Articles

Urea and Thermal Equilibrium Denaturation Studies on the Dimerization Domain of *Escherichia coli Trp* Repressor†

Lisa M. Gloss and C. Robert Matthews*

Department of Chemistry and Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, Pennsylvania 16802

Received January 10, 1997; Revised Manuscript Received March 5, 1997⊗

ABSTRACT: The urea-induced equilibrium unfolding of the *Escherichia coli Trp* repressor (TR) is a two-state process, involving the native dimeric and unfolded monomeric species. Kinetic studies, however, reveal the presence of transient intermediates that appear only during the folding of the 107-residue protein [Gittelman, M. G., & Matthews, C. R. (1990) *Biochemistry 29*, 7011−7020]. In order to gain insight into the complex kinetic folding mechanism, the sequence of TR was reduced to the amino-terminal 66 residues, corresponding to the dimerization domain. Two polypeptides, 2−66 and NHis-7−66, were shown to be dimeric at 25 °C by size exclusion chromatography and to retain native-like spectroscopic features as evidenced by near- and far-UV circular dichroism and fluorescence spectroscopy. The equilibrium properties of the urea-induced folding of these core fragments were examined by intrinsic tryptophan fluorescence and circular dichroism and found to be well described by a two-state model. At 25 °C, the stabilities of both fragments are 14 kcal mol^{-1}, as compared to the 24 kcal mol^{-1} observed for full-length TR. In contrast, the thermal denaturation of $[2-66]_2$ and full-length TR are three-state processes; the midpoint of the transition monitored by absorbance at 292 nm precedes that monitored by circular dichroism at 222 nm. Global analysis of the thermal data as a function of monomer concentration suggests that both the full-length and $[2-66]_2$ TR variants unfold via a dimeric intermediate. Taken together, these results demonstrate that the $[2-66]_2$ fragment constitutes a well-structured, independently folding subdomain of TR that may be useful in elucidating the properties of the transient intermediates observed in the folding of the full-length protein. The dimeric intermediate observed in the thermal denaturation of $[2-66]_2$ suggests that it may be possible to further reduce the core sequence while maintaining the ability to dimerize.

Many of the structural features and thermodynamic properties of stable, well-folded proteins have been elucidated from studies of small, monomeric proteins (Baldwin, 1993; Matthews, 1993; Dill *et al.,* 1995). Understanding the relevant intramolecular forces has also provided insights into the folding and stability of oligomeric protein systems (Jaenicke, 1987). However, an accurate description of such higher order structures requires study of the features unique to intermolecular associations. For example, does the formation of an intermolecular hydrophobic interface differ from that of an intramolecular hydrophobic core? How does an oligomeric structure become efficiently and accurately assembled from individual polypeptide chains through a reaction governed by second or higher order kinetics? What are the thermodynamic and functional implications of the dependence of the midpoint of the equilibrium transition on protein concentration for an oligomeric system? Recent studies have elucidated some of the features important for the formation and stabilization of oligomeric protein complexes; these studies include both heterodimeric systems, such as the human growth hormone−receptor complex (Wells, 1996; Clackson & Wells, 1995), and homodimeric systems, such as the *Arc* repressor (Milla & Sauer, 1994, 1995) or the GCN4 leucine zipper peptide (Zitzewitz *et al*., 1995; Harbury *et al*., 1993).

Another homodimeric system that has yielded insights into protein−protein association reactions is the *Escherichia coli Trp* repressor (TR) (Gittelman & Matthews, 1990; Mann & Matthews, 1993; Mann *et al*., 1995). This prototypical member of the helix−turn−helix family of DNA-binding proteins regulates the transcription of four operons involved in the biosynthesis of tryptophan in *E. coli*. The primary structure contains 108 residues per monomer, and the secondary structure is composed entirely of α-helices separated by turns or short loops (Zhang *et al*., 1987; Zhao *et al*., 1993). The hydrophobic core/dimerization domain of TR consists primarily of the A, B, and C helices of each monomer intertwined about each other (Figure 1). The F helices are docked against this core and also contribute intermonomer contacts. The remaining two pairs of helices, DE and D′E′, flank the core domain and, with the intervening turn, constitute the DNA reading heads.

The reversible folding of WT TR, induced by urea, is a two-state equilibrium process, with concerted disruption of

† Supported by NIH Grants GM 23303 and GM 54836 to CRM LMG was supported in part by NIH Postdoctoral Fellowship GM 16685.

* Author to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* April 15, 1997.

S0006-2960(97)00056-1 CCC: $14.00 © 1997 American Chemical Society

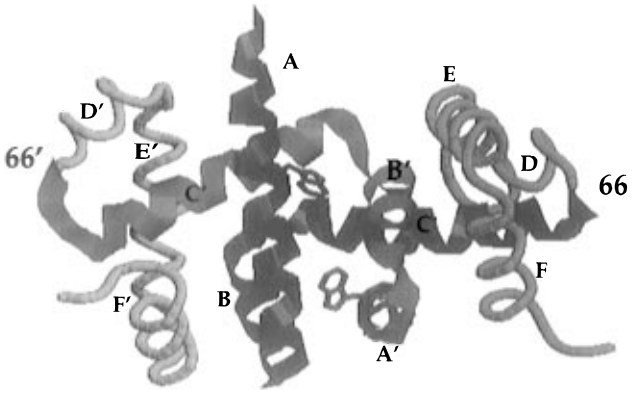

FIGURE 1: Three-dimensional structure of full-length WT *Trp* repressor from the X-ray crystal structure (Zhang *et al*., 1987). The $[2-66]_2$ core variant spans helices A, B, C, A′, B′, and C′, drawn as ribbons; the terminal residues, 66 and 66′, are indicated. The remaining helices, which are not in the [2−66] core, are drawn as thin coils. The side chains of Trp-19 and Trp-19′ are shown extending from the ribbon of helices A and A′.

secondary and tertiary structure (Gittelman & Matthews, 1990). The kinetic folding process, however, is much more complicated. The proposed mechanism postulates three parallel channels, each of which contains multiple steps involving monomeric and dimeric intermediates and both proline and non-proline isomerization reactions (Mann & Matthews, 1993; Mann *et al*., 1995). The absence of equilibrium intermediates facilitates the quantitative measurement of thermodynamic parameters but provides no insights into the structures of the transient kinetic intermediates that would define a folding pathway.

In an effort to populate and characterize equilibrium intermediates that may be relevant to those seen transiently in kinetic studies, two approaches were taken in the experiments reported herein: studying the behavior of fragments corresponding to the dimeric, core subdomain of the full-length protein and applying an alternative denaturation method to full-length TR and its fragments.

The first approach, fragmentation, is an emerging and powerful technique for breaking down the cooperativity of an unfolding transition—by identifying autonomous folding units—or for dissecting a kinetic folding reaction in order to elucidate events along the folding pathway. For example, NMR and CD[1] studies of a 131-residue fragment of staphylococcal nuclease have revealed features of a partially folded structure not populated by the full-length protein (Alexandrescu *et al*., 1994; Shortle *et al*., 1996). Fragments have also been made of the small, monomeric proteins, barnase and barley chymotrypsin inhibitor 2, by cyanogen bromide cleavage at a methionine residue (Kippen *et al*., 1994; Kippen & Fersht, 1995; Ruiz-Sanz *et al*., 1995). The equilibrium and kinetic studies of these two fragment systems have complemented and enhanced mutational analyses of the transition states of the folding mechanisms of the full-length, intact proteins (Kippen *et al*., 1995; Ruiz-Sanz *et al*., 1995). Other recent examples include fragments from yeast phosphoglycerate kinase (Pecorari *et al*., 1996) and cytochrome *c* (Wu *et al*., 1993).

A core fragment of *E. coli* TR, spanning residues 8−71, had previously been generated by limited chymotryptic cleavage of full-length TR (Carey, 1989). A preliminary structural characterization of this polypeptide demonstrates that it retains much of its helical structure, exhibits native-like dispersion of NMR resonances, and is dimeric (Tasayco & Carey, 1993). The exact boundaries of the 8−71 core are a compromise between the divisions of the three-dimensional structure of TR into discrete domains and the substrate specificity of chymotrypsin. For example, the 8−71 fragment extends beyond the C helix and into the middle of the D helix (Figure 1). Although, proteolytic digestion is a powerful technique for identifying domain boundaries in multidomain proteins (Wu *et al*., 1994), the introduction of stop codons by site-directed mutagenesis permits a more precise coordination of fragment boundaries with units of secondary structure. In this report, a core fragment was constructed by introduction of a pair of stop codons at positions 67 and 68 in the *TrpR* gene. Ser-67 is an optimal site for stop codon introduction because it maintains the full sequence comprising the C helix (Zhang *et al*., 1987) and its canonical C-cap sequence (Aurora *et al*., 1994), without further extension of the polypeptide chain into the following secondary structural unit. At the DNA level, the core fragment spans residues 1−66; however, it was anticipated that, similar to WT TR, the N-terminal formyl-Met would be processed (Joachimiak *et al*., 1983), yielding a protein product corresponding to $[2-66]_2$.

The second approach employed for populating equilibrium intermediates that may be relevant to kinetic intermediates was thermal denaturation. Discrepancies between the observed calorimetric and van't Hoff enthalpies in differential scanning calorimetry experiments on WT TR (Bae *et al*., 1988) suggest that a partially folded intermediate is populated during thermal denaturation. However, the remarkable stability of WT TR ($T_{1/2} \geq 90$ °C at dimer concentrations above 70 μM) did not permit an unambiguous determination of the oligomeric state of the populated thermal intermediate. With far more sensitive optical techniques, *e.g.*, circular dichroism and absorbance, it is possible to monitor thermal denaturation at much lower protein concentrations. The anticipated decrease in the midpoint for the thermal unfolding transition should permit a more accurate thermodynamic analysis, including determination of the oligomeric state of the suspected thermal intermediate.

