5086 *Biochemistry* **2010,** *49,* 5086–5096
DOI: 10.1021/bi100222h

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Osmolyte-Induced Folding of an Intrinsically Disordered Protein: Folding Mechanism in the Absence of Ligand†

Yu-Chu Chang and Terrence G. Oas*

Department of Biochemistry, Box 3711, Duke University Medical Center, Durham, North Carolina 27710

Received February 12, 2010; Revised Manuscript Received May 15, 2010

ABSTRACT: Understanding the interconversion between thermodynamically distinguishable states present in a protein folding pathway provides not only the kinetics and energetics of protein folding but also insights into the functional roles of these states in biological systems. The protein component of the bacterial RNase P holoenzyme from *Bacillus subtilis* (P protein) was previously shown to be unfolded in the absence of its cognate RNA or other anionic ligands. P protein was used in this study as a model system to explore general features of intrinsically disordered protein (IDP) folding mechanisms. The use of trimethylamine *N*-oxide (TMAO), an osmolyte that stabilizes the unliganded folded form of the protein, enabled us to study the folding process of P protein in the absence of ligand. Transient stopped-flow kinetic traces at various final TMAO concentrations exhibited multiphasic kinetics. Equilibrium "cotitration" experiments were performed using both TMAO and urea during the titration to produce a urea−TMAO titration surface of P protein. Both kinetic and equilibrium studies show evidence of a previously undetected intermediate state in the P protein folding process. The intermediate state is significantly populated, and the folding rate constants are relatively slow compared to those of intrinsically folded proteins similar in size and topology. The experiments and analysis described serve as a useful example for mechanistic folding studies of other IDPs.

Protein folding kinetics have been extensively explored in many model protein systems (*1*, *2*), but most previous studies focused primarily on natively structured globular proteins. Recently, a biologically important class of proteins called intrinsically disordered proteins (IDP) have become widely recognized (*3*−*6*). It had long been accepted that the folded structure of a protein determines its function, but most IDPs lack definite structure in the absence of bound ligand(s). Many IDPs adopt active conformations through disorder-to-order transitions coupled to binding of other cellular components, and this coupled folding and binding phenomenon is a major feature of the IDP family (*7*). Coupling a folding transition to a molecular interaction has been considered functionally advantageous because it may allow a protein to recognize different target substrates, could enhance the rate of intermolecular interactions, and could provide for large intermolecular interfaces with a relatively small protein (*8*). Recent studies have shown that many IDPs participate in transcriptional and translational regulation (*9*). Because of their structural flexibility, some IDPs may function as multimolecular hubs that interact with multiple binding partners (*3*).

Ribonuclease P holoenzyme (RNase P) is the endonuclease that catalyzes the maturation of pre-tRNA via removal of the 5′ leader sequence (*10*−*12*). The crystal structure of the 119-residue protein subunit of *Bacillus subtilis* RNase P (P protein hereafter) has been determined to 2.6 Å resolution (*13*, *14*). Our previous biophysical studies of P protein demonstrated that it is intrinsically disordered in the absence of ligand and at low ionic strengths (*15*). The major functional role of P protein in the holoenzyme is to enhance substrate recognition (*16*) and stimulate the intrinsic ribonuclease activity of the catalytic RNA subunit (*17*). Several structural roles for P protein have also been proposed, including stabilizing the structure of P RNA or compensating for charge repulsion (*18*). Because P protein is intrinsically disordered, holoenzyme assembly includes folding of the protein. Thus, the mechanism of P protein folding is important for understanding the assembly of the RNase P ribonucleoprotein complex. Moreover, it is interesting to determine the sequence of events in multistep folding and RNA binding reactions.

Previous studies have demonstrated that many small proteins exhibit simple two-state folding kinetics (*1*). In this situation, the two thermodynamically distinguishable ensembles are in equilibrium and undergo cooperative conformational change as they interconvert. However, many proteins do not exhibit two-state folding but instead fold via one or more partially folded intermediates. These intermediates can be broadly divided into two classes (*19*). One class consists of obligatory folding intermediates, which usually, but not always, have a partially formed folded structure. Under some circumstances, they can also accelerate the search for the folded state (*20*, *21*). The second class consists of nonobligatory intermediates, which are usually kinetically trapped misfolded species (*22*). However, these intermediate states can also be steps in an alternative folding pathway (*23*). Although intermediate states can be crucial in the overall folding process, it is sometimes challenging to detect and characterize them. Many intermediates form rapidly, within the dead time of a conventional stopped-flow instrument. Transient kinetic intermediates are generally not sufficiently stable to be observed at equilibrium under strongly denaturing conditions. Intermediates may also be spectroscopically silent,

†Supported by National Institutes of Health Grants 5RO1GM061367 and 5RO1GM081666 (to T.G.O.).

*To whom correspondence should be addressed: Department of Biochemistry, Box 3711, Duke University Medical Center, Durham, NC 27710. Telephone: (919) 684-4363. Fax: (919) 684-8885. E-mail: oas@duke.edu.

 Published on Web 05/18/2010

making them difficult to detect. Thus, it is important to search for conditions that maximize the population and stability of possible intermediate states in order to determine their thermodynamic, kinetic, and even structural properties. In this study, we detected a state that apparently is an obligatory, on-pathway intermediate in the folding of P protein.

There have been many studies of biological functions of IDPs (*8*), but fewer fundamental biochemical and biophysical studies of properties such as their folding mechanisms. Renaturation of IDPs and denaturation of natively folded proteins are two sides of the same experimental coin, and both provide insights into the balance of forces and intrinsic steps of the folding−unfolding reaction. The theory and methods behind these two kinds of experiments are fundamentally the same. The main difference is the experimental design: whereas natively structured proteins require a denaturant such as urea or guanidinium chloride to perturb the equilibrium, IDPs require a "protection osmolyte" (*24*). In this study, an osmolyte, trimethylamine *N*-oxide (TMAO), accompanied by the denaturant urea was used to perturb folding equilibrium to nearly fully deplete both the unfolded (high TMAO) and folded (high urea) states. Interestingly, as described below, we find no conditions that fully deplete the intermediate state. The highly basic amino acid composition (11 Arg, 19 Lys, and 3 His out of 119 residues) of P protein gives a net charge of +16 to +17 at neutral pH and a pI of 10.8. Electrostatic repulsion in the absence of the cognate polyanionic ligands may explain why folding to a compact folded state is unfavorable.

Osmolytes are small organic molecules, including certain amino acids, polyols, and methylamines, that are used in many organisms to maintain the intracellular water level (*25*). This water stress can result from osmotic pressure, desiccation, extreme temperatures, or high intracellular urea concentrations. Water stress resistant cells have evolved to use osmolytes to maintain macromolecular stability and function in the presence of these stresses. Many studies have used osmolytes to stabilize the folded state of proteins (*26*, *27*) and RNA (*28*) secondary and tertiary structures. The major mechanism of protein stabilization by an osmolyte is unfavorable interaction with the backbone atoms, which increases the free energy of the denatured state (*29*). Most osmolytes are interchangeable and additive, which means that stabilization can be achieved in combination. Studies also show that the free energy of unfolding is linearly dependent on osmolyte concentration, which is analogous to the linear dependence of the unfolding free energy on chemical denaturant, albeit with a slope with the opposite sign (*30*). This property is particularly useful for estimation of the unfolding free energy of IDPs and partly folding proteins (*31*). IDPs can be titrated with an osmolyte to generate a refolding curve from which the (positive) folding free energy can be estimated using the same linear extrapolation method used to analyze a urea or guanidine denaturation curve. The stability of a partly folded protein can be measured by mixing both urea and osmolyte solvent in a protein sample to generate complete denatured or native state baselines in the titration curve (*31*).

In this study, P protein was used as a model system to elucidate the folding−unfolding process of intrinsically disordered proteins. Both kinetic and equilibrium data show the presence of an intermediate state in the P protein folding pathway. A sequential three-state mechanism was modeled to fit the experimental data. Kinetic and equilibrium data were fitted globally to extract the rate constants, thermodynamic parameters, and spectroscopic properties of these three states. To the best of our knowledge, these are the first folding and unfolding rate constants determined by TMAO-induced folding kinetics for an intrinsically disordered protein.

EXPERIMENTAL PROCEDURES

Chemicals and Reagents. Ultrapure urea was purchased from Nacalai Tesque Inc. Trimethylamine *N*-oxide (TMAO) was from Fluka. Cacodylic acid was from Sigma. To remove the background fluorescence from TMAO, a 3 M stock TMAO solution was mixed with activated charcoal (0.025 g/mL) for at least 5 h. The charcoal was then removed by filtration using a 0.22 μm syringe filter. Urea and TMAO concentrations were measured via the refractive index using the equations from refs *32* and *26*, respectively.

