( r_{H} \)         | \( r_{H} \)         | \( r_{H} \)         | \( r_{H} \)         | \( r_{H} \)         |
| 0                      | −4.5                 | −30.7                | −8.2                 | −10.6                | −15.1                |
| 1                      | −5.1                 | 31.9                 | −8                  | 22.5                 | 22.5                 |
| 3                      | −2.7                 | 31.1                 | −5                  | 25.5                 | 25.5                 |
| 4                      | −2.4                 | 32.3                 | −3.6                | 25                  | 25                  |

**TABLE 1**

The values of hydrodynamic radius (\( r_{H} \)) and ellipticity at 222 nm of \( U, I_c, I_s, \) and \( N \) in the presence of different urea concentration.

\( N \) corresponds to the native state of TMR-cytochrome c in 20 mM sodium phosphate buffer at pH 7.4. Ellipticity and \( r_{H} \) of \( U \) change in the presence of different concentrations of urea, and the change depends on the extent of the residual structure present in \( U \), degree.

**TABLE 1**

The values of hydrodynamic radius (\( r_{H} \)) and ellipticity at 222 nm of \( U, I_c, I_s, \) and \( N \) in the presence of different urea concentration.

\( N \) corresponds to the native state of TMR-cytochrome c in 20 mM sodium phosphate buffer at pH 7.4. Ellipticity and \( r_{H} \) of \( U \) change in the presence of different concentrations of urea, and the change depends on the extent of the residual structure present in \( U \), degree.
The presence of urea (Fig. 8) may now be addressed. The present data explains the apparent versus of the initial protein folding events in the refolding buffer. Thus, the long-standing specific versus nonspecific controversy of the initial protein folding events in the refolding buffer may now be addressed. The present data explains the apparent disagreement between distance measurements (light scattering or fluorescence data indicating specific nature) and secondary structure measurements (CD experiments suggesting the presence of nonspecific nature) for the early folding processes. We show that, if the collapse is probed by the contraction of the unfolded state, the initial collapse formation would seem to be specific. This conclusion comes directly from Fig. 4a because the change in IC with sodium perchlorate concentration in the absence of urea is cooperative and rapid, as expected for a specific transition. On the other hand, if the early formation of the secondary structure is monitored, it would show both nonspecific and specific contributions. The present data are also in agreement with the far UV circular dichroism measurements using microsecond continuous flow mixer, which show early formation of ~20% of the native secondary structure at the earliest time point monitored (5).

It has been shown that pronounced biases toward the native or non-native structures exist in the unfolded states of a protein (43, 44). Because U and IC maintain an equilibrium in the unfolded state, the nature of contact formation in the unfolded state would dictate the nature of bias present in IC. The presence of non-native or misfolded bias in IC, for example, would be kinetically inefficient for the subsequent folding events because these misfolded contacts needed to be broken before the protein could be folded properly. These processes of breaking of misfolding contacts would slow down the formation of the secondary structure (formation of IS). Although it is difficult to monitor the nature of bias in the unfolded state of a protein experimentally, several successful attempts are available. The presence of high content of β-structures has been observed in the collapsed unfolded state of CspTm (45). The possibility of the formation of short and highly dynamic β-segments resulting in non-native hydrogen bond interaction has been postulated within the unfolded state, which could contribute to the chain contraction (46, 47). The use of NMR has been particu-
Fluorescence Correlation Spectroscopy on Cytochrome c Folding

FIGURE 9. The calculated variation of the hydrodynamic radii of different proteins with the number of residues for the extended unfolded states in the good solvent (black) and the native states in the poor solvent (red). The $r_H$ values determined from the present FCS measurements at different solution conditions are also shown. These are unfolded state (black open circles), native state (black filled triangles), $I_p$ in the absence (red open squares) and presence of 1 M urea (red open circles), 3 M urea (black filled circles) and 4 M urea (red filled squares).