In this report, the urea-induced denaturations of the fragment, $[2-66]_2$, and a derivative of it, $[NHis\text{-}7-66]_2$, are presented. The thermal denaturation behavior of $[2-66]_2$ and WT TR are also reported herein.

MATERIALS AND METHODS

Materials

Ultrapure urea was purchased from ICN Biomedicals Inc. (Aurora, OH). Mutagenic oligonucleotides were purchased from Midland Certified Reagent Co. (Midland, TX). Iminodiacetic acid resin was obtained from Sigma Chemical Co. (St. Louis, MO). All other chemicals were of reagent grade.

[1] Abbreviations: $[2-66]_2$, dimeric *Trp* repressor fragment containing residues 2−66; $[NHis\text{-}7-66]_2$, dimeric *Trp* repressor fragment with six histidine residues on the N-terminus and containing WT residues 7−66; IAEW, fluorescence intensity-averaged emission wavelength; Abs, UV−visible absorbance; CD, circular dichroism; DSC, differential scanning calorimetry; F_{app}, apparent fraction of unfolded protein; FL, fluorescence; FPLC, fast protein liquid chromatography; MRE, mean residue ellipticity; MW, molecular weight; TR, *Trp* repressor; WT, dimeric wild-type *Trp* repressor containing residues 2−108.

Methods

Site-Directed Mutagenesis. Expression of the core fragments was not observed when double-stop mutations were introduced into the TR expression plasmid, pJPR2 (Pulah & Yanofsky, 1986), in which transcription is under the control of a *tac* promoter. Instead, a plasmid in which the *Trp* repressor gene (*TrpR*) is under the control of a T_7 promoter was employed. Such a plasmid was constructed from pET13atrpr3, a generous gift from Dr. Catherine L. Lawson (N. C. Combatti and C. L. Lawson, unpublished results); pET13atrpr3 is a derivative of the kanamycin-resistance, T_7 plasmid pET13a (Rosenberg *et al*., 1987), containing the *TrpR* gene. A smaller pUC119-derived plasmid that encodes ampicillin resistance and has an M13 origin of replication, pT_7RUC, was made by ligating the *Bgl*II−*Eco*RV fragment of pET13atrpr3 (containing the T_7 promoter and *TrpR* open reading frame) into *Bam*HI−*Hin*cII-linearized pUC119 (Viera & Messing, 1987). All further molecular biology manipulations were performed on pT_7-RUC.

The codons for Ser-67 and Gln-68 were mutated to *opal* and *ochre* stop codons, respectively, by Kunkel mutagenesis (Kunkel, 1985) to generate pT_7RUC·67op68oc for the production of the 2−66 peptide. Additionally, a second core peptide was engineered to take advantage of polyhistidine affinity chromatography on Ni^+−iminodiacetic agarose resin (Porath *et al*., 1975). The eight N-terminal residues of full-length TR appear to be disordered in both the NMR (Zhao *et al*., 1993; Zheng *et al*., 1995) and the X-ray structures (Zhang *et al*., 1987). This interpretation is supported by the susceptibility of the peptide bond at Tyr-7 to chymotryptic cleavage (Carey, 1989). Folded 8−108 TR (armless) has a far-UV CD spectrum and urea denaturation behavior similar to those of WT TR (Carey, 1989), suggesting that the N-terminal arms of TR play little role in protein stability. Thus, to aid in purification, a tag of six histidines was engineered onto the N-terminus of the core fragment. The N-terminal sequence of the *TrpR* gene was mutated from Met-Ala-Gln-Gln-Ser-Pro-Tyr to Met-$(His)_6$-Tyr to produce NHis-7−66. The presence of the desired mutations and the absence of any unintended mutations were confirmed by dideoxy sequencing (Sequenase kit, United States Biochemical Co., Cleveland, OH) of the entire *TrpR* gene.

Protein Expression and Purification. WT TR was expressed from pJPR2 in the *E. coli* host CY-15070 and purified as described by Pulah and Yanofsky (1986). The $[2-66]_2$ and $[NHis\text{-}7-66]_2$ fragments were expressed from the above-described pT_7RUC plasmids in the *E. coli* host BL21(DE3) (Studier *et al*., 1990). Cells were grown to mid-log phase, either in 1 L fluted flasks or an 80 L fermentor, and protein production was induced by the addition of IPTG to 0.5 mM. After growth to saturation, the cells were harvested. The core fragments were expressed as inclusion bodies.

To recover the insoluble fragments, a method modified from that of Tsuji *et al*. (1993) was employed. In brief, the cells were resuspended in 100 mM potassium phosphate and 1 mM EDTA, pH 7.6, and lysed by sonication. The soluble protein fraction was separated by centrifugation at 4 °C. The insoluble protein pellet was solubilized by stirring in 6 M urea, 20 mM potassium phosphate, and 1 mM EDTA, pH 7.6 at 4 °C. After centrifugation, the supernatant contained the core fragments. All subsequent purification steps were also performed at 4 °C.

The supernatant of $[NHis\text{-}7-66]_2$ was then applied directly to an immobilized nickel column (iminodiacetic acid resin with chelated Ni^{2+}, 60 mL resin volume) equilibrated in 6 M urea, 100 mM Tris, and 100 mM KCl, pH 7.5. The $[NHis\text{-}7-66]_2$ was eluted with a 0−300 mM imidazole gradient in the equilibration buffer.

$[2-66]_2$ was precipitated from the 6 M urea solubilization buffer by addition of ammonium sulfate to 50%. The protein pellet was resuspended in and dialyzed against 5 M urea, 100 mM Tris, and 1 mM EDTA, pH 7.6. If necessary, the dialysate was concentrated with an Amicon flow cell (YM-3 membrane) before application to a Sephadex G-100 size-exclusion column (2 × 100 cm) equilibrated in dialysis buffer. After chromatography, the purified fragments were folded by dialysis into 10 mM potassium phosphate and 0.5 mM EDTA, pH 7.6, and then lyophilized for storage. The purity of the TR variants was examined by sodium dodecyl sulfate−polyacrylamide gel electrophoresis and judged to be ≥95%. The extinction coefficients at 280 nm of $[2-66]_2$ and $[NHis\text{-}7-66]_2$ fragments were determined by the method of Gill and von Hippel (1989) and found to be 1.02 and 1.10 mL mg^{-1} in 8 and 0 M urea, respectively.

Analytical size-exclusion chromatography was performed with a Pharmacia FPLC system, using a Superose 12 column, equilibrated at 25 °C in 100 mM potassium phosphate buffer, pH 7.6. Molecular weight standards were aldolase (160 kD), bovine serum albumin (66.3 kD), ovalbumin (44 kD), carbonic anhydrase (28.8 kD), hen egg white lysozyme (14.3 kD), ribonuclease A (13.7 kD), and insulin (5.8 kD). Samples for electrospray mass spectrometry were submitted to the Protein Chemistry Laboratory at the School of Medicine, Washington University, in St. Louis, MO.

Urea Denaturation Experiments. The buffer conditions for equilibrium folding experiments were 10 mM potassium phosphate and 0.1 mM EDTA, pH 7.6, with varying concentrations of ultrapure urea at 25 °C. CD and fluorescence spectra were collected on an Aviv Model 62DS CD spectrophotometer and a Spex-Fluorolog 1.81 m spectrophotometer, respectively. The excitation wavelength for the fluorescence experiments was 295 nm, with a detection cutoff filter of 305 nm. For investigation of the equilibrium spectral properties, individual samples were made for each urea concentration and incubated at 25 °C for ≥1 h to ensure that the samples were completely equilibrated.

Data Analysis of Urea Denaturation Experiments. The urea-induced unfolding of TR variants was monitored by CD, using the ellipticity at 222 nm, or fluorescence emission. The fluorescence data were analyzed by two methods: fluorescence intensity at 330 nm, the emission maximum of all folded TR variants, and intensity-averaged emission wavelength (IAEW) (Royer *et al*., 1993). The latter parameter, defined by

$$\langle\gamma\rangle = \sum_{i=1}^{N}(I_i\gamma_i)/\sum_{i=1}^{N}(I_i) \qquad (1)$$

reflects urea-induced changes in the shape and the λ_{max} of the spectra. As it is an integrated measure over the entire spectrum, it is less prone to error from instrumental noise

or vagaries in sample preparation than CD or FL monitored at a single wavelength.

Local fits, *i.e.*, fits of individual data sets at a single monomer concentration monitored by either CD or FL, were performed with the NLIN program of the SAS statistical package (SAS Institute, Cary, NC). The data were fitted to the equation describing the two-state dimer model, for native dimer (N_2) in equilibrium with unfolded monomers (2U):

$$N_2 \overset{K_{eq}}{\rightleftharpoons} 2U \tag{2}$$

where $K_{eq} = [U]^2/[N_2]$. A linear dependence of $\Delta G°$ on the concentration of urea was assumed (Pace, 1986): $\Delta G° = \Delta G°(H_2O) - m[\text{urea}]$, where $\Delta G°(H_2O)$ is the free energy difference between N_2 and U in the absence of denaturant and m reflects the sensitivity of the transition to denaturant concentration. For comparison of data at different monomer concentrations or collected by different spectral methods, data were normalized to F_{app}, the apparent fraction of unfolded protein, where

$$F_{app} = \frac{(Y_o - Y_f)}{(Y_u - Y_f)} \tag{3}$$

Y_o is the observed optical signal at a given urea concentration; Y_f and Y_u are the observed values for the folded and unfolded species at the same urea concentration. The terms Y_f and Y_u are, in turn, defined by linear equations accounting for the urea dependence of the native and unfolded baselines, respectively. For a detailed description of the derivation of these relationships and the model, see Gittelman and Matthews (1990). Global fits, *i.e.*, simultaneous fits of multiple data sets collected by CD and FL at varying concentrations of monomer, were performed with in-house software based on the analysis package Savuka 5.0 (D. Lambright and O. Bilsel, personal communication; Zitzewitz *et al.*, 1995). The values of $\Delta G°(H_2O)$ and m were globally constrained, but those of Y_f and Y_u were treated as adjustable local parameters.