Expression and Purification of P Protein. The gene for P protein variant F107W used in this study was constructed by the QuickChange site-directed mutagenesis protocol (Strategene). The variant was overexpressed in *Escherichia coli* and purified as described previously (*13*) with the following modifications. The pooled fractions of P protein that eluted from the second CM-Sepharose column were concentrated using a Centriprep filter (Amicon) until the volume of the pooled fractions was reduced to 1−2 mL. The buffer of the protein stock was changed to 6 M guanidinium chloride with 10 mM Tris (pH 7.5) during this concentrating step. The concentrated stock was then loaded on an S-100 column (GE Healthcare) equilibrated with 6 M Gnd-HCl and 10 mM Tris (pH 7.5) to remove EDTA from P protein. Fractions containing P protein were pooled and stored at −80 °C. Before equilibrium titration or stopped-flow kinetic experiments, the P protein sample was dialyzed against water extensively, and then 20 mM sodium cacodylate (pH 7). In all experiments, the protein concentration was determined using the Edelhoch method using an extinction coefficient of 11200 M^{-1} cm^{-1} at 276 nm (*33*).

Circular Dichroism (CD) and Fluorescence Spectrum. An Aviv model 202 CD spectrometer was employed to collect CD wavelength spectra. Samples were scanned from 260 to 190 nm in 1 nm increments with a bandwidth of 1.5 mm and averaged for 3 s at each wavelength. The protein concentration was 10 μM in 20 mM sodium cacodylate (pH 7) in the presence or absence of TMAO using a cuvette with a path length of 1 mm at 25 °C.

Fluorescence spectra were recorded on a Fluorolog-3 spectrometer from HORIBA Jobin Yvon using a cuvette with a path length of 1 cm. The sample was 5 μM P protein in 20 mM cacodylate (pH 7). The temperature was controlled at 25 °C by a circulating water bath. The excitation wavelength was 285 nm with a 4 nm bandpass, and emission spectra were collected from 300 to 400 nm at 1 nm increments.

CD and Fluorescence Equilibrium Cotitration Experiments. The urea−TMAO equilibrium cotitration experiments were performed on an Applied Photophysics PiStar instrument equipped with both fluorescence and CD detection. A total of six renaturation equilibrium titration curves were collected in each set of experiments, and all samples were in 20 mM sodium cacodylate buffer (pH 7.0) at 25 °C. The starting sample protein concentration was 5 μM in 0, 2, 4, or 6 M urea. The two additional starting samples were 5 μM protein in 6 M urea and 0.5 M TMAO or in 6 M urea and 0.9 M TMAO. The titrant solution was 5 μM protein in 1.4 M TMAO. An automated Hamilton titrator was used to perform all titrations. The CD signal was recorded at 222 nm, and the fluorescence signal was recorded using a 320 nm cutoff filter with an excitation wavelength of

285 nm and a bandwidth of 4 nm. The CD and fluorescence signals were collected simultaneously from the same sample. Each urea−TMAO concentration pair data point was measured for 15 s to produce a mean signal. The equilibrium time between two titration points was 3−3.5 min.

Stopped-Flow Fluorescence Kinetic Experiments. The stopped-flow fluorescence kinetic experiments were performed on an Applied Photophysics SX20 instrument at 25 °C with a circulating water bath. The excitation wavelength was 285 nm with a slit width of 1 mm. The path length of the observation cell was 2 mm. The fluorescence signal was collected through a 320 nm high-pass filter. For single-mixing folding and unfolding experiments, a 6- or 11-fold dilution was used to initiate the reactions. In the refolding experiments, the protein concentration was 6 μM in 20 mM sodium cacodylate buffer (pH 7.0), and TMAO concentrations ranged from 0.15 to 1.4 M in 20 mM sodium cacodylate buffer (pH 7.0). The final protein concentration was 1 μM, with final TMAO concentrations ranging from 0.125 to 1.167 M. For unfolding experiments, the protein concentration was 11 μM in 1 M TMAO with 20 mM sodium cacodylate, and final TMAO concentrations ranged from 0 to 0.5 M. Following 11-fold dilution in the stopped-flow instrument, the final protein concentration was 1 μM in 0.091−0.55 M TMAO. The time traces of the fluorescence change for the reaction were collected for 20 s in a split time base mode. Five thousand points were collected for the first 10 s, and 1000 points were collected from 10 to 20 s.

Fluorescence signals of the protein from equilibrium and kinetic experiments are normalized using signals between two experimental conditions. The signal is arbitrarily set to 1 when the protein is in 0 M TMAO and 0 M urea, and the signal is set to 0 when the protein is in 1.2 M TMAO and 0 M urea.

Data Analysis. (*a*) *Kinetic Folding and Unfolding Data*. TMAO-induced folding and unfolding kinetic time traces from stopped-flow experiments were fitted by the nonlinear least-squares method with the multiple-exponential function

$$F(t) = F(\infty) + \sum_i A_i e^{k_{obs} t} \quad (1)$$

where $F(t)$ and $F(\infty)$ are the observed fluorescence at time t and infinite time, respectively. A_i and k_{obs} are the fluorescence amplitudes and observed rate constants, respectively, of the different phases. The amplitudes and observed rate constants were plotted against the final TMAO concentration.

(*b*) *Equilibrium Titration Data*. Equilibrium urea−TMAO cotitration surface data were fit to a linear three-state model assuming that ΔG_{UI} and ΔG_{IN} are linearly dependent on urea and TMAO concentration (*31*), according to eqs 2−6:

$$\Delta G_{UI}(\text{TMAO}, \text{urea}) = \Delta G^0_{UI} + m^T_{UI}[\text{TMAO}] + m^U_{UI}[\text{urea}] \quad (2)$$

$$\Delta G_{IN}(\text{TMAO}, \text{urea}) = \Delta G^0_{IN} + m^T_{IN}[\text{TMAO}] + m^U_{IN}[\text{urea}] \quad (3)$$

$$K_{UI} = \exp\left(-\frac{\Delta G_{UI}}{RT}\right)$$
$$K_{IN} = \exp\left(-\frac{\Delta G_{IN}}{RT}\right) \quad (4)$$

$$F_U = 1/(1 + K_{UI} + K_{UI}K_{IN})$$
$$F_I = K_{UI}/(1 + K_{UI} + K_{UI}K_{IN})$$
$$F_N = K_{UI}K_{IN}/(1 + K_{UI} + K_{UI}K_{IN}) \quad (5)$$

$$\text{Sig}^{CD} = S^{CD}_U F_U + S^{CD}_I F_I + S^{CD}_N F_N$$
$$\text{Sig}^{FL} = S^{FL}_U F_U + S^{FL}_I F_I + S^{FL}_N F_N \quad (6)$$

where m^T_{ij} and m^U_{ij} are the m values that describe the linear osmolyte concentration dependence of the free energy, ΔG_{ij}; K_{ij} is the equilibrium constant of the i−j equilibrium; F_i is the fractional population of state i at a specific osmolyte concentration; S^{CD}_i and S^{FL}_i are the CD and fluorescence intensities, respectively, of the unfolded (U), intermediate (I), and folded (N) states; and Sig^{CD} and Sig^{FL} are the observed CD and fluorescence signals, respectively.

Scheme 1

$$\mathbf{U} \underset{k_{IU}}{\overset{k_{UI}}{\rightleftharpoons}} \mathbf{I} \underset{k_{NI}}{\overset{k_{IN}}{\rightleftharpoons}} \mathbf{N}$$

(*c*) *Global Fitting*. The CD and fluorescence equilibrium titration surfaces were fit using eqs 2−6 and 8. The amplitudes and observed rate constants were fit to a linear three-state model (Scheme 1) using the analytical solutions described by Bujalowski (*34*). The TMAO dependence of the observed rate constants and associated amplitudes assumes a linear dependence of the logarithm of the microscopic rate constants in Scheme 1 on TMAO concentration in the form

$$k_{ij} = \exp\left[\log k^0_{ij} + (\beta_{ij} - 1)\frac{m^T_{ij}}{RT}[\text{TMAO}]\right] \quad (7)$$

where k_{ij} is the microscopic rate constant for the transition from state i to j at a specific TMAO concentration. β_{ij} is the Tanford value (*35*) which is a measure of the average degree of solvent exposure of the transition state relative to the difference in solvent exposure between states i and j. Equilibrium data, kinetic amplitudes, and observed rates were then fit simultaneously using Mathematica (Wolfram) to obtain the equilibrium and kinetic parameters for the model.