The observed value of $r_H$ in the presence of 4 M urea at pH 2 (the completely unfolded protein in good solvent) obtained by the present FCS experiments falls exactly on the predicted line (Fig. 9). Similarly, the value of $r_H$ in the presence of sodium perchlorate and no urea at pH 2 (the collapsed protein in poor solvent) also is in exact agreement with the prediction. Not surprisingly, the $r_H$ values of $I_p$ in the presence of different urea concentrations deviates significantly from both the predicted lines: they are more compact than the completely unfolded state in urea, whereas they are more extended than the collapsed/native state in aqueous buffer. Depending on the concentration of urea, the $r_H$ values of $I_p$ would be either near the predicted line for the extended state in good solvent ($n = 0.6$), or near the predicted line for the collapsed/native state in poor solvent ($n = 0.3$) or distant from both. These partially folded/unfolded states of different compactness or secondary structure are not defined well, although they may well have significant implications in the mechanism of protein folding in general and in the field of the intrinsically disordered proteins in particular. The present data support Muller-Spath et al. (53) who show that the dimensions of an intrinsically disordered protein (or the unfolded state of a globular protein) would depend on several factors, including hydrophobicity, charge interactions, solution conditions, screening effects of charged denaturants such as guanidinium hydrochloride, and the binding of denaturants with the protein chain.

Acknowledgments—We thank Professor S. Maiti of the Tata Institute of Fundamental Research for the Maximum Entropy Method analysis software. We also thank Professors Elliot Elson and Carl Frieden of the Washington University School of Medicine for critical comments on this work. Professor S. Roy of the Indian Institute of Chemical Biology is acknowledged for help and support.

REFERENCES

1. Eaton, W. A., Thompson, P. A., Chan, C. K., Hage, S. J., and Hofrichter, J. (1996) Fast events in protein folding. Structure 4, 1133–1139.
2. Hagen, S. J., Hofrichter, J., Szabo, A., and Eaton, W. A. (1996) Diffusion-limited contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. Proc. Natl. Acad. Sci. U.S.A. 93, 11615–11617.
3. Chattopadhyay, K., Elson, E. L., and Frieden, C. (2005) The kinetics of conformational fluctuations in an unfolded protein measured by fluorescence methods. Proc. Natl. Acad. Sci. U.S.A. 102, 2385–2389.
4. Chen, H., Rhoades, E., Butler, J. S., Loh, S. N., and Webb, W. W. (2007) Dynamics of equilibrium structural fluctuations of apomyoglobin measured by fluorescence correlation spectroscopy. Proc. Natl. Acad. Sci. U.S.A. 104, 10459–10464.
5. Akiyama, S., Takahashi, S., Ishimori, K., and Morishima, I. (2000) Stepwise formation of α-helices during cytochrome c folding. Nat. Struct. Biol. 7, 514–520.
6. Shatrov, M. C., and Roder, H. (1998) Evidence for barrier-limited protein folding kinetics on the microsecond time scale. Nat. Struct. Biol. 5, 385–392.
7. Chattopadhyay, K., Zhong, S., Yeh, S. R., Rousseau, D. L., and Frieden, C. (2002) The intestinal fatty acid binding protein: The role of turns in fast and slow folding processes. Biochemistry 41, 4040–4047.
8. Takahashi, S., Yeh, S. R., Das, T. K., Chan, C. K., Gottfried, D. S., and Rousseau, D. L. (1997) Folding of cytochrome c initiated by submillisecond mixing. Nat. Struct. Biol. 4, 44–50.
9. Goldbeck, R. A., Chen, E., and Kliger, D. S. (2009) Early events, kinetic intermediates, and the mechanism of protein folding in cytochrome c. Int. J. Mol. Sci. 10, 1476–1499.
10. Chen, E., Goldbeck, R. A., and Kliger, D. S. (2004) The earliest events in protein folding: A structural requirement for ultrafast folding in cytochrome c. J. Am. Chem. Soc. 126, 11175–11181.
11. Uversky, V. N., and Fink, A. L. (2002) The chicken-egg scenario of protein folding revisited. FEBS Lett. 515, 79–83.
12. Sosnick, T. R., Mayne, L., and Englander, S. W. (1996) Molecular collapse: The rate-limiting step in two-state cytochrome c folding. Proteins 24, 413–426.
13. Sosnick, T. R., Shilberman, M. D., Mayne, L., and Englander, S. W. (1997) Ultrafast signals in protein folding and the polypeptide contracted state. Proc. Natl. Acad. Sci. U.S.A. 94, 8545–8550.
14. Haldar, S., Mitra, S., and Chattopadhyay, K. (2010) Role of protein stabilizers on the conformation of the unfolded state of cytochrome c and its early folding kinetics: Investigation at single molecular resolution. J. Biol. Chem. 285, 25314–25323.
15. Ghosh, R., Sharma, S., and Chattopadhyay, K. (2009) Effect of arginine on protein aggregation studied by fluorescence correlation spectroscopy and other biophysical methods. Biochemistry 48, 1135–1143.
16. Haldar, S., and Chattopadhyay, K. (2011) Effects of arginine and other solution additives on the self-association of different surfactants: An investigation at single-molecule resolution. Langmuir 27, 5842–5849.
17. Haupts, U., Maiti, S., Schwille, P., and Webb, W. W. (1998) Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. Proc. Natl. Acad. Sci. U.S.A. 95,
Fluorescence Correlation Spectroscopy on Cytochrome c Folding