Thermal Denaturation Experiments. The buffer conditions for the thermal denaturation experiments were the same as those used for urea unfolding (10 mM potassium phosphate, 0.1 mM EDTA, pH 7.6). The monomer concentration ranges were 2.6−11 μM and 7−30 μM for WT TR and $[2-66]_2$, respectively. Far-UV CD data at 222 nm were collected with an Aviv Model 62DS CD spectrophotometer with a Hewlett-Packard 89100A temperature controller. UV absorbance data at 292 nm were collected on a Cary 14DS UV−vis spectrophotometer modified by Aviv with a thermal electric cell holder. For CD data, cell path lengths of 0.1−0.5 cm were used, depending on the monomer concentration; for absorbance data, a 1 cm cell was employed. For both types of spectroscopy, the cells were sealed with Teflon stoppers and parafilm so that no evaporation occurred. Samples were heated in increments of 1 °C from 15 to 95 °C and equilibrated for 3−6 min at each temperature before the spectral data were recorded. The complete equilibration of the TR samples was demonstrated by the independence of the unfolding transition on the equilibration time used. The accuracy of the temperature control and rate of temperature equilibration were verified in control experiments in which a thermocouple was inserted into the cuvettes. Near-UV CD spectra of $[2-66]_2$, at several temperatures, were obtained with a 10 cm cell containing 15 μM monomer. The cell was equilibrated at a given temperature for 15−30 min before data acquisition and scanned three to five times. Equilibration was verified by the superimposition of the first and last spectra, collected over the course of 1 h.

Thermal Data Analysis. Individual thermal unfolding transitions monitored by either CD 222 nm or Abs 292 nm, at a single monomer concentration, could be fit to a two-state model as described by eq 2. However, for both $[2-66]_2$ and WT TR, the F_{app} curves (eq 3) for CD and Abs data were not coincident, demonstrating that at least one populated intermediate exists. If the intermediate is dimeric, the following three-state model is appropriate:

$$N_2 \overset{K_1}{\rightleftharpoons} I_2 \overset{K_2}{\rightleftharpoons} 2U \tag{4}$$

where $K_1 = [I_2]/[N_2]$ and $K_2 = [U]^2/[I_2]$. Given that the total protein, P_t, in terms of monomers is $P_t = 2N_2 + 2I_2 + U$, the concentration of unfolded monomer, U, can be expressed as

$$[U] = \frac{-K_1K_2 + \sqrt{(K_1K_2)^2 + 8P_tK_1K_2(1 + K_1)}}{4(1 + K_1)} \tag{5}$$

Thus, the fraction of each species is defined as

$$X_u = [U]/P_t \tag{6a}$$

$$X_{I_2} = 2[U]^2/K_2P_t \tag{6b}$$

$$X_{N_2} = 1 - X_{I_2} - X_u \tag{6c}$$

and the expression for F_{app} in a three-state model is

$$F_{app} = X_u + ZX_{I_2} \tag{7}$$

where Z is a normalized parameter that describes the extent to which the intermediate state optically resembles the unfolded state; *i.e.*, if, by a given optical measure, I_2 resembles N_2, $Z = 0$, but if I_2 resembles U, then $Z = 1$.

If the thermal intermediate is monomeric, the appropriate model is

$$N_2 \overset{K_1}{\rightleftharpoons} 2I \overset{K_2}{\rightleftharpoons} 2U \tag{8}$$

where $K_1 = [I]^2/[N_2]$ and $K_2 = [U]^2/[I]^2$. Given that the total protein, P_t, in terms of monomers is $P_t = 2N_2 + I + U$, the concentration of unfolded monomer, U, can be expressed as

$$[U] = \frac{-K_1K_2 + \sqrt{(K_1K_2)^2 + 4P_tK_1K_2(2 + K_1)}}{2(2 + K_1)} \tag{9}$$

The definitions for X_u, X_{N_2}, and F_{app} are those given in eqs 6a, 6c, and 7, but the fraction of monomeric intermediate is

$$X_I = \frac{\sqrt{U^2/K_2}}{P_t} \tag{10}$$

For both three-state models (eqs 4 and 8), the observed

optical signal, Y_o, is described in terms of folded and unfolded baselines, Y_f and Y_u, respectively, by the expression:

$$Y_o = Y_f + F_{app}(Y_u - Y_f) \quad (11)$$

which is readily derived from eq 3.

The equilibrium constants, K_1 and K_2, can be cast in terms of thermodynamic parameters, such that

$$K_i = \exp(-\Delta G^\circ_i/RT) \quad (12a)$$

$$K_i = \exp\left\{\frac{T^\circ_i/T(\Delta C_{pi} - S^\circ(T^\circ)_i) + \Delta C_{pi} + S^\circ(T^\circ)_i + \Delta C_{pi}\ln(T/T^\circ_i)}{R}\right\} \quad (12b)$$

where *i* represents the parameters for either the native-to-intermediate or intermediate-to-unfolded transition. This expression assumes that ΔC_p is independent of temperature; this independence is a reasonable assumption over the temperature range examined, as discussed by others (Becktel & Schellman, 1987; Makhatadze & Privalov, 1995). More detailed or alternative derivations of eqs 4−12 have been presented elsewhere (Becktel & Schellman, 1987; Thompson *et al*., 1993; Eftink *et al*., 1994).

The thermal data were fitted by global analysis with the Savuka 5.0 software. In the fitting procedure, the thermodynamic parameters, as well as Z_{CD222} and Z_{Abs292}, were globally constrained. The baselines of each trace, Y_f and Y_u, were treated as local adjustable parameters in the fitting of $[2-66]_2$. Fitting of the WT TR data was more difficult because of the lack of an unfolded baseline in the CD transitions above 2.6 μM monomer. Initially, the Abs Y_f and Y_u and CD Y_f baselines were fixed at the values determined for each scan from linear fits of the data from 15 to 40 °C and 85 to 95 °C for Y_f and Y_u, respectively. The unfolded CD baselines were fitted in one of three ways: (1) the baseline was fixed at the values obtained by linear fits to the 2.6 μM data from 87 to 95 °C; (2) the baseline was fixed at values calculated from the thermally unfolded mean residue ellipticity of $[2-66]_2$, adjusted for the difference in polypeptide length between WT TR and $[2-66]_2$; or (3) the unfolded MRE at 25 °C was fixed at the value determined in 8 M urea (Figure 2), and the slope of the baseline was treated as a globally fitted parameter between the CD traces. The thermodynamic values obtained from the fits were the same within error for all methods of treating the baselines and converged to reduced χ^2 values of 4−6. In the final analysis of the WT TR data set, all values of Y_f and Y_u, for both CD and Abs scans, were treated as locally adjustable parameters. The reduced χ^2 decreased to 2.01, and all thermodynamic parameters (given in Table 2) converged to values similar to those observed for fits without locally adjustable baselines.

RESULTS

Structural Characteristics of the Core Fragments. The molecular weights of the monomeric units of core fragments were determined by electrospray mass spectrometry. The MW of $[NHis\text{-}7-66]_2$ was found to be 8067 ± 2, compared to a value of 8057 calculated for the expected gene product, with an N-terminal methionine present. $[2-66]_2$ exhibited a MW of 7626.1 ± 0.3, which is in good agreement with the calculated value of 7614 for a mature protein, lacking an N-terminal Met. The N-terminal processing of the Met-1 residue is also seen for WT TR (Joachimiak *et al*., 1983).

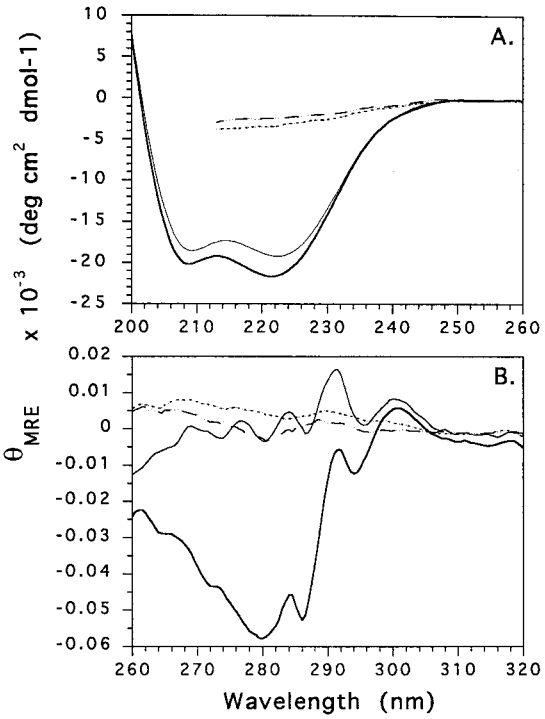

FIGURE 2: CD spectra of full-length WT and $[2-66]_2$ core *Trp* repressors. (A) Far-UV CD spectra of $[2-66]_2$ (0 M urea, bold solid line; 7 M urea, dotted line) and WT TR (0 M urea, thin solid line; 8 M urea, dashed line). (B) Near UV-CD spectra of $[2-66]_2$ (0 M urea, bold lower solid line; 8 M urea, dotted line) and WT (0 M, thin solid line; 8 M urea, dashed line). Conditions: 15 μM monomer, 25 °C, 10 mM potassium phosphate, 0.1 mM EDTA, pH 7.6.