RESULTS

Tryptophan Substitution in P Protein as a Probe for Studying the Kinetics and Thermodynamics of P Protein. Wild-type *B. subtilis* P protein does not contain any tryptophan residues, which are often used as intrinsic fluorescent probes to monitor protein conformational changes. To utilize this property for fluorescence-detected folding kinetic and equilibrium studies, a tryptophan-containing variant of P protein, F107W, was generated. The site for mutagenesis was selected on the basis of a sequence comparison between *B. subtilis* P protein and the highly homologous (similarity) *E. coli* P protein, C5. In *E. coli* P protein, tryptophan 107 is highly conserved among different Gram-negative bacteria. The Trp residue fits in a hydrophobic pocket and undergoes a fluorescence change upon urea unfolding (*36*). Figure 1 shows a model of the m0 rotamer of a tryptophan side chain with a −26° adjustment of χ_2 that inserted into the crystal structure of wild-type *B. subtilis* P protein (*14*). The Trp side chain fits well in this conformation using the structural biology tool in KING (*37*) which indicates good van der Waals contacts with surrounding side chains. The F107W variant shares the IDP property with wild-type P protein and can be induced to fold by binding of small anions as observed in

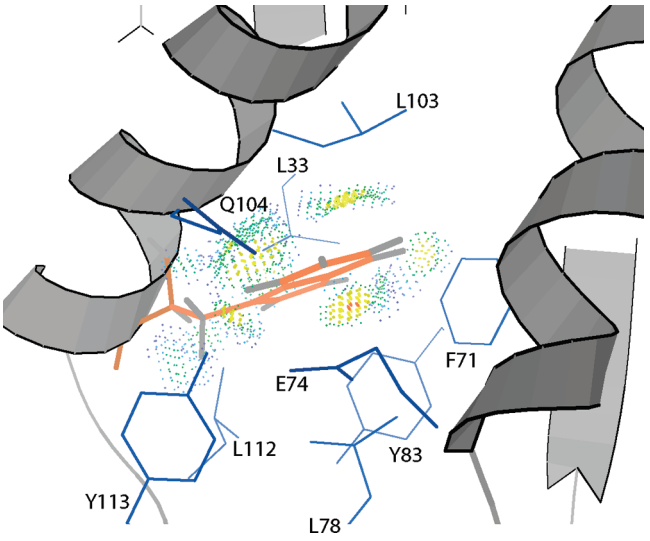

FIGURE 1: Model of Trp 107 side chain contacts with surrounding side chains. The model was constructed using the crystal structure of *B. subtilis* P protein (Protein Data Bank entry 1A6F) and the side chain mutator tool in KING (*50*). The dots generated by Probe (*51*) indicate the contact surface of the Trp side chain with the surrounding atoms. Red spikes represent steric clashes between atoms. The blue to green dots indicate favorable van der Waals contacts between each atom. The model shows that the engineered Trp fits well in this hydrophobic pocket and probably does not greatly perturb the overall structure. The modeled Trp side chain rotamer is m0 with a minor adjustment of the χ_2 angle.

previous studies of wild-type P protein (*15*). Addition of TMAO to the protein sample also shifts the folding equilibrium to the folded state. The fluorescence emission spectrum (285 nm excitation) of F107W at various TMAO concentrations from 0 to 1 M is shown in Figure 2A. Unfolded F107W has a λ_{max} at 358 nm, and the peak is blue-shifted to 308 nm when the protein folds to the folded state, indicating that the tryptophan side chain is located in a nonpolar region of the folded protein (*38*, *39*). The total fluorescence also decreases upon folding, suggesting that there is additional quenching in the native protein. The wild-type P protein crystal structure supports these conclusions, but the Trp is not in a completely nonpolar environment. As shown in Figure 1, the modeled Trp side chain is surrounded by four Leu residues, one Phe, two Tyr residues, one Gln, and one Glu. Adjacent aromatic side chains are possible quenchers of Trp fluorescence. Figure 2B shows the CD spectrum of the variant also in 0 to 1 M TMAO. The appearance of distinct peaks around 220 and 208 nm indicates the formation of α-helical secondary structure at >0.2 M TMAO. The CD spectrum of P protein in 1 M TMAO is similar to that found for 20 mM sulfate, in which P protein is predominantly folded (*15*). The NMR HSQC spectrum of F107W (Figure 2C) under folded conditions is similar to the wild-type spectrum, and more than 90% of the peaks in the F107W spectrum can be assigned via comparison with the wild-type spectrum. Most of the shifted resonances correspond to amides in the proximity of the site of the Trp substitution.

Taken together, these results suggest that the fluorescence of the Trp residue in the F107W variant is sensitive to the structural change between folded and unfolded states, with a minimum perturbation of the overall protein structure. For these reasons, the variant was used for the equilibrium and kinetic studies described below.

TMAO Stopped-Flow Kinetic Studies and Chevron Plots. Transient fluorescence traces of P protein folding−unfolding

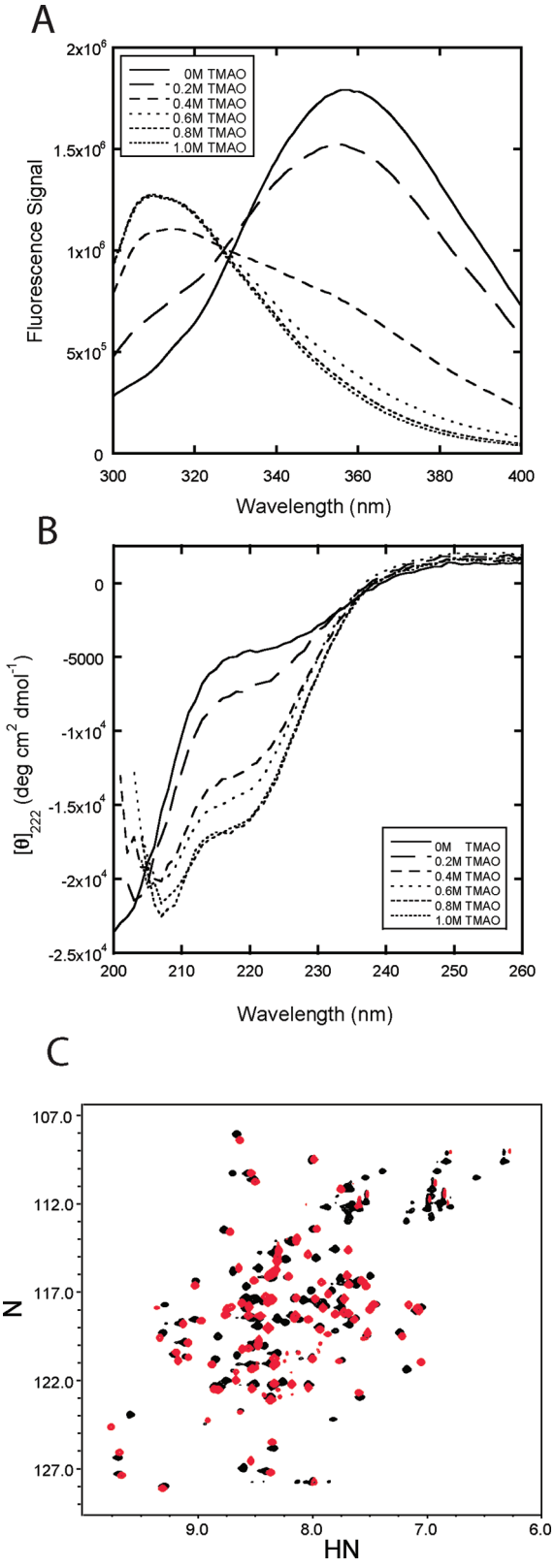

FIGURE 2: Fluorescence and circular dichroism spectra of the F107W variant of P protein at various TMAO concentrations. (A) Fluorescence spectrum of the F107W variant of P protein (5 μM, 1 cm cuvette) at the same TMAO concentrations used for CD. The excitation wavelength was 285 nm, and the emission spectra were scanned from 400 to 300 nm in 1 nm increments. (B) Far-UV CD spectrum of the F107W variant of P protein (10 μM, 1 mm cuvette) in 20 mM cacodylate buffer in 0, 0.2, 0.4, 0.6, 0.8, and 1 M TMAO. The spectra were scanned from 260 to 200 nm in 1 nm increments. (C) HSQC spectrum of F107W (red) overlaid on the wild-type spectrum (black).

reactions induced by changes in TMAO concentrations were collected using the stopped-flow method. As expected from the fluorescence emission spectra described above, the magnitude of the fluorescence signal decreases when folding is induced by mixing protein in buffer with TMAO solutions (TMAO up-jump), while the magnitude of the signal increases when unfolding is induced by mixing protein in TMAO with buffer alone (TMAO down-jump). Panels A and B of Figure 3 show representative up-jump and down-jump traces, respectively. The kinetic traces were fitted to multiple-exponential functions by a nonlinear least-squares method using Pro Data Viewer from Applied Photophysics. The minimum complexity model that fits both up-jump and down-jump kinetic traces is a sum of three exponentials (eq 1 in Experimental Procedures), so there are three observable phases in both the folding and unfolding reactions. Each fit was obtained by minimization of the residuals as shown in the insets of panels A and B of Figure 3. The observed rate constants (k_{obs}) of the fast phase (ranging from 2 to 7 s^{-1}) and middle phase (ranging from 0.8 to 3 s^{-1}) are dependent on TMAO concentration, as are the amplitudes associated with each phase, as shown in panels A and B of Figure 4, respectively. Because the sum of the fast and middle phase amplitudes account for more than 95% of the total raw amplitude, these two phases describe the major folding kinetics of P protein. The third phase, which is around 0.10−0.15 s^{-1}, is independent of TMAO concentration, and the amplitude of the phase is small (< 5% of the total raw amplitude). We have assigned this slowest phase to cis−trans isomerization of Pro39 and/or Pro90 on the basis of the absence of the third phase in folding and unfolding of a P39A/P90A/F107W triple substitution variant of P protein (data not shown). All subsequent kinetic and equilibrium studies employed the proline-containing F107W variant, whose transients were fit to three exponentials, but the slowest (third) phase was not included in subsequent data analysis.