15. O'Brien, E. P., Dima, R. I., Brooks, B., and Thirumalai, D. (2007) Interactions between hydrophobic and ionic solutes in aqueous guanidinium chloride and urea solutions: Lessons for protein denaturation mechanism. J. Am. Chem. Soc. 129, 7346–7353

16. Lim, W. K., Rösgen, J., and Englander, S. W. (2009) Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. Proc. Natl. Acad. Sci. U.S.A. 106, 2595–2600

17. Hagen, S. I., and Eaton, W. A. (2000) Two-state expansion and collapse of a polypeptide. J. Mol. Biol. 301, 1019–1027

18. Chu, B., Ying, Q., and Grosberg, Y. A. (1995) Two-stage kinetics of single-chain collapse polystyrene in cyclohexane. Macromolecules 28, 180–189

19. Zhu, P. W., and Napper, D. H. (1997) The longer time collapse kinetics of interfacial poly (N-isopropylacrylamide) in water. J. Chem. Phys. 106, 6492–6498

20. Sinha, K. K., and Udgaonkar, J. B. (2007) Dissecting the non-specific and specific components of the initial folding reaction of barstar by multi-site FRET measurements. J. Mol. Biol. 370, 385–405

21. Srinivasan, R., and Rose, G. D. (1999) A physical basis for protein secondary structure. Proc. Natl. Acad. Sci. U.S.A. 96, 14258–14263

22. Baldwin, R. L., and Rose, G. D. (1999) Is protein folding hierarchical? II. Folding intermediates and transition states. Trends Biochem. Sci. 24, 77–83

23. Hofmann, A., Kane, A., Nettels, D., Hertzog, D. E., Baumgaertel, P., and Schuler, B. (2007) Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. Proc. Natl. Acad. Sci. U.S.A. 104, 105–110

24. Möglich, A., Joder, K., and Kiefhaber, T. (2006) End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. Proc. Natl. Acad. Sci. U.S.A. 103, 12394–12399

25. Nettels, D., Müller-Spaeth, S., Soranno, A., Hirschfeld, V., Hofmann, H., Kiefhaber, T., Reymond, L., Nettels, D., and Schuler, B. (2010) From the cover: Charge density measurements by pulse-field gradient NMR techniques. J. Am. Chem. Soc. 132, 9401–9407

26. Nettels, D., Müller-Spaeth, S., Soranno, A., Hirschfeld, V., Hofmann, H., Kiefhaber, T., Reymond, L., Nettels, D., and Schuler, B. (2010) Mapping protein collapse with single-molecule fluorescence and kinetics synchrotron radiation circular dichroism spectroscopy. Proc. Natl. Acad. Sci. U.S.A. 107, 20740–20745

27. Kristjansdottir, S., Lindorff-Larsen, K., Fieber, W., Dobson, C. M., Vendruscolo, M., and Poulos, F. M. (2005) Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. J. Mol. Biol. 347, 1053–1062

28. Mukrasch, M. D., Markwick, P., Biermt, J., Bergen, M., Bernadò, P., Griesinger, C., Mandelkow, E., Zweckstetter, M., and Blackledge, M. (2007) Highly populated turn conformations in natively unfolded Tau protein identified from residual dipolar couplings and molecular simulations. J. Am. Chem. Soc. 129, 5235–5243

29. Meng, W., and Raleigh, D. P. (2011) Analysis of electrostatic interactions in the denatured state ensemble of the N-terminal domain of L9 under native conditions. Proteins 79, 3500–3510

30. Zhang, O., and Forman-Kay, J. D. (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. Biochemistry 36, 3959–3970

31. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. Biochemistry 38, 16424–16431

32. Müller-Späth, S., Soranno, A., Hirschfeld, V., Hofmann, H., Rückger, S., Reymond, L., Nettels, D., and Schuler, B. (2010) From the cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. Proc. Natl. Acad. Sci. U.S.A. 107, 14609–14614