Because residues 8−66 in the primary structure of WT TR contribute the majority of the intermolecular contacts in the full-length dimer (Zhang *et al*., 1987), it was reasonable to suppose that $[2-66]_2$ might form a dimer. Retention of a dimeric structure by the $[2-66]_2$ fragment was verified by size-exclusion chromatography. The observed molecular weight of $[2-66]_2$, 14 ± 1 kD determined from its elution volume compared to those of proteins of known molecular weight (see Methods), is in reasonable agreement with the calculated dimeric molecular weight of 15.5 kD.

The circular dichroism and fluorescence spectra of TR variants are a convenient, low-resolution method for comparing secondary and tertiary structure. Comparison of the far-UV CD spectra of WT and $[2-66]_2$ under folding conditions (0 M urea, Figure 2A) demonstrates that $[2-66]_2$, like WT, has a CD spectrum typical of an α-helical protein, with double minima at 208 and 222 nm. The mean residue ellipticity at 222 nm of the core fragments (−21.7 and −21.0 mdeg cm dmol^{-1} for $[2-66]_2$ and $[NHis\text{-}7-66]_2$, respectively) is slightly higher than that of the full-length protein (−19.4 mdeg cm dmol^{-1}), implying that residues 67−108 have a lower average helical content than residues 2−66. The increased MRE_{222} of the core fragments most likely reflects the frayed structures of the D and E helices in the full-length protein. The NMR structure of *Trp* repressor, in the absence of corepressor, L-tryptophan, shows that the D and E helices are poorly defined and very mobile (Zhao *et al*., 1993). These helices are not part of the $[2-66]_2$ fragment, which also lacks the two turns between the D−E and E−F helices.

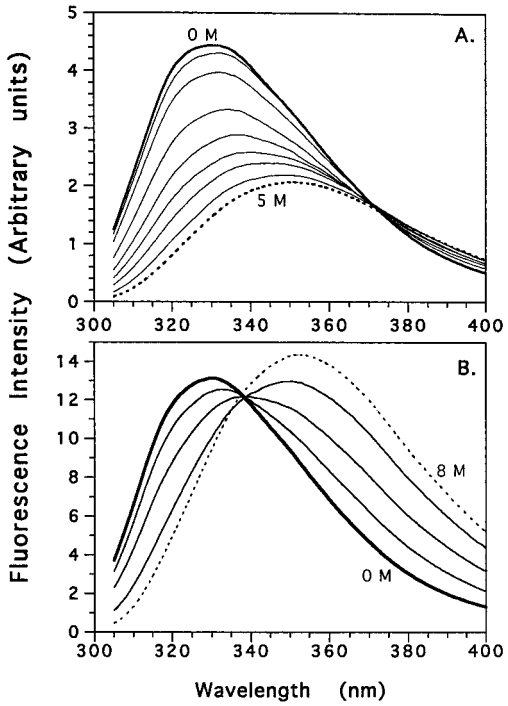

FIGURE 3: Fluorescence spectra of $[2-66]_2$ core and full-length *Trp* repressors as a function of urea concentration. (A) $[2-66]_2$ at 0 (bold line), 2.0, 2.5, 3.0, 3.25, 3.5, 3.75, 4.0, and 5.0 M (dotted line) urea, in order of decreasing intensity below 370 nm. (B) WT at 0 (bold line), 5, 5.5, 6, and 8 M (dotted line) urea, in order of decreasing intensity at 330 nm and increasing intensity at 360 nm. Excitation was at 295 nm. Conditions are given in the legend of Figure 2.

The near-UV CD spectrum examines the chirality of the environment of the aromatic residues in a polypeptide. Per monomer, WT TR contains two tryptophans (residues 19 and 99), two tyrosines (residues 7 and 30), and one phenylalanine (residue 22); all but Trp-99 are also present in the $[2-66]_2$ fragment. Both $[2-66]_2$ and WT have near-UV CD spectra with fine structure, indicative of restrictive conformations about their aromatic residues (Figure 2B). The unfolded structures of WT and $[2-66]_2$, at 8 M urea, as monitored by near- and far-UV CD are similar (Figure 2), with featureless spectra that suggest random coil structures. The spectrum of folded WT TR shows more distinct fine structure from 270 to 295 nm than the spectrum of $[2-66]_2$, and the ellipticities of the folded WT spectrum are generally more positive bands than the unfolded spectrum. In contrast, the fine structure of the folded $[2-66]_2$ is generally of more negative ellipticity than that of the featureless unfolded spectrum, particularly the large negative band centered at 280 nm. This difference in the CD spectra of WT TR and $[2-66]_2$ may reflect somewhat different packing about the aromatic residues retained in $[2-66]_2$, relative to WT TR, but reflects predominantly the lack of Trp-99 in the $[2-66]_2$ fragment. The near-UV CD spectra of the mutant W99F shows similar fine structure to that of $[2-66]_2$ from 270 to 295 nm, with negative ellipticities relative to the unfolded spectrum (Gloss and Matthews, unpublished data).

The intrinsic fluorescence of the tryptophan residues is another spectral probe of tertiary structure, yielding information about the solvent accessibility and hydrophobicity of the tryptophan environments. Spectra of the intrinsic tryptophan fluorescence, as a function of urea, for $[2-66]_2$ and WT TR are shown in Figure 3. The WT spectra reflect contributions from Trp-19 and Trp-99, while the core spectra only reflect the behavior of Trp-19. The unfolded WT fluorescence spectrum is red shifted 22 nm relative to the folded spectrum. At 8 M urea, the intensity at 352 nm (the λ_{max} of the urea-denatured fluorescence emission) is ~9% greater than that at 330 nm in 0 M urea (the λ_{max} of the folded fluorescence emission; Figure 3B). The λ_{max} values of the unfolded and folded spectra of $[2-66]_2$ are also 352 and 330 nm, respectively. However, the intensity at 352 nm of unfolded $[2-66]_2$ is ~40% less than that at 330 nm in 0 M urea (Figure 3A). The same spectral behavior is seen for [NHis-7−66]$_2$ (data not shown).

The relative contributions and behavior in response to urea denaturation of Trp-19 and Trp-99 have been previously dissected by site-directed mutagenesis (Royer *et al*., 1993). The spectra of single tryptophan mutants (W99F and W19F) are red shifted 22 and 12 nm, respectively, upon unfolding; however, the fluorescence intensity of Trp-19 (W99F) decreases at high urea by ~30% while the intensity of Trp-99 (W19F) increases to a greater extent upon unfolding. Thus the $[2-66]_2$ and [NHis-7−66]$_2$ core fragments exhibit urea-dependent fluorescence spectra similar to that reported for W99F (Royer *et al*., 1993). These data demonstrate that the intrinsic tryptophan fluorescence response of Trp-19 in $[2-66]_2$ to unfolding is very similar to that observed for Trp-19 in full-length TR.

Urea Denaturation. The reversibility of the urea-induced unfolding of WT TR has been previously reported (Gittelman & Matthews, 1990). To test the reversibility of the urea denaturation of [NHis-7−66]$_2$ and $[2-66]_2$, lyophilized protein stocks were dissolved in either 0 or 8 M urea and diluted to varying final concentrations of urea from 0.5 to 7.5 M. The CD and fluorescence spectra were the same within error whether the samples had been prepared from folded (0 M) or unfolded (8 M) protein (data not shown). Thus, the urea denaturations of [NHis-7−66]$_2$ and $[2-66]_2$ are fully reversible processes that can be thermodynamically analyzed.

The equilibrium unfolding of WT TR is highly cooperative, as evidenced by the quality of the fits to a two-state model (eq 2) and the coincidence of the unfolding transitions measured by CD and FL (Gittelman & Matthews, 1990). The urea-induced unfolding transitions of $[2-66]_2$ and [NHis-7−66]$_2$ monitored by CD at 222 nm, fluorescence intensity at 330 nm (λ_{max} of the native proteins), and fluorescence intensity-averaged emission wavelength are coincident (Figure 4; [NHis-7−66]$_2$ data not shown) and are also well fitted by the simple, sigmoidal behavior expected for a two-state process. A more qualitative probe of the validity of the two-state model for the equilibrium urea-induced unfolding of TR variants comes from an examination of the isoemissive point observed in fluorescence spectra at various urea concentrations. WT TR exhibits an isoemissive point at 338 nm, indicative of the two-state nature of this transition (Figure 3B). The isoemissive points of the $[2-66]_2$ (Figure 3A) and [NHis-7−66]$_2$ core (data not shown) are apparent at about 370 nm.