In many multiphase protein folding systems, the observed rate constants are usually at least 1 order of magnitude different from each other. Under these conditions, the nonlinear least-squares fitting method is robust and capable of accurately estimating both the rates and the associated amplitudes of each phase. However, an intrinsic challenge in kinetic data analysis of the P protein system originates from the similarity of the observed rate constants associated with the two major phases. The observed rate constants for the fast and medium phases in P protein folding processes are relatively close to each other (between 2- and 6-fold). When rate constants of two phases are close, the uncertainty of the parameters obtained from nonlinear square fits is large because the fitting parameters have a correlation coefficient close to 1. To best represent the uncertainty for each data point in Figure 4A, the same kinetic experiment was performed at least three times. The uncertainty was calculated for each condition from the fitting results of these independent experiments.

Because there are two major folding phases, the simplest mechanism consistent with the data is a three-state reaction with an on-pathway intermediate state, as shown in Scheme 1. Four kinetic rate constants, k_{UI}, k_{IU}, k_{IN}, and k_{NI}, are associated with this mechanism. To obtain the extrapolated values for all four rate constants, we attempted to fit their TMAO concentration dependence using eq 7. The initial fitting analyses from the kinetic data alone did not yield convergent estimations of the parameters because the two observed rate constants were similar and the uncertainty in the fastest rate constant was high. Equilibrium titration experiments were performed to provide more constraints for the fitting process.

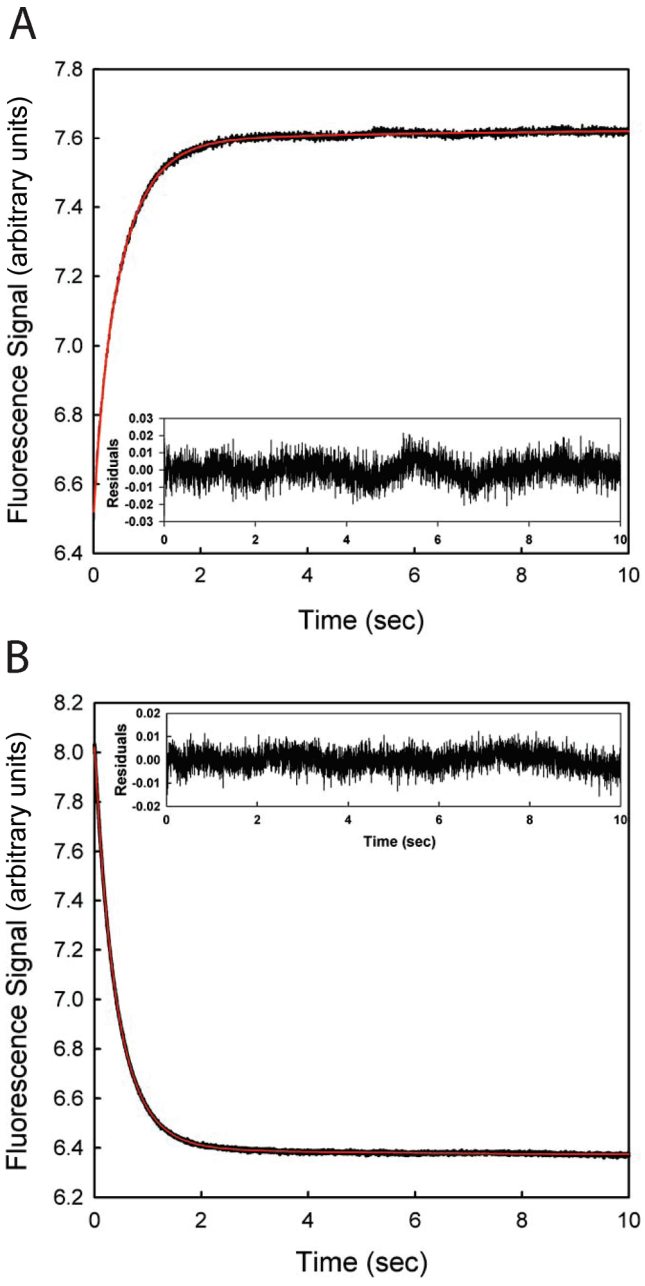

FIGURE 3: Folding and unfolding of P protein induced by changes in TMAO concentration and monitored by stopped-flow fluorescence. (A) Unfolding (down-jump) of P protein from 1 to 0.09 M TMAO. The kinetic trace of the unfolding process monitored by fluorescence is the black line. The red line through the data points is the best fit to a triple-exponential function. The inset shows the residuals of the curve fits. (B) Folding (up-jump) of P protein from 0 to 0.54 M TMAO. The red line through the black data points is the best fit to a triple-exponential function. The inset shows the residual of the fitting. Both data sets were collected at a rate of 500 points/s.

Urea−TMAO Fluorescence and CD Cotitration. Urea and guanidinium chloride denaturation titrations are widely used to estimate the protein stability, but the method is suitable only for stable proteins with both folded and unfolded baselines present in the titration curve. When this method is applied to IDPs, the absence of a complete cooperative transition and folded baseline in the titration curve makes it impossible to estimate the thermodynamic parameters using denaturation titration. To estimate the stability of P protein, we used mixtures of urea and TMAO as the denaturant and renaturant, respectively, to populate the folded or

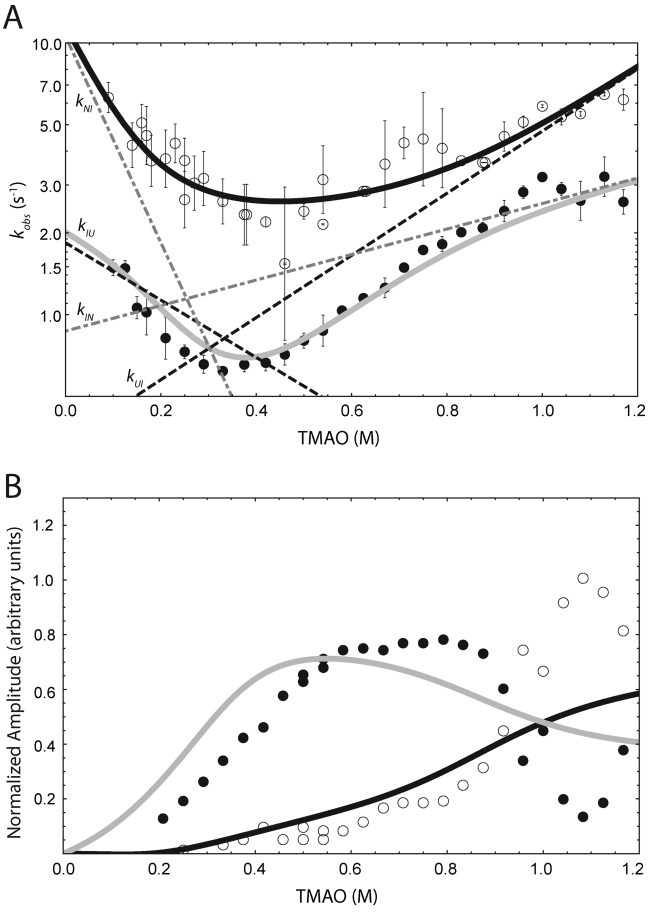

FIGURE 4: Folding and unfolding kinetics of F107W P protein as a function of final TMAO concentration. (A) TMAO concentration dependence of the natural logarithm of observed rate constants for both up-jump and down-jump experiments. The empty and filled circles are the fast and slow observed rate constants obtained from the folding (up-jump) experiments, respectively. The black dashed lines give rate constants k_{UI} and k_{IU} values as a function of TMAO concentration, and the gray dotted−dashed lines give k_{IN} and k_{NI} values as a function of TMAO concentration. (B) Raw amplitudes associated with fast and slow observed rate constants plotted as a function of TMAO concentration. Only the amplitudes of folding (up-jump) experiments are shown. Smooth lines through the observed rate constants are calculated from the parameters determined in the global fit of k_{obs} values and the equilibrium data (listed in Table 1) using the analytical solutions described by Bujalowski and Jezewska (*34*). The poor agreement between calculated and observed amplitudes at high TMAO concentrations is likely due to large uncertainties in the experimental amplitudes when the two phases have similar rate constants.

unfolded states of P protein. Titrations were initiated with a starting sample consisting of protein in urea or various urea/TMAO mixtures. Protein at the same concentration, in 1.4 M TMAO, was then added using an automated titrator, which resulted in progressive dilution of the starting urea concentration while the TMAO concentration increased. Six such titration curves allowed us to construct a three-dimensional urea−TMAO cotitration surface, which accessed a wide range of folded and unfolded populations. Each titration step yielded both fluorescence and CD data from the same protein sample using the same instrument. This procedure eliminated errors from discrepancies in protein, urea, or TMAO concentrations that might occur if the two types of data were collected separately. Panels A and B of Figure 5 show the TMAO dependence of the equilibrium fluorescence signal (λ_{ex} = 285 nm; λ_{em} > 320 nm) and CD signal (λ = 222 nm) for the F107W variant, respectively. Taken separately, fluorescence and CD titration curves can be fit to a two-state model, but the resulting fits give statistically significant different $C_{1/2}$ values. This observation reveals the presence of one or more significantly populated equilibrium intermediates in the folding process, consistent with the kinetic data. Our previous reports on P protein folding described the mechanism as two-state, based on statistically insignificant differences in $C_{1/2}$ values (*15*). The small difference, combined with the fact that the near- and far-UV CD data were collected on different samples, led us to erroneously suggest that the equilibrium intermediate is not present, when in fact it is (see below). However, by combining the equilibrium titration data with the stopped-flow data, we can globally fit the equilibrium data to a model based on the three-state mechanism shown in Scheme 1.

The experimental data were analyzed using eqs 2−6 (Experimental Procedures), which assume that the free energy of each transition is linearly dependent on both urea and TMAO concentration. In general, the CD and fluorescence signals of each state may vary with the urea or TMAO concentration. In accordance with convention, we assumed that the CD and fluorescence signals of states U, I, and N vary linearly with urea and TMAO concentration:

$$Sig_j^i([\mathbf{Urea}],[\mathbf{TMAO}]) = Sig_j^i(0) + s_j^{i,\mathrm{u}}[\mathbf{urea}] + s_j^{i,\mathrm{T}}[\mathbf{TMAO}] \quad (8)$$

where Sig represents the signal; *i* indicates the U, I, or N state; *j* indicates the CD or fluorescence signal; and $s_j^{i,\mathrm{u}}$ and $s_j^{i,\mathrm{T}}$ are the slopes of the baseplanes that describe the TMAO and urea dependence, respectively, of the signal of each state. Therefore, nine parameters for the fluorescence signals and nine parameters for the CD signals for a total of 18 spectroscopic parameters were used in the model. Fluorescence and CD titration data were simultaneously fitted using a three-state equilibrium model. Six thermodynamic parameters, ΔG^0_{UI}, ΔG^0_{IN}, $m^{\mathrm{T}}_{\mathrm{UI}}$, $m^{\mathrm{U}}_{\mathrm{UI}}$, $m^{\mathrm{T}}_{\mathrm{IN}}$, and $m^{\mathrm{U}}_{\mathrm{IN}}$, were used as global fitting parameters between the fluorescence and CD data sets. Fitting of the equilibrium data alone did not yield well-determined results for every parameter, especially the spectroscopic parameters associated with the intermediate state. For this reason, the uncertainties of free energies ΔG^0_{UI} and ΔG^0_{IN} are large. The poor estimation of these parameters mainly results from the large parameter space (total of 24 parameters) and the lack of two distinct transitions on the titration surfaces. This latter limitation indicates that either the intermediate state is poorly populated throughout the titration range or its spectroscopic signals are very similar to those of state U or N. To overcome these limitations in the data analysis, we estimated ΔG^0_{UI} and m^0_{UI} from NMR spectra recorded at various urea concentrations as described below. These parameters were held fixed in a global nonlinear least-squares fit of the remaining parameters to the combined kinetic and equilibrium data.

ΔG^0_{UI} and m^0_{UI} from NMR Urea Titration Experiments. NMR experiments were performed to estimate the equilibrium constant between the U and I states. More than 95% of the backbone amide ^{15}N−^{1}H resonances of the folded state P protein were assigned in a previous study (*40*). The presence of at least one alternative state was suggested by the observation of ~50 extra peaks in the ^{15}N−^{1}H HSQC spectrum of sulfate-folded P protein. In addition, the ^{15}N−^{1}H HSQC spectrum of P protein in the absence of sulfate [20 mM sodium cacodylate buffer (pH 7)] has some resonances whose positions are

Table 1: Kinetic and Equilibrium Parameters of Three-State Folding and Unfolding of P Protein

kinetic parameters[a]		equilibrium parameters[a]		spectroscopic parameters	
k_{UI}[b]	0.34 ± 0.05	ΔG^0_{UI}	1.0 ± 0.2	Sig^U_{CD}[f]	-3899 ± 339
k_{IU}[c]	1.8 ± 0.3	ΔG^0_{IN}	1.48 ± 0.1	Sig^I_{CD}	$-1 \times 10^4 \pm 1017$
k_{IN}	0.86 ± 0.1	m^{TMAO}_{UI}[e]	-3 ± 0.2	Sig^N_{CD}	$-1.29 \times 10^4 \pm 508$
k_{NI}[d]	10.5 ± 0.3	m^{TMAO}_{IN}	-5.8 ± 0.3	Sig^U_{FL}[g]	1 ± 0.01
β_{UI}	0.48 ± 0.02	m^{urea}_{UI}	1.5 ± 0.5	Sig^I_{FL}	0.1 ± 0.02
β_{IN}	0.89 ± 0.02	m^{urea}_{IN}	2.14 ± 0.2	Sig^N_{FL}	0.05 ± 0.01
		$m^{TMAO}_{UI}/m^{TMAO}_{UN}$[h]	0.34		
		$m^{urea}_{UI}/m^{urea}_{UN}$	0.41		

[a]All kinetic and equilibrium parameters are the values in the absence of urea and TMAO. [b]Units of inverse seconds. [c]Calculated using k_{IU}/K_{UI}. [d]Calculated using k_{IN}/K_{IN}. [e]Units of kilocalories per mole per molar. [f]CD signals in absence of urea or TMAO, units of $[\theta]_{222}$ (mean residue ellipticity). [g]CD signals in absence of urea or TMAO, arbitrary units. [h]Unitless.

similar to those found for the sulfate-folded protein. Both of these results suggest the existence of an alternative state in the P protein folding transition, which gives rise to extra HSQC peaks under both folding and unfolding conditions. As a working hypothesis, we assumed that this alternative state corresponds to the intermediate state observed in the kinetic and equilibrium experiments. In addition, we assumed that the folded state peaks observed in the sulfate-free sample correspond to residues that are in a foldedlike magnetic environment in the intermediate state. If correct, this interpretation suggests that the intermediate state has a significant population, comparable to that of the unfolded state. The equilibrium data suggest that the population of the folded state is negligible under these conditions, so addition of urea alters the NMR spectrum by perturbing the equilibrium between the U and I states. Thus, $^{15}N-^{1}H$ HSQC spectra were recorded at various urea concentrations. The relative populations of U and I states were estimated by peak intensity or volume change and used to determine the ΔG_{UI} as a linear function of urea concentration.

Figure 6A shows a region of the unfolded P protein HSQC spectrum, which contains some of the observed cross-peaks of the folded structure of P protein. These peaks were monitored in spectra obtained at urea concentrations ranging from 0 to 1 M. As the urea concentration increases, the intensities of these peaks decrease, indicating a decrease in the population of the intermediate state to which they are tentatively assigned. To estimate the population of the intermediate from these spectra, the folded state [in 20 mM sulfate (pH 7.0)] HSQC spectrum was used as the reference to determine the intensity of each peak in fully folded P protein. The peak intensities in all spectra were normalized to the intensity of the K119 amide peak, whose position is independent of sample conditions, making it an excellent internal standard. The normalized intensity of each intermediate peak was divided by the normalized intensity of the corresponding folded spectrum peak to produce an estimate of the fractional population of the intermediate at each urea concentration. These populations were used to calculate K_{UI}, which was converted to free energy using eq 2. To avoid line shape artifacts, peak volume was used to estimate peak intensity. Figure 6B shows the free energy calculated in this fashion for each monitored peak plotted versus urea concentration and fit to a linear function. The intercept of each line is free energy ΔG^0_{UI}, and the slope of the urea m value is m^0_{UI} for the U to I transition. The data points from four peaks gave ΔG^0_{UI} values ranging from 1.0 to 1.3 kcal/mol. The fact that urea m values exhibit larger variations between different peaks might be due to the large variation in the estimation of the peak volume for different peaks. Combined estimates of ΔG^0_{UI} (1.0 kcal/mol) and m^0_{UI} (1.5 kcal mol^{-1} M^{-1}) were obtained from a linear regression of all the data points, shown as a red line in Figure 6B. These two values were used as fixed parameter values in the global fitting of the combined equilibrium and kinetic data.