The apparent stability of a dimeric protein is expected to be dependent upon the concentration of monomer, as has been previously reported for WT TR (Gittelman & Matthews, 1990). For the $[2-66]_2$ fragment, the urea-induced transitions are also dependent upon monomer concentration (Figure 4). C_M is the urea concentration at which the unfolding

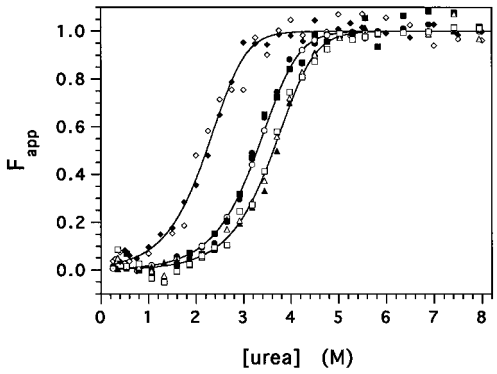

FIGURE 4: Urea dependence of the equilibrium unfolding of $[2-66]_2$ core *Trp* repressor monitored by CD and fluorescence at different monomer concentrations. For clarity of presentation, only a subset of the data sets used for the global analysis (Table 1) is presented. 3.5 μM monomer data from fluorescence intensity-averaged emission wavelength, IAEW (◆), and fluorescence intensity at 330 nm, the λ_{max} of emission for native protein, FL_{330} (◇). 7 μM monomer data from IAEW (●), FL_{330} (○), and CD_{222} (■). 21 μM monomer data from IAEW (▲), FL_{330} (△), and CD_{222} (□). The lines represent the global fits of the data to the model of eq 2. Conditions: 25 °C, 10 mM potassium phosphate, 0.1 mM EDTA, pH 7.6.

Table 1: Thermodynamic Parameters for the Urea-Induced Unfolding of *Trp* Repressor Variants[a]

TR variant	spectral method[b]	fitting method[c]	$\Delta G°(H_2O)$ (kcal mol^{-1})	−m (kcal mol^{-1} M^{-1})
WT TR[d]	CD	local	23 (2)	3.0 (0.3)
	IAEW	local	20.5 (0.5)	2.5 (0.1)
	FL 330 nm	local	23 (3)	3.0 (0.5)
		global	24 (2)	3.1 (0.3)
$[2-66]_2$ core[e]	CD	local	12.5 (0.9)	1.8 (0.2)
	IAEW	local	14.3 (0.2)	2.2 (0.4)
	FL 330 nm	local	13 (2)	2.1 (0.3)
		global	14.4 (0.3)	2.0 (0.1)
$[NHis\text{-}7-66]_2$[f]	CD	local	14 (2)	2.1 (0.5)
	IAEW	local	15.3 (0.8)	2.4 (0.4)
	FL 330 nm	local	13 (1)	2.0 (0.5)
		global	14.5 (0.5)	2.2 (0.1)

[a] Conditions: 10 mM potassium phosphate, 0.1 mM EDTA, pH 7.6 and 25 °C. The standard deviations are in parentheses. [b] CD, ellipticity at 222 nm; IAEW, fluorescence intensity-averaged emission wavelength (eq 1); FL 330 nm, fluorescence intensity at 330 nm, the λ_{max} of the folded proteins. [c] Local fits are the average value obtained for fits of individual data sets to a two-state model (eq 2). Global fits are the results of the simultaneous fitting of all the data sets collected by different spectral methods and at different [monomer] using Savuka 5.0. [d] The monomer concentration of the WT TR data was 15 μM monomer. Global fitting was performed on three individual data sets. [e] $[2-66]_2$ core monomer concentrations were 3.5, 7, 11.6, 15, and 21 μM for fluorescence and 7, 11.6, 15, and 21 μM for CD. Global fits were performed on 14 data sets. [f] $[NHis\text{-}7-66]_2$ data were collected at 15 and 21.5 μM monomer for both CD and FL; global fits were performed on seven data sets.

transition is half-completed, *i.e.*, $F_{app} = 0.5$. The C_M values of $[2-66]_2$ were 2.1, 3.3, and 3.7 M urea for monomer concentrations of 3.5, 7, and 21 μM monomer, respectively. A similar dependence was observed for $[NHis\text{-}7-66]_2$ (data not shown). The free energy of folding in the absence of denaturant, $\Delta G°(H_2O)$, and the sensitivity of the transition to denaturant, *m*, for $[2-66]_2$ and $[NHis\text{-}7-66]_2$, at 25 °C, are given in Table 1. The thermodynamic parameters for $[2-66]_2$ were measured at monomer concentrations covering a 7-fold range (3.5−21 μM). The average values of the local fits are given in Table 1; the individual fitted values varied randomly around the average, within the indicated standard deviation. The average of the local fits for multiple data sets for the CD and fluorescence data are the same, within error. Furthermore, the parameters determined by globally fitting all of the data sets, at different monomer concentrations and by different spectral techniques, are in excellent agreement with the local fits. The agreement provides further evidence for a two-state unfolding process. Similar agreement was found for fitting of the $[NHis\text{-}7-66]_2$ data (Table 1).

The thermodynamic parameters describing the unfolding of WT TR (Table 1) agree well with those published previously (Gittelman & Matthews, 1990). The $\Delta G°(H_2O)$ values and the *m* values of the two core variants are very similar to each other and are approximately 60% of those of WT. However, the magnitude of these thermodynamic parameters and the quality of the fits to a two-state model demonstrate that the core fragments constitute independently stable, cooperatively folded domains within the global structure of full-length TR.

Thermal Denaturation. The reversibility of the thermal denaturation of WT TR and $[2-66]_2$ was examined by three methods. First, after the thermal transition was recorded up to 90−95 °C, the sample was quickly cooled (less than 5 min) to 25 °C; >95% of the native ellipticity or absorbance was recovered. Second, some samples were melted, cooled quickly, and melted a second time. The magnitudes of the signal changes recorded for the two unfolding transitions were within 5% and 15% for $[2-66]_2$ and WT TR, respectively. Third, the absence of hysteresis was shown by recording the unfolding (heating) transition, followed immediately by the refolding (cooling) transition. When 7 μM samples of $[2-66]_2$ were heated to 80 °C (99% unfolded) and then cooled (a process requiring up to 30 min), the unfolding and refolding transitions were completely superimposable (data not shown). Heating to 95 °C and then cooling resulted in transitions whose optical properties were within 6% of each other over the entire temperature range examined. WT TR displayed somewhat lower reversibility. Folding transitions recorded after unfolding to 95 °C (95% unfolded for the 2.6 μM samples employed) had spectral values up to 15% less than those recorded prior to unfolding; the average difference was 8%.

The thermal transitions of WT TR and $[2-66]_2$ were monitored by absorbance and circular dichroism spectroscopy at various monomer concentrations. The decrease in Abs at 292 nm with increasing temperature was employed to monitor the exposure to solvent of Trp-19 and Trp-99 (in WT TR) and, thereby, the disruption of tertiary structure; the loss of CD signal at 222 nm monitored the disruption of α-helical structure. The CD_{222} and Abs_{292} signals depended linearly on temperature from 15 to at least 40 °C and from 25 to at least 65 °C for $[2-66]_2$ and WT, respectively, indicative of a native baseline region. The native baseline region extended to higher temperatures for samples with higher monomer concentrations. Above these temperatures, the spectral signals varied in a sigmoidal manner with temperature. For both $[2-66]_2$ and WT TR, the midpoints of the Abs transitions were only slightly dependent on monomer concentration while the CD midpoints were clearly dependent on monomer concentration; the CD midpoints were consistently higher than those observed for Abs transitions at a given monomer concentration. For $[2-66]_2$, above 80 °C, the CD and Abs signals displayed a linear

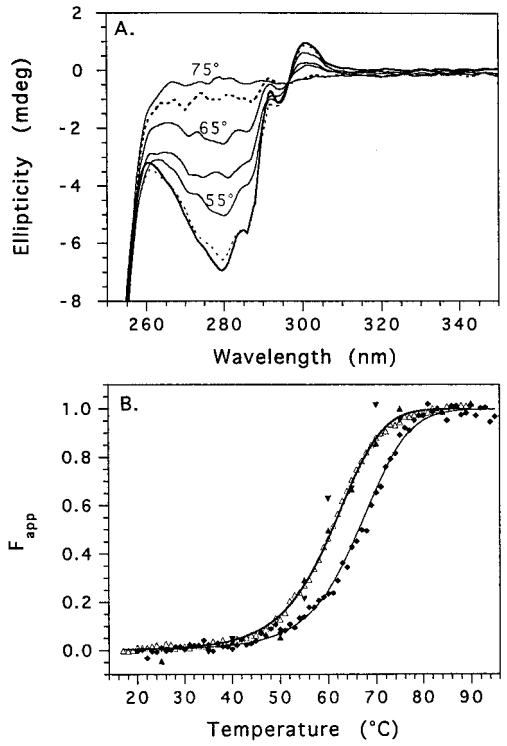

FIGURE 5: Thermal denaturation of $[2-66]_2$ core monitored by near-UV and far-UV CD and Abs_{292}. (A) Temperature dependence of the near-UV CD spectra of $[2-66]_2$. Spectra were taken at 25 (bold line), 50 (lower dashed line), 55, 60, 65, 70 (upper dashed line), and 75 °C. (B) Apparent fraction of thermally unfolded $[2-66]_2$. The solid lines represent the global fits of the Abs_{292} (△) or CD_{222} (◆) data to a three-state model with a dimeric intermediate (eq 4). Also shown are the near-UV CD data at 280 (▲) and 300 nm (▼). Conditions: 15 μM monomer, 10 mM potassium phosphate, 0.1 mM EDTA, pH 7.6.

variation with temperature, yielding an unfolded baseline region. For WT TR, an unfolded baseline could be observed only by CD for the lowest monomer concentration, 2.6 μM, but the Abs_{292} data showed a baseline region above about 85 °C for all monomer concentrations. The slopes of the baselines observed for both proteins were of similar magnitude; the slopes of CD_{222} folded and unfolded baselines were positive, while the Abs_{292} slopes were negative.