Global Fitting of Kinetic and Equilibrium Data. Because the kinetic or equilibrium data alone cannot provide an accurate estimation of all of the parameters for the model, a global analysis of both data sets was performed to estimate the kinetic and spectroscopic parameters together. The experimental data included the two observed rate constants and the two associated amplitudes at 33 different TMAO concentrations and the fluorescence and CD titration surfaces. Because preliminary NMR studies (data not shown) indicated that the magnetic environment of the Trp residue in the intermediate state may be similar to that of the unfolded state, we further simplified the fit by assuming that the TMAO dependence of fluorescence of the intermediate state is the same as that of the unfolded state. The two solid lines in Figure 4A are the best-fit results for the observed rate constants. As depicted in Figure 4A, the usual chevron shape of the plot of k_{obs} versus TMAO concentration is not observed because the two reaction steps in Scheme 1 have similar rate constants and are coupled to each other over most of the TMAO concentration range. The dashed lines and dotted−dashed lines in the plot are individual rate constants calculated as a function of TMAO concentration. At high TMAO concentrations (>1.1 M), the fast and slow phase k_{obs} values approximately equal k_{UI} and k_{IN}, respectively. The amplitudes were also analyzed using the model. The fitting results have similar general trends but do not fit well to the data points. This discrepancy might result from the large uncertainty in the fluorescence signal of each state from the fitting (Table 1) and is also likely due to large uncertainties in the experimental amplitudes when the two phases have similar rate constants. As shown in Figure 5, the best-fit model fit the equilibrium fluorescence and CD data quite well, although the denatured fluorescence baseplane did not fit as well as the CD. Inspection of the data in the denatured baseplane region shows differences in the denatured state fluorescence in different titrations. Although these differences could be caused by day-to-day variations in the fluorescence detection electronics, they also could result from small perturbations in the actual (not read) pH of the buffer caused by urea sensitivity of the pH electrode (*41*). The fluorescence of the F107W variant is sensitive to pH between pH 5 and 7 (data not shown), presumably because it is near the His105 residue whose ionization state could affect the fluorescence signal of the tryptophan fluorophore. Differences in the starting fluorescence of each titration curve could be due to slightly different pH values at each starting urea concentration.

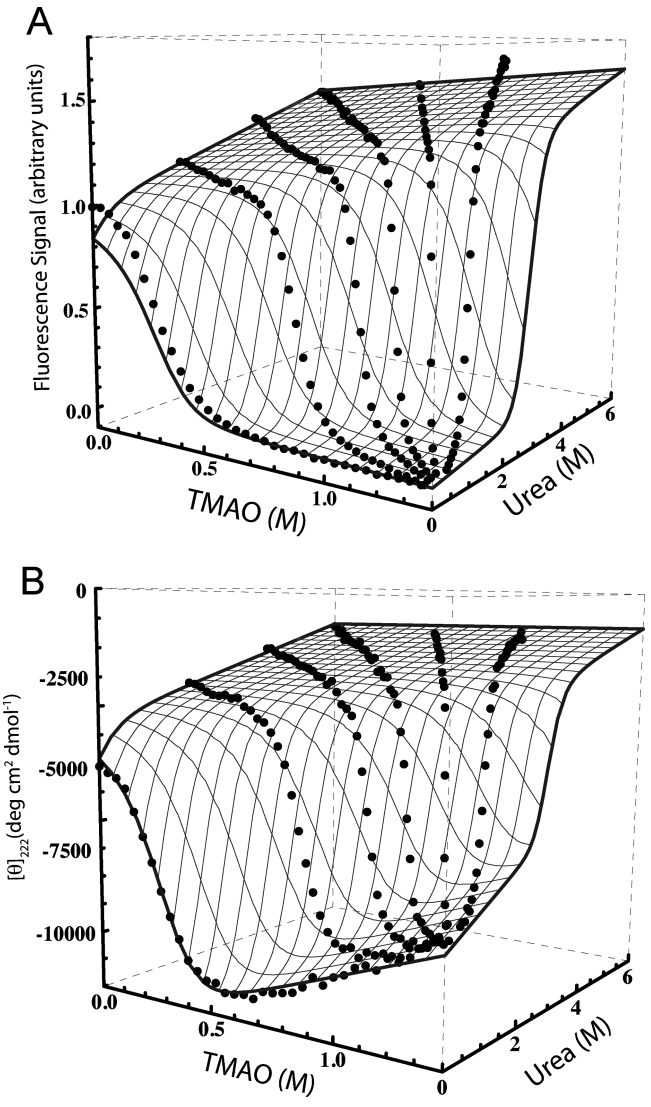

FIGURE 5: Fluorescence- and circular dichroism-detected equilibrium cotitration surface of the P protein (5 μM, 10 mm path length cuvette). The gray mesh surfaces on the two data sets are the best global fit results using a three-state model described by eq 4. (A) Fluorescence titration surface constructed by six titration curves plotted as a function of urea and TMAO concentration. The protein sample was excited at 285 nm, and the fluorescence signal was collected through a 320 nm high-pass filter. (B) CD titration surface constructed by six titration curves plotted as a function of urea and TMAO concentration. The CD signal was monitored at 222 nm. The best-fit equilibrium parameters from the model are listed in Table 1.

The kinetic and equilibrium parameters for the best-fit model are listed in Table 1. The free energy differences between U and I ($\Delta G^0_{UI} \cong 1$ kcal/mol) and I and N ($\Delta G^0_{IN} \cong 1.48$ kcal/mol) are approximately the same and predict populations for unfolded, intermediate, and native state populations of 83 ± 5, 16 ± 3, and 1 ± 0.2%, respectively. The ratios of TMAO and urea *m* values are −2.0 ± 0.3 and −2.7 ± 0.5 for the U ⇔ I and I ⇔ N transitions, respectively. Auton and Bolen have proposed a method for calculating *m* values on the basis of differences in the backbone and side chain solvent accessible surface areas (*42*). Because this calculation depends on side chain composition, there is no single expected ratio of TMAO to urea *m* values for different equilibria. Previous studies have shown that Barnase and Nank1-7* have ratios of −1.1 and −2.4, respectively (*31*). However, in general, TMAO has a stronger opposite effect on protein stability than urea (larger absolute *m* value).

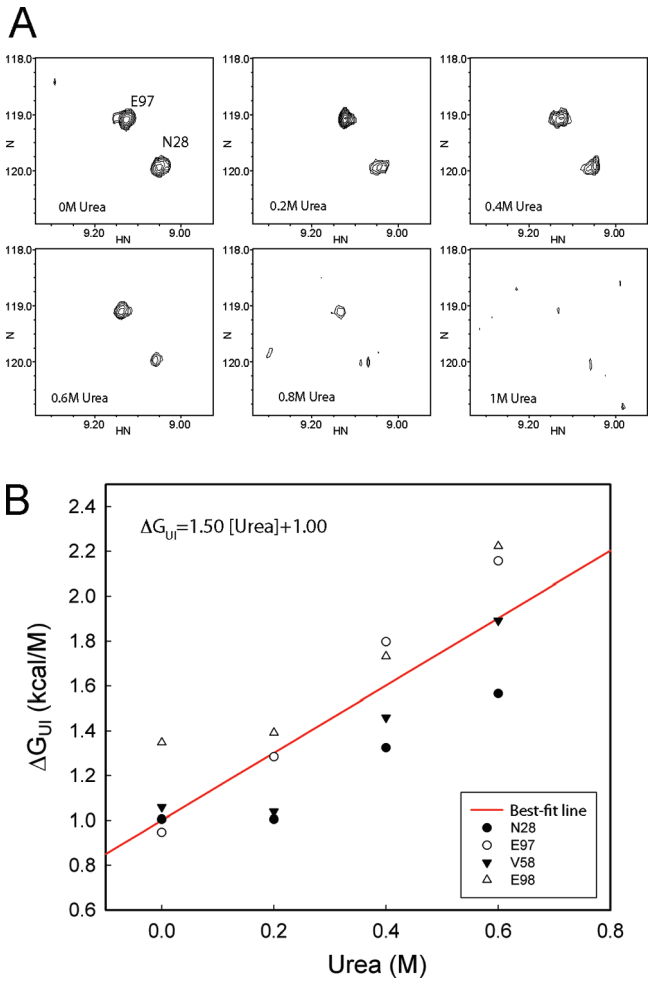

FIGURE 6: Urea titration of P protein in the absence of ligand [20 mM cacodylate buffer (pH 7)] monitored by the HSQC spectrum and estimation of the free energy difference between the U and I states. (A) A region of the HSQC spectrum is shown, and two peaks corresponding to N28 and R97 in a nativelike environment are labeled. The urea concentrations are given in each plot. (B) Estimation of the free energy between the U and I states from different residues vs urea concentration. The free energy was calculated from the population of the U and I states from the HSQC urea titration experiment. The line gives the best-fit results of all the data points in the plot using the linear function.

The best-fit values of both folding rate constants (k_{UI} = 0.34 s^{-1}, and k_{IN} = 0.86 s^{-1}) are relatively slow compared to those of other proteins that are similar in size (*1*). Further kinetic studies on other IDPs are necessary to determine if the slow refolding rate constant is a general property of the IDP folding reactions.

The best-fit model parameters include the intermediate state CD and fluorescence signals, which provide information about the structure of I. The best-fit value for the fluorescence signal of I is close to that of N, indicating that the environment surrounding Trp107 is similar in the intermediate and folded states. Moreover, the CD signal ($[\theta]_{222}$) of I is midway between those of U and N, suggesting that approximately 60−70% of the native state secondary structure, particularly helix, is formed in the intermediate state.

DISCUSSION

Comparison between Intrinsically Disordered Protein and Natively Folded Protein Kinetics. These kinetic studies of P protein show that the folding rate constant is lower than the

unfolding rate constant, as expected for an equilibrium that favors the unfolded state. Interestingly, the folding rate constant is quite low when compared to those of natively folded proteins that are similar in size and topology (*1*). Contact order is a measure of the topological complexity of a given protein's structure and a measurement of the number of contacts a residue makes with other residues that are local in sequence, relative to the number of contacts with residues that are distant (*43*). Thus, the contact order is related to secondary structure but is also influenced substantially by the tertiary fold. Plaxco et al. showed a correlation between the contact order and natural logarithm of the refolding rate constant and the position of the transition state. These correlations suggested that the lower the contact order the faster the folding rate constant and less foldedlike the transition state. From this correlation, the contact order of the P protein (14.3) predicts a folding rate constant (1000 s^{-1}) at least 2−3 orders of magnitude faster than our experimental estimates. This major discrepancy suggests that other factors besides topology, such as the charge−charge interactions in the protein, can be a major determinant of the actual folding kinetics. There are not many folding kinetic studies of IDPs in the literature, so it is too early to conclude whether slow folding (as opposed to fast unfolding) is a general property of such proteins. In the case of P protein, the low stability is at least partially due to unfavorable interactions in the more compact intermediate and folded state ensembles, many of which persist in the transition state ensembles ($TS^{\ddagger}_{UI}$ and $TS^{\ddagger}_{IN}$), thus raising the barrier to folding. This conclusion is supported by the high $\beta_{Tanford}$ values determined from the TMAO dependence of the rate constants (see Table 1), which indicate that the loss in solvent accessible surface area (Δ_{SASA}) for the U → $TS^{\ddagger}_{UI}$ transition is 48% of the Δ_{SASA} for the U → I transition and the Δ_{SASA} for the I → $TS^{\ddagger}_{IN}$ transition is 89% of that for the I → N transition (see Table 1). Thus, both transition state ensembles are significantly more compact than the preceding ground states, and unfavorable electrostatic interactions in these compact transition states are a likely explanation for slow folding.

Utility of TMAO (renaturant) Stopped-Flow and Cotitration Experiments as a General Method for Determining IDP Folding Mechanisms. Various protective osmolytes have been used as cosolutes to stabilize proteins and favor the folded state (*26*, *44*). Many studies have shown that some IDPs can fold to their folded states in the presence of their binding partners. Are the folded states of proteins induced by protective osmolytes the same as those induced by their binding partner? In this and previous studies (*15*), the CD spectrum and the fluorescence spectrum of the ligand-induced native state and the TMAO-induced native state of P protein are very similar. Furthermore, most of the recognizable peaks in the $^{1}H-^{15}N$ HSQC spectrum of the native state under these two conditions have similar chemical shifts, indicating the structures of the protein under the two conditions are very similar. This conclusion is consistent with previous studies of the effect of TMAO on the stabilities and structures of other proteins, which noted little perturbation on the folded structure (*27*, *45*). Other studies have shown that TMAO enhances the enzymatic activity (*46*, *47*), which makes large structural perturbations unlikely. For these reasons, addition of TMAO appears to be a useful way to shift the folding equilibrium of an IDP to favor the same folded conformation that is present at much lower concentrations in the absence of TMAO.

With the increasing number of reported IDPs and partially folded proteins, there is a need for systematic biophysical methods to measure the kinetic and thermodynamic properties of these proteins. Although some previous studies employed denaturant-induced protein unfolding kinetic experiments in the presence of various TMAO concentrations to investigate the stability effect (*48*, *49*), there have been no previous kinetic studies using TMAO to initiate IDP refolding. The stopped-flow measurements of TMAO-induced IDP folding and unfolding described in this study are the equivalent of traditional denaturant-induced folding and unfolding kinetics of natively structure proteins. The experimental approach described here should work well for the study of the folding mechanism of other IDPs.

Previous studies have demonstrated that TMAO combined with urea or guanidinium chloride denaturation can be used to properly estimate the free energy of the partially folded protein (*31*). The cotitration stability measurement method we developed in this study should be applicable to other IDPs. The highest TMAO concentration in the titration was 1.4 M, but in many circumstances, it could be used at much higher concentrations. The upper limit in this study was determined by the weak buffering capacity of our low-affinity buffer (20 mM sodium cacodylate). As shown in panels A and B of Figure 6, the six titration curves formed diagonal transects across the titration surface and effectively sampled the unfolded and folded baseplanes and the transition region between them. Although the presence of an intermediate is not immediately apparent on these surfaces, the simultaneous detection of CD and fluorescence during the equilibrium titration made it possible for the global fitting procedure to detect an intermediate in the folding reaction, measure its population, and provide information about its spectroscopic properties.

Role of the Partially Folded Intermediate in the Folding Mechanism. On the basis of the parameter values listed in Table 1, the population of the P protein intermediate state ensemble is around 16−20% of the total protein concentration in the absence of urea and TMAO. In addition, the m_{UI}/m_{UN} values for both urea and TMAO in Table 1 indicate that ~40% of the area buried in N is also buried in I, suggesting that I is significantly folded. The best-fit CD signal for I listed in Table 1 is significantly more negative than that for U, indicating that I has significant secondary structure. The secondary structure of the intermediate state of P protein explains the shape of the plot of the CD signal versus urea concentration along the 0 M TMAO line in Figure 2B. The curvature from 0 to 2 M urea can be explained by the equilibrium transition between the unfolded state and the intermediate state of the P protein. The population of the three states as a function of TMAO concentration is plotted in Figure 7 using the equilibrium parameters from the fitting. The population of the intermediate state reaches its maximum at 0.25 M TMAO. At that TMAO concentration, the population of I is comparable to those of the U and N states, and the relative amounts of three states are approximately equal. By these criteria, I should be considered an important subensemble in the P protein energy landscape, distinguishable from U and N by the kinetic barriers among the three states.

In addition to its role in the folding mechanism, the intermediate state ensemble may play a functional role in vivo in a process such as protein−RNA recognition during the holoenzyme assembly process. Given the highly basic nature of P protein, there is almost certainly some nonspecific binding of protein on the negatively charged RNA molecule. The partially folded P protein intermediate state may more efficiently search the binding pocket on the RNA than either fully folded or

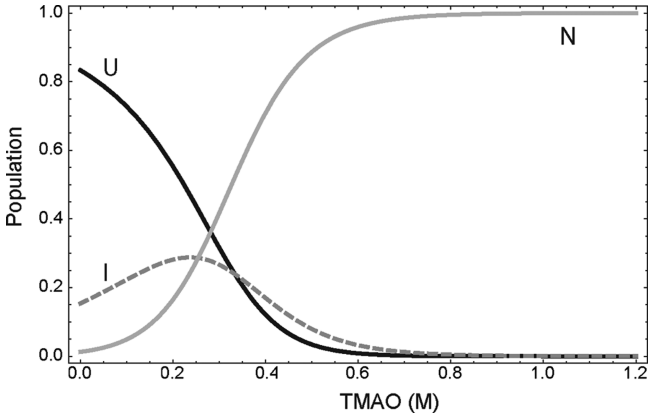

FIGURE 7: Population of U, I, and N states as a function of TMAO concentration. The population of each state was calculated using eq 3 and the equilibrium parameters (ΔG^0_{UI}, ΔG^0_{IN}, m^{TMAO}_{UI}, and m^{TMAO}_{IN}) listed in Table 1.

unfolded P protein, leading to more productive binding events in the formation of the holoenzyme. Future studies will be aimed at determining the structure of this important species and whether it has significant affinity for the two anionic ligands that bind with high affinity to N (*15*). Kinetic studies of P protein with small anionic ligands and a cognate binding partner, P RNA, will also provide insights about the holoenzyme assembly process.

ACKNOWLEDGMENT

We thank Drs. K. V. Rajagopalan and Michael Fitzgerald for generously sharing their instruments. We also thank Dr. Gordon Hammes for critical comments and for reading the manuscript and various members of the Oas laboratory for insightful discussion and advice.