Individual thermal transitions for WT TR and $[2-66]_2$, by either CD 222 nm or Abs 292 nm were well described by local fits to a two-state model (eq 2) using the thermodynamic formalism in eq 12. However, the F_{app} curves (eq 3) generated from two-state fits of Abs and CD, for a given monomer concentration, were not coincident (data not shown). Therefore, thermal denaturation must proceed through at least three states, with tertiary structure, as monitored by tryptophan absorbance at 292 nm, melting prior to the secondary structure, monitored by CD at 222 nm. Tertiary structure was also probed by monitoring absorbance changes at 280 nm. The relative contributions of Trp, Tyr, and Phe to the absorption spectrum differ between 280 and 292 nm, providing a test of the concerted nature of the disruption of tertiary structure at high temperature. The thermal unfolding of $[2-66]_2$ monitored by Abs at 280 nm has the same midpoint as that observed for the Abs 292 nm transition (data not shown). Another probe for the melting of tertiary structure is the near-UV CD spectrum of the aromatic residues; spectra at several temperatures are shown in Figure 5A. The thermal dependence of the maxima and minima of the folded spectra, at 300 and 280 nm, respectively, display midpoints similar to that observed by Abs_{292} (Figure 5B). The coincidence of the thermal transitions monitored by these three probes of tertiary structure (Abs_{292}, Abs_{280}, and near-UV CD) demonstrates that the disruption of tertiary structure at high temperature is a global phenomenon.

$[2-66]_2$ and WT TR data were fitted to two, alternative three-state models, assuming either a dimeric or monomeric intermediate (eqs 4 and 8, respectively). The quality of the fits to the latter model were substantially poorer for the following reasons. Either the reduced χ^2 values were at least 5-fold higher than those obtained for fits with a dimeric intermediate or the thermodynamic parameters converged to values with no physical meaning. For example, a negative value of ΔC_p and/or a value of $T°$ less than 260 K for the $N_2 \rightarrow 2I$ transition were obtained.

The thermodynamic parameters derived from the global fits of WT TR and $[2-66]_2$ data to a three-state model with a dimeric intermediate (eq 4) are given in Table 2. Representative Abs_{292} and CD_{222} F_{app} curves for $[2-66]_2$ at several concentrations, derived from fits to a three-state, dimeric intermediate model are shown in Figure 6. Similar fits for WT TR are shown in Figure 7. The calculated relative populations of the N_2, I_2, and U species, as a function of temperature and monomer concentration, are shown in Figure 8. Because of the relative stabilities of N_2, I_2, and U, the dimeric intermediate is more highly populated for the WT TR transitions than for those of $[2-66]_2$. The temperatures for maximal population of the dimeric intermediate are 84 °C for WT TR at 11 μM monomer and 65 °C for $[2-66]_2$ at 15 μM monomer.

The Z values in eq 7 have a theoretical range of 0−1 for Abs and CD and relate the optical properties of the intermediate to those of the unfolded species. The Z values for CD and Abs for WT TR and $[2-66]_2$ fit to the limits of this parameter (Table 2). The Z_{CD222} values of ∼0 show that the dimeric intermediates retain all the helical structure of the native dimers. Conversely, the Z_{Abs292} values of ∼1 demonstrate that, optically, the tryptophan environments are as solvent accessible in the I_2 intermediates as they are in the unfolded monomers. Given the limiting values obtained for the Z values, it is clear that Abs and CD are reporting two separate unfolding transitions. The F_{app} curves of Abs_{292} are representative of a lower thermal transition for the tertiary structure. For $[2-66]_2$, the near-UV CD signals at 300 and 280 nm (the major bands in the spectra) display a temperature dependence similar to that of the Abs_{292} signal (Figure 5B), demonstrating that the tertiary packing and solvent accessibility of all of the aromatic residues in $[2-66]_2$ display the same thermal dependence. The F_{app} curves of CD_{222} reflect secondary structure melting, occurring almost entirely in the higher thermal transition. The lack of coincidence of the Abs_{292} and CD_{222} transitions is most clearly demonstrated by the F_{app} curves shown in Figures 5B and 7 for a subset of the monomer concentrations included in the global analyses of $[2-66]_2$ and WT TR data.

The $T°$ parameters, indicated in Table 2, are reference temperatures at which the $\Delta G°$ value is zero for the specified transition. In a monomeric system, $T°$ would also be a $T_{1/2}$ or T_M value, the temperature at which the transition was half-completed. This identity is not valid for a dimeric system, where the unfolding transition is coupled to a dissociation

Table 2: Thermodynamic Parameters for the Thermal-Induced Unfoldng of Wild-Type *Trp* Repressor and the $[2-66]_2$ Fragment[a]

	$T°$ (°C)[b]	$\Delta S°$ (T°) [cal/(mol K)]	$\Delta H°$ ($T°$) (kcal/mol)	ΔC_p [cal/(mol K)]	*Z* values	
					Abs 292 nm	CD 222 nm
WT TR[c]					0.9 (0.1)	0.0 (0.01)
$N_2 \rightleftarrows I_2$	79.9 (0.5)	197 (8)	70 (4)	100 (100)		
$I_2 \rightleftarrows 2U$	120 (6)	318 (60)	125 (25)	1700 (700)		
$[2-66]_2$ core[d]					1.0 (0.1)	0.0 (0.1)
$N_2 \rightleftarrows I_2$	69 (2)	102 (10)	35 (3)	200 (200)		
$I_2 \rightleftarrows 2U$	105 (5)	233 (36)	88 (18)	1300 (400)		

[a] Parameters were obtained from global fits of scans of MRE_{222} and extinction coefficient at 292 nm as a function of temperature. Conditions: 10 mM potassium phosphate, 0.1 mM EDTA, pH 7.6. [b] $T°$ is a reference temperature at which $\Delta G° = 0$ and is independent of monomer concentration. [c] WT TR data were collected at 2.6, 5.3, and 11 μM monomer. The temperature range was 25−95 °C. The global fits of 13 individual data sets converged to a reduced χ^2 value of 2.01. [d] $[2-66]_2$ concentrations were 3.5, 7, and 15 μM monomer for the CD data and 7, 15, 21, and 30 μM monomer for the absorbance data. The temperature range was 15−95 °C. The global fits for 15 individual scans converged to a reduced χ^2 value of 1.95.

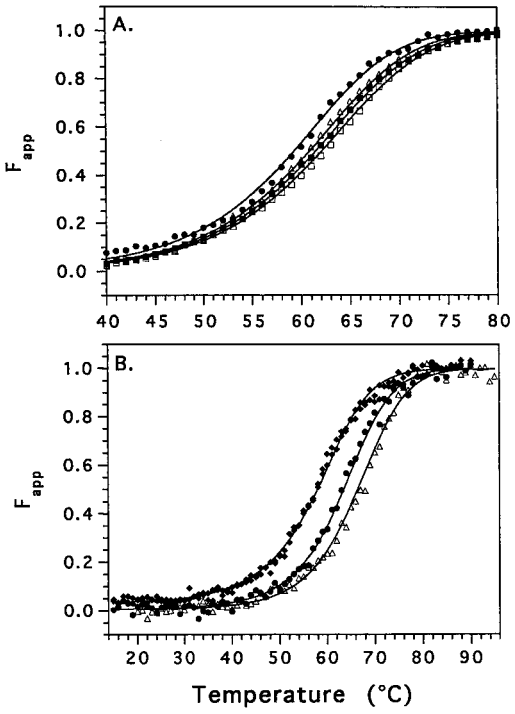

FIGURE 6: Apparent fraction unfolded of $[2-66]_2$ core *Trp* repressor as a function of temperature. The solid lines are the global fits of the data to a three-state model with a dimeric intermediate (eq 4). (A) Thermal denaturation monitored by absorbance at 292 nm. Monomer concentrations: 7 μM (●), 15 μM (△), 21 μM (■), and 30 μM (□). To enhance the differentiation between the various data sets, only the transition region is shown although data were collected from 15 to 95 °C. (B) Thermal denaturation monitored by circular dichroism at 222 nm. Monomer concentrations: 3.5 μM (◆), 7 μM (●), and 15 μM (△). Conditions: 10 mM potassium phosphate, 0.1 mM EDTA, pH 7.6.

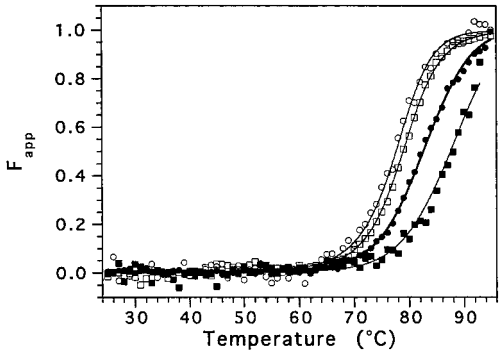

FIGURE 7: Apparent fraction thermally unfolded WT TR monitored by Abs_{292} and CD. The solid lines represent the global fits of the Abs_{292} (open symbols) or CD_{222} (closed symbols) data to a three-state model with a dimeric intermediate (eq 4). WT TR monomer concentrations are 2.6 μM (●, ○) and 11 μM (■, □). Conditions are described in the legend of Figure 6.

step (Thompson *et al.*, 1993). Although the $T°$ value for a dimer is independent of monomer concentration, the $T_{1/2}$ value (the temperature at which unfolded monomer constitutes 50% for the total protein) should increase with higher monomer concentrations, as does C_M for urea denaturation (*e.g.*, Figure 4). Inspection of the fraction unfolded, X_u, as a function of monomer concentration (Figure 8) demonstrates the expected dependence for $T_{1/2}$ for both $[2-66]_2$ and WT TR in the concentration ranges used in these experiments. Comparison of the concentration-independent parameter, $T°$, shows that WT TR is more stable to thermal denaturation than is $[2-66]_2$, similar to the results for urea denaturation. For both the N_2/I_2 and $I_2/2U$ transitions, the $T°$ values of WT TR are 15−17 °C higher than those observed for $[2-66]_2$. Similarly, the values of $\Delta S°$ and $\Delta H°$ for both transitions are significantly larger for WT than are the corresponding values for $[2-66]_2$. For both TR variants, the thermodynamic parameters $\Delta S°$ and $\Delta H°$ are much larger for the $I_2/2U$ transition than for the N_2/I_2 transition, ∼1.6-fold and ∼2.4 fold higher for WT TR and $[2-66]_2$, respectively. The larger values reflect, in part, the larger values of ΔC_p and $T°$ for the $I_2/2U$ transition.

DISCUSSION

Structural Integrity of Core Fragments. Several lines of evidence demonstrate that $[2-66]_2$ is a dimeric protein with extensive secondary structure and well-defined tertiary structure. Size-exclusion FPLC confirmed the ability of the fragment to assemble into a dimeric species. Far-UV CD spectra showed that, on a per residue basis, $[2-66]_2$ has a slightly higher α-helical content than WT TR (Figure 2A). Near-UV CD spectra demonstrated that the aromatic residues of [2−66] are in asymmetric environments, indicative of a well-packed hydrophobic core (Figure 2B). The intrinsic fluorescence of Trp-19 in the core exhibits a urea dependence similar to that observed for this same fluorophore in full-length TR dimers (Figure 3 and Results), demonstrating that the solvent inaccessibility and hydrophobicity of the environment of Trp-19 in WT TR is preserved in $[2-66]_2$.

Urea-Induced Unfolding. The isolated core domains, which contain ∼60% of the primary sequence of full-length TR, are quite stable to urea denaturation, with $\Delta G°(H_2O)$ values that are ∼60% that of WT TR (Table 1). The

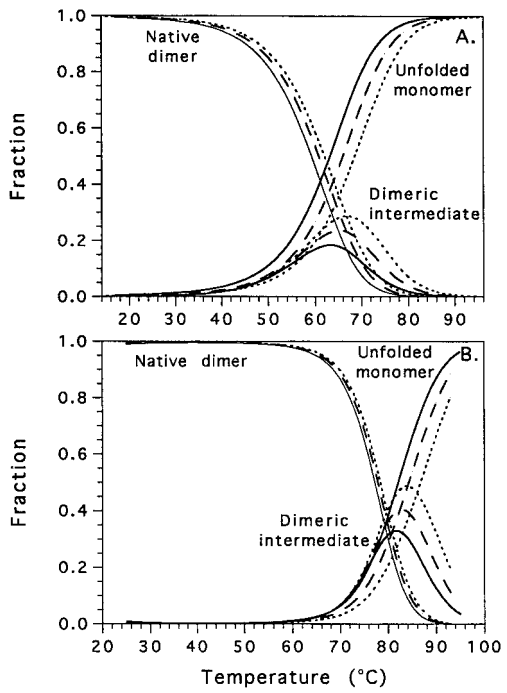

FIGURE 8: Calculated fractions of the native dimer, dimeric intermediate, and unfolded monomer as a function of temperature and monomer concentration. (A) $[2-66]_2$ core at 7 μM (solid), 15 μM (dashed), and 30 μM (dotted) monomer. (B) WT *Trp* repressor at 2.6 μM (solid), 5.3 μM (dashed), and 11 μM (dotted) monomer.

observations that the replacement of the unstructured N-terminal residues with a tag of six histidines has no measurable effect on stability of the $[2-66]_2$ fragment and the removal of the seven N-terminal residues (by chymotryptic cleavage) has little effect on the stability of the full-length protein (Carey, 1989) demonstrate that only residues 8−66 are required to develop the observed 14 kcal mol^{-1} of stability.

The unfolding transitions of the core fragments are highly cooperative as judged by the quality of the two-state fits to individual transitions, the coincidence of transition curves monitored by different spectral probes, and the quality of global fits to data sets at various monomer concentrations (Figure 4). However, the *m* values of the unfolding transitions of these cores are substantially less than that of WT (Table 1). The previously predicted (Schellman, 1978) and observed (Myers *et al*., 1995) correlation between *m* value and the amount of hydrophobic and polar surface area buried during folding leads to the expectation that decreasing the sequence length of a protein may reduce the *m* value for a two-state process. Inspection of the X-ray structure of full-length TR (Zhang *et al*., 1987) provides support for this hypothesis. The core fragments lack the D, E, and F helices which, in WT TR, abut the ABC helical core domain, interacting predominantly via hydrophobic residues with the core domain. The solvent exposure of the reduced number of buried hydrophobic residues contained in the core fragments would be expected to lead to a decrease in the *m* value for the core fragments, relative to WT, as observed.

Taken together, the spectroscopic and thermodynamic data clearly demonstrate that the isolated core fragments of *Trp* repressor constitute a well-structured, independent folding domain. Although the details of the structure of $[2-66]_2$ may be somewhat different than those of WT, this fragment does not fold to a molten globule-like state (Kuwajima, 1989). While the $[2-66]_2$ fragment is less stable than full-length TR, it exhibits a stability that is comparable to other dimers of similar size. For example, the *E. coli* Arc repressor with only 53 residues per monomer exhibits a stability of 10.6 kcal mol^{-1} (Milla & Sauer, 1994).

Thermal Denaturation and Detection of a Dimeric Intermediate. In agreement with a previous report (Bae *et al*., 1988), WT TR possesses a very high thermal stability. At a concentration of 2.6 μM monomer, WT TR melts with a $T_{1/2}$ of 82 °C. The $T_{1/2}$ value increases with monomer concentration, as expected for a dimeric protein (Figure 8). The $T_{1/2}$ values of $[2-66]_2$ also depend on monomer concentration; however, the $T_{1/2}$ is reduced to 63 °C at 3.5 μM monomer.

The striking finding of the thermal studies is the presence of a well-populated equilibrium intermediate in the thermal denaturation of both WT TR and $[2-66]_2$. Global analysis of the transition curves obtained at a series of protein concentrations indicates that the intermediate is dimeric. Transient dimeric intermediates have been postulated for the kinetic folding mechanism of WT TR (Mann & Matthews, 1993; Mann *et al*., 1995), and a transient dimeric intermediate has been detected in the refolding of $[2-66]_2$ (Gloss and Matthews, manuscript in preparation). The spectral properties of the WT TR and $[2-66]_2$ dimeric intermediates appear to be very similar. In both cases, the aromatic side chains lose their close packing contacts and become exposed to solvent while the helical structure remains intact. These structural features are characteristic of a molten globule, a commonly observed transient folding intermediate, or an equilibrium intermediate populated at extreme pH or temperature (Kuwajima, 1989; Matthews, 1993; Griko *et al*., 1994; Ptitsyn, 1996).

Given that these dimeric intermediates are stable at micromolar monomer concentrations and temperatures above 60 °C, the ≥9 kcal mol^{-1} of binding free energy suggests a rather specific mode of association. Specific recognition in the absence of native-like, solvent-exclusive packing of the aromatic side chains might appear to be contradictory; however, such structures have been observed in a molten globule form of α-lactalbumin (Peng *et al*., 1995; Wu *et al*., 1996) and transiently in the pulsed hydrogen-exchange NMR studies of *E. coli* dihydrofolate reductase (Jones & Matthews, 1995) and other proteins (Baldwin, 1993). One feature of these *Trp* repressor dimeric intermediates not usually associated with molten globules is a substantial ΔC_p for their thermal unfolding. However, recent studies have yielded similar ΔC_p values for the molten globule intermediates of α-lactalbumin, cytochrome *c*, and apomyoglobin (Haynie & Freire, 1993; Griko *et al*., 1994).

The extent of exposure of buried surface area to solvent upon unfolding is reflected in the ΔC_p values and is generally dominated by the positive contribution of solvating nonpolar surfaces (Spolar *et al*., 1989; Privalov & Makhatadze, 1992; Thompson *et al*., 1993; Myers *et al*., 1995). ΔC_p values from differential scanning calorimetry have been reported for a few dimeric proteins: MetJ, another dimeric *E. coli* DNA-binding protein (Johnson *et al*., 1992); λ Cro repressor (Griko *et al*., 1992); the four-helix bundle, ROP (Steif *et al*., 1993); and the histone H2A−H2B core dimer (Karantza *et al*., 1995). The range of ΔC_p values for $N_2/2U$ transitions determined from these studies and the present report is 1.5−2.5 kcal mol^{-1} K^{-1}, the same range observed for monomeric

proteins (Privalov & Gill, 1988). Thus, the intersubunit hydrophobic interfaces in oligomeric proteins must, on the average, have a distribution of nonpolar and polar surfaces similar to that found in the interior of monomeric proteins.

A puzzling aspect of the spectral and thermodynamic properties of the thermal I_2 species is the apparently complete exposure of the aromatic side chains to solvent with a very small value of ΔC_p for the N_2/I_2 transition (Table 2). A possible explanation for this behavior may be the counterbalancing effects of ΔC_p for the exposure of nonpolar and polar groups to water. The ΔC_p value contains positive contributions from the exposure to solvent of hydrophobic surface area and a smaller, negative contribution due to the exposure of hydrophilic surfaces (Murphy & Freire, 1992). Also relevant may be the observation that the exposure of aromatic residues to an aqueous solvent is less positive than exposure of aliphatic residues, possibly because of the favorable interactions of the π electrons with hydrogen-bonding moieties (Privalov & Makhatadze, 1992). Therefore, the small magnitude of $\Delta C_{pN_2/I_2}$ may result from a near-canceling of the heat capacity changes accompanying the exposure of these two types of surface. The approximately 1.7-fold larger contribution, per Å^2, of nonpolar surface area (Murphy & Freire, 1992) suggests that approximately 4-fold more polar surface area is exposed in the dimeric intermediate to account for the low ΔC_p of the N_2/I_2 transition.