REFERENCES

1. Jackson, S. E. (1998) How do small single-domain proteins fold? *Folding Des. 3*, R81–R91.
2. Clark, A. C. (2008) Protein folding: Are we there yet? *Arch. Biochem. Biophys. 469*, 1–3.
3. Dunker, A., Oldfield, C., Meng, J., Romero, P., Yang, J., Chen, J., Vacic, V., Obradovic, Z., and Uversky, V. (2008) The unfoldomics decade: An update on intrinsically disordered proteins. *BMC Genomics 9*, S1.
4. Fink, A. L. (2005) Natively unfolded proteins. *Curr. Opin. Struct. Biol. 15*, 35–41.
5. Uversky, V. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci. 11*, 739–756.
6. Wright, P., and Dyson, H. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol. 293*, 321–331.
7. Dyson, H. J., and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol. 12*, 54–60.
8. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197–208.
9. Wright, P. E., and Dyson, H. J. (2009) Linking folding and binding. *Curr. Opin. Struct. Biol. 19*, 31–38.
10. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell 35*, 849–857.
11. Frank, D. N., and Pace, N. R. (1998) Ribonuclease P: Unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem. 67*, 153–180.
12. Kurz, J., and Fierke, C. (2000) Ribonuclease P: A ribonucleoprotein enzyme. *Curr. Opin. Chem. Biol. 4*, 553–558.
13. Niranjanakumari, S., Kurz, J., and Fierke, C. (1998) Expression, purification and characterization of the recombinant ribonuclease P protein component from *Bacillus subtilis*. *Nucleic Acids Res. 26*, 3090–3096.
14. Stams, T., Niranjanakumari, S., Fierke, C., and Christianson, D. (1998) Ribonuclease P protein structure: Evolutionary origins in the translational apparatus. *Science 280*, 752–755.
15. Henkels, C. H., Kurz, J. C., Fierke, C. A., and Oas, T. G. (2001) Linked folding and anion binding of the *Bacillus subtilis* ribonuclease P protein. *Biochemistry 40*, 2777–2789.
16. Crary, S., Niranjanakumari, S., and Fierke, C. (1998) The protein component of *Bacillus subtilis* ribonuclease P increases catalytic efficiency by enhancing interactions with the 5′ leader sequence of pre-tRNAAsp. *Biochemistry 37*, 9409–9416.
17. Buck, A. H., Dalby, A. B., Poole, A. W., Kazantsev, A. V., and Pace, N. R. (2005) Protein activation of a ribozyme: The role of bacterial RNase P protein. *EMBO J. 24*, 3360–3368.
18. Hsieh, J., Andrews, A., and Fierke, C. (2004) Roles of protein subunits in RNA-protein complexes: Lessons from ribonuclease P. *Biopolymers 73*, 79–89.
19. Baldwin, R. L. (1996) On-pathway versus off-pathway folding intermediates. *Folding Des. 1*, R1–R8.
20. Wagner, C., and Kiefhaber, T. (1999) Intermediates can accelerate protein folding. *Proc. Natl. Acad. Sci. U.S.A. 96*, 6716–6721.
21. Capaldi, A. P., Shastry, M. C., Kleanthous, C., Roder, H., and Radford, S. E. (2001) Ultrarapid mixing experiments reveal that Im7 folds via an on-pathway intermediate. *Nat. Struct. Biol. 8*, 68–72.
22. Kathuria, S. V., Day, I. J., Wallace, L. A., and Matthews, C. R. (2008) Kinetic traps in the folding of $\beta\alpha$-repeat proteins: CheY initially misfolds before accessing the native conformation. *J. Mol. Biol. 382*, 467–484.
23. Kiefhaber, T. (1995) Kinetic traps in lysozyme folding. *Proc. Natl. Acad. Sci. U.S.A. 92*, 9029–9033.
24. Bolen, D. W., and Baskakov, I. V. (2001) The osmophobic effect: Natural selection of a thermodynamic force in protein folding. *J. Mol. Biol. 310*, 955–963.
25. Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., and Somero, G. N. (1982) Living with water stress: Evolution of osmolyte systems. *Science 217*, 1214–1222.
26. Wang, A., and Bolen, D. W. (1997) A naturally occurring protective system in urea-rich cells: Mechanism of osmolyte protection of proteins against urea denaturation. *Biochemistry 36*, 9101–9108.
27. Lin, S. L., Zarrine-Afsar, A., and Davidson, A. R. (2009) The osmolyte trimethylamine-N-oxide stabilizes the Fyn SH3 domain without altering the structure of its folding transition state. *Protein Sci. 18*, 526–536.
28. Lambert, D., and Draper, D. E. (2007) Effects of osmolytes on RNA secondary and tertiary structure stabilities and RNA-Mg^{2+} interactions. *J. Mol. Biol. 370*, 993–1005.
29. Bolen, D. W., and Rose, G. D. (2008) Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annu. Rev. Biochem. 77*, 339–362.
30. Baskakov, I., and Bolen, D. W. (1998) Forcing thermodynamically unfolded proteins to fold. *J. Biol. Chem. 273*, 4831–4834.
31. Mello, C. C., and Barrick, D. (2003) Measuring the stability of partly folded proteins using TMAO. *Protein Sci. 12*, 1522–1529.
32. Warren, J. R., and Gordon, J. A. (1966) On the Refractive Indices of Aqueous Solutions of Urea. *J. Phys. Chem. 70*, 297–300.
33. Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*, 1948–1954.
34. Bujalowski, W., and Jezewska, M. J. (2000) Kinetic mechanism of the single-stranded DNA recognition by *Escherichia coli* replicative helicase DnaB protein. Application of the matrix projection operator technique to analyze stopped-flow kinetics. *J. Mol. Biol. 295*, 831–852.
35. Fersht, A. (1999) Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding, W. H. Freeman, New York.
36. Gopalan, V., Golbik, R., Schreiber, G., Fersht, A. R., and Altman, S. (1997) Fluorescence properties of a tryptophan residue in an aromatic core of the protein subunit of ribonuclease P from *Escherichia coli*. *J. Mol. Biol. 267*, 765–769.
37. Chen, V. B., Davis, I. W., and Richardson, D. C. (2009) KING (Kinemage, Next Generation): A versatile interactive molecular and scientific visualization program. *Protein Sci. 18*, 2403–2409.
38. Szabo, A. G., Stepanik, T. M., Wayner, D. M., and Young, N. M. (1983) Conformational heterogeneity of the copper binding site in azurin. A time-resolved fluorescence study. *Biophys. J. 41*, 233–244.
39. Lakowicz, J. R. (1999) Principles of fluorescence spectroscopy, 2nd ed., Kluwer Academic/Plenum, New York.
40. Henkels, C. H. (2005) Biophysical studies of *Bacillus subtilis* ribonuclease P protein. Ph.D. Thesis, Duke University, Durham, NC.
41. Acevedo, O. (2002) pH Corrections in Chemical Denaturant Solutions. *Anal. Biochem. 306*, 158–161.

42. Auton, M., and Bolen, D. (2007) Application of the Transfer Model to Understand How Naturally Occurring Osmolytes Affect Protein Stability. *Methods Enzymol. 428*, 397–418.
43. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol. 277*, 985–994.
44. Russo, A. T., Rosgen, J., and Bolen, D. (2003) Osmolyte effects on kinetics of FKBP12 C22A folding coupled with prolyl isomerization. *J. Mol. Biol. 330*, 851–866.
45. He, H. L., Chen, X. L., Zhang, X. Y., Sun, C. Y., Zou, B. C., and Zhang, Y. Z. (2009) Novel Use for the Osmolyte Trimethylamine N-oxide: Retaining the Psychrophilic Characters of Cold-Adapted Protease Deseasin MCP-01 and Simultaneously Improving its Thermostability. *Mar. Biotechnol. 11*, 710–716.
46. Mashino, T., and Fridovich, I. (1987) Effects of urea and trimethylamine-N-oxide on enzyme activity and stability. *Arch. Biochem. Biophys. 258*, 356–360.
47. Kumar, R., Serrette, J. M., and Thompson, E. B. (2005) Osmolyte-induced folding enhances tryptic enzyme activity. *Arch. Biochem. Biophys. 436*, 78–82.
48. Banks, D. D., and Gloss, L. M. (2004) Folding mechanism of the $(H3\text{-}H4)_2$ histone tetramer of the core nucleosome. *Protein Sci. 13*, 1304–1316.
49. Topping, T. B., and Gloss, L. M. (2004) Stability and folding mechanism of mesophilic, thermophilic and hyperthermophilic archael histones: The importance of folding intermediates. *J. Mol. Biol. 342*, 247–260.
50. Davis, I. W., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2004) MOLPROBITY: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res. 32*, W615–W619.
51. Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., Richardson, J. S., and Richardson, D. C. (1999) Visualizing and quantifying molecular goodness-of-fit: Small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol. 285*, 1711–1733.