Alternative denaturation methods to urea or guanidine hydrochloride are often useful means to populate equilibrium intermediates that resemble transient kinetic intermediates. One example is the use of low pH, which populates the molten globule of apomyoglobin (Jennings & Wright, 1993) and other monomeric proteins (Kuwajima, 1989), as well as a monomeric form of Trp repressor (Eftink *et al*., 1994). For WT TR and $[2-66]_2$, thermal denaturation breaks down the cooperativity of the unfolding reactions of WT TR and $[2-66]_2$, allowing the detection of a well-populated equilibrium intermediate. The thermal denaturations of most wild-type dimeric proteins reported in the literature are predominantly two-state processes, including the four dimers cited above and the dimeric fragment of GCN4 containing the leucine zipper motif, bZIP (Thompson *et al*., 1993). Thus, the three-state thermal unfolding of $[2-66]_2$ and WT TR is a rather unique behavior, which permits a glimpse of stable, partially folded structures, not detectable by equilibrium urea denaturation.

Implications for Protein Folding. Simplification of the complicated tertiary structure of proteins by characterization of fragments corresponding to putative subdomains has been performed successfully for many years. Initial work utilized protease sensitivity to define domain boundaries [*e.g.*, Kato and Anfinsen (1969)]. The advent of molecular biology has permitted a refinement of this approach through expression of truncated gene products or the insertion of residues susceptible to chemical cleavage, namely, methionine (Kippen & Fersht, 1995) and cysteine (Gegg *et al*., 1996). Fragmentation has also been employed to better characterize the mechanisms by which proteins fold (see introduction).

This report demonstrates that the $[2-66]_2$ fragment of TR is a promising minimized dimerization domain that will serve as a fruitful model system for investigation of the properties of dimer formation. The $[2-66]_2$ sequence contains all of the information necessary to direct the stable formation of dimeric structure and yet folds via a kinetic response that is far simpler than that of WT TR (Gloss and Matthews, manuscript in preparation). Therefore, this fragment model system is a useful vehicle through which to explore the equilibrium and kinetic features of a dimerization reaction. The detection of a dimeric intermediate in the thermal unfolding of $[2-66]_2$ suggests that a core dimer can be partially unfolded and still retain a dimeric structure. This implies that a further truncation of the 2−66 fragment may define a minimal sequence required for dimerization. Such studies are in progress.

ACKNOWLEDGMENT

We thank Dr. William DeW. Horrocks, Jr., and his laboratory for the use of their Spex−Fluorolog fluorometer. We appreciate Dr. Osman Bilsel's efforts in the development of the in-house global fitting routines and his assistance in their utilization. We are indebted to Dr. Jill A. Zitzewitz for many fruitful discussions and critical review of the manuscript. We also thank Drs. Jannette Carey and Darryl Abriola for discussions about the 8−71 *Trp* repressor core fragment.

REFERENCES

Alexandrescu, A. T., Abeygunawardana, C., & Shortle, D. (1994) *Biochemistry 33*, 1063−1072.
Aurora, R., Srinivasan, R., & Rose, G. D. (1994) *Science 264*, 1126−1130.
Bae, S.-J., Chou, W.-Y., Matthews, K., & Sturtevant, J. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6731−6732.
Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 84−91.
Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859−1877.
Carey, J. (1989) *J. Biol. Chem. 264*, 1941−1945.
Clackson, T., & Wells, J. A. (1995) *Science 267*, 383−386.
Dill, K. A., Bromberg, S., Yue, Fiefg, K. M., Yee, D. P., Thomas, P. D., & Chan, H. S. (1995) *Protein Sci. 4*, 561−602.
Eftink, M. R., Helton, K. J., Beavers, A., & Ramsay, G. D. (1994) *Biochemistry 33*, 10220−10228.
Gegg, C. V., Bowers, K. E., & Matthews, C. R. (1996) in *Techniques in Protein Chemistry, 7th*, pp 439−448. Academic Press, San Diego.
Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
Gittelman, M. S., & Matthews, C. R. (1990) *Biochemistry 29*, 7011−7020.
Griko, Y. V., Rogov, V. V., & Privalov, P. L. (1992) *Biochemistry 31*, 12701−12705.
Griko, Y. V., Freire, E., & Privalov, P. L. (1994) *Biochemistry 33*, 1889−1899.
Harbury, P. B., Zhang, T., Kim, P. S., & Alber, T. (1993) *Science 262*, 1401−1407.
Haynie, D. T., & Freire, E. (1993) *Proteins 16*, 115−140.
Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117−237.
Jennings, P. A., & Wright, P. E. (1993) *Science 262*, 892−896.
Joachimiak, A., Kelley, R. L., Gunsalus, R. P., Yanofsky, C., & Sigler, P. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 668−672.
Johnson, C. M., Cooper, A., & Stockley, P. G. (1992) *Biochemistry 31*, 9717−9724.
Jones, B. E., & Matthews, C. R. (1995) *Protein Sci. 4*, 167−177.
Karantza, V., Baxevanis, A. D., Freire, E., & Moudrianakis, E. N. (1995) *Biochemistry 34*, 5988−5996.
Kato, I., & Anfinsen, C. B. (1969) *J. Biol. Chem. 244*, 1004−1007.
Kippen, A. D., & Fersht, A. R. (1995) *Biochemistry 34*, 1464−1468.
Kippen, A. D., Sancho, J., & Fersht, A. R. (1994) *Biochemistry 33*, 3778−3786.
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
Kuwajima, K. (1989) *Proteins 6*, 87−103.

Makhatadze, G. I., & Privalov, P. L. (1995) *Adv. Protein Chem. 47*, 307−425.
Mann, C. J., & Matthews, C. R. (1993) *Biochemistry 32*, 5282−5290.
Mann, C. J., Shao, X., & Matthews, C. R. (1995) *Biochemistry 34*, 14573−14580.
Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.
Milla, M. E., & Sauer, R. T. (1994) *Biochemistry 33*, 1125−1133.
Milla, M. E., & Sauer, R. T. (1995) *Biochemistry 34*, 3344−3351.
Murphy, K. P., & Freire, E. (1992) *Adv. Protein Chem. 43*, 313−361.
Myers, J. K., Pace, C. N., & Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
Pace, C. N. (1986) *Methods Enaymol. 131*, 266−280.
Pecorari, F., Guilbert, C., Minard, P., Desmadril, M., & Yon, J. M. (1996) *Biochemistry 35*, 3465−3476.
Peng, Z. Y., Wu, L. C., Schulman, B. A., & Kim, P. S. (1995) *Philos. Trans. R. Soc. London, B, Biol. Sci. 348*, 43−47.
Porath, J., Carlsson, J., Olsson, I., & Belfrage, G. (1975) *Nature 258*, 598−599.
Privalov, P. L., & Gill, S. J. (1988) *Adv. Protein Chem. 39*, 191−214.
Privalov, P. L., & Makhatadze, G. I. (1992) *J. Mol. Biol. 224*, 715−723.
Ptitsyn, O. (1996) *Nat. Struct. Biol. 3*, 488−490.
Pulah, J. L., & Yanofsky, C. (1986) *Nucleic Acids Res. 14*, 7851−7860.
Rosenberg, A. H., Lade, B. N., Chui, D.-S., Lin, S.-W., Dunn, J. J., & Studier, F. W. (1987) *Gene 56*, 125−135.
Royer, C. A., Mann, C. J., & Matthews, C. R. (1993) *Protein Sci. 2*, 1844−1852.
Ruiz-Sanz, J., de Prat Gay, G., Otzen, D. E., & Fersht, A. R. (1995) *Biochemistry 34*, 1695−1701.
Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
Shortle, D., Wang, Y., Gillespie, J. R., & Wrabl, J. O. (1996) *Protein Sci. 5*, 991−1000.
Spolar, R. S., Ha, J.-H., & Record, T. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8382−8385.
Steif, C., Weber, P., Hinz, H.-J., Flossdorf, J., & Kokkinidis, M. (1993) *Biochemistry 32*, 3867−3876.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.
Tasayco, M. L., & Carey, J. (1992) *Science 255*, 594−597.
Thompson, K. S., Vinson, C. R., & Freire, E. (1993) *Biochemistry 32*, 5491−5496.
Tsuji, T., Chrunyk, B. A., Chen, X., & Matthews, C. R. (1993) *Biochemistry 32*, 5566−5575.
Viera, J., & Messing, J. (1987) *Methods Enzymol. 153*, 3−11.
Wells, J. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 1−6.
Wu, L. C., Laub, P. B., Elöve, G. A., Carey, J., & Roder, H. (1993) *Biochemistry 32*, 10271−10276.
Wu, L., Grandori, R., & Carey, J. (1994) *Protein Sci. 3*, 369−371.
Wu, L. C., Schulman, B. A., Peng, Z.-y., & Kim, P. S. (1996) *Biochemistry 35*, 859−863.
Zhang, R.-G. Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., & Sigler, P. B. (1987) *Nature 327*, 591−597.
Zhao, D., Arrowsmith, C. H., Jia, X., & Jardetzky, O. (1993) *J. Mol. Biol. 229*, 735−746.
Zheng, Z., Czaplicki, J., & Jardetzky, O. (1995) *J. Mol. Biol. 34*, 5215−5223.
Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., & Matthews, C. R. (1995) *Biochemistry 34*, 12812−12819.

BI970056E