Folding of Chymotrypsin Inhibitor 2. 1. Evidence for a Two-State Transition

Sophie E. Jackson[‡] and Alan R. Fersht*

MRC Unit for Protein Function and Design, Cambridge IRC for Protein Engineering, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.

Received April 25, 1991; Revised Manuscript Received July 18, 1991

ABSTRACT: The reversible folding and unfolding of barley chymotrypsin inhibitor 2 (CI2) appears to be a rare example in which both equilibria and kinetics are described by a two-state model. Equilibrium denaturation by guanidinium chloride and heat is completely reversible, and the data can be fitted to a simple two-state model involving only native and denatured forms. The free energy of folding in the absence of denaturant, ΔG_{H_2O}, at pH 6.3, is calculated to be 7.03 ± 0.16 and 7.18 ± 0.43 kcal mol^{-1} for guanidinium chloride and thermal denaturation, respectively. Scanning microcalorimetry shows that the ratio of the van't Hoff enthalpy of denaturation to the calorimetric enthalpy of denaturation does not deviate from unity, the value observed for a two-state transition, over the pH range 2.2–3.5. The heat capacity change for denaturation is found to be 0.789 kcal mol^{-1} K^{-1}. The rate of unfolding of CI2 is first order and increases exponentially with increasing guanidinium chloride concentration. Refolding, however, is complex and involves at least three well-resolved phases. The three phases result from heterogeneity of the unfolded form due to proline isomerization. The fast phase, 77% of the amplitude, corresponds to the refolding of the fraction of the protein that has all its prolines in a native trans conformation. The rate of this major phase decreases exponentially with increasing guanidinium chloride concentration. The unfolding and refolding kinetics can also be fitted to a two-state model. Importantly, ΔG_{H_2O} and m, the constant of proportionality of the free energy of folding with respect to guanidinium chloride concentration, calculated from the kinetic experiments, 7.24 ± 0.22 kcal mol^{-1} and 1.86 ± 0.05 kcal mol^{-1} M^{-1}, respectively, agree, within experimental error, with the values measured from the equilibrium experiments. This is perhaps the strongest evidence that the unfolding of CI2 follows a simple two-state transition.

Small monomeric proteins lacking disulfides bonds and cis proline peptide bonds provide the simplest starting point for investigating the rules that govern protein folding. In proteins that contain disulfide bridges or proline peptide bonds in a cis conformation, the rate-determining step of the protein-folding process is frequently that of disulfide bond rearrangement (Creighton, 1988) or proline isomerization (Kiefhaber et al., 1990a,b,c), respectively. The folding of larger multidomain proteins is still more complex (Jaenicke, 1987), and frequently there is more than one transition between native and denatured states as domains fold independently. This complicates both analysis and interpretation of results. Many small monomeric proteins have simple, single, unfolding transitions between native and denatured states. However, there is now increasing evidence for complex behavior within such simple systems. Studies on staphylococcal nuclease, for example, have shown that there is residual structure in the denatured state (Shortle, 1989), and scanning microcalorimetry studies on T4 lysozyme have shown that there is deviation from a two-state model at some values of pH (Kitamura & Sturtevant, 1989). Both events complicate the analysis of denaturation curves. Here, we present the characterization of the denaturation and renaturation of chymotrypsin inhibitor 2 and show it is a good model system for both equilibrium and kinetic studies.

Chymotrypsin inhibitor 2, found in the albumin fraction of seeds from the *Hiproly* strain of barley, is a member of the potato inhibitor I family of serine protease inhibitors (Jonassen, 1980). It is a small monomeric protein of 83 residues, with M_r = 9250 (Svendsen et al., 1980). It is a single-domain protein containing both α-helix and β-sheet structural elements. The tertiary structure of CI2 consists of a four-stranded mixed parallel and antiparallel β-sheet against which an α-helix packs to form the hydrophobic core. Between parallel strands 2 and 3, on the opposite side of the β-sheet to the α-helix, is a wide loop in extended conformation that contains the reactive site bond. Unlike many other serine protease inhibitors, CI2 contains no disulfide bonds. Four of the five prolines are well defined in the crystal structure and are in the trans conformation. The fifth proline is in a region not defined in either the crystal structure or the NMR solution structure. CI2 has one tryptophan residue that is buried in the hydrophobic core and that is responsible for a large change in the fluorescence spectrum of the protein on denaturation.

The crystal structure of CI2 has been solved to high resolution, 2.0 Å (McPhalen & James, 1987), which is a prerequisite for the rational design of site-directed mutants. The solution structure has been determined by NMR spectroscopy (Kjaer et al., 1987; Kjaer & Poulsen, 1987; Clore et al., 1987a,b). The gene encoding for the protein has been cloned and can be expressed at high levels in *Escherichia coli*. (Longstaff et al., 1990). This will allow future studies to probe the nature of protein stability and folding by using protein engineering techniques.

We now present a study on the equilibrium unfolding transition of CI2. The nature of the transition between unfolded and folded protein is characterized by a number of different techniques in order to detect stable folding intermediates. Guanidinium chloride and heat are used to denature the protein, and fluorescence is used as a probe of the state of the protein. In addition, differential scanning microcalorimetry experiments are performed. The kinetics of unfolding and refolding can be used to detect intermediates which are not stable enough to be detected under equilibrium conditions. The kinetics of unfolding and refolding are investigated using [GdnHCl] jump and pH-jump experiments.

* Author to whom correspondence should be addressed.

‡ S.E.J. was supported by a SERC studentship.

0006-2960/91/0430-10428$02.50/0

Experimental Procedures

Materials

Chemicals. The buffer used in denaturation experiments was 2-(*N*-morpholino)ethanesulfonic acid (MES) purchased from Sigma. Guanidinium chloride was sequanal grade purchased from Pierce Chemicals. Water used in equilibrium and kinetic experiments was purified to 15 MΩ resistance by an Elgastat system. Subtilisin BPN′ used in the activity measurements was expressed and isolated from *E. coli* harboring the plasmid pPT30 (Thomas et al., 1985). Yeast extract and tryptone were purchased from Lab M, Bury, U.K. Ammonium sulfate was enzyme grade from BDH. All other chemicals were purchased from Sigma.

Recombinant CI2. The wild-type CI2 gene has been cloned into the high level expression vector pINIIIompA3, in which the structural gene is fused to the N-terminal signal sequence of *ompA* and is under control of the *llp* promoter and *lac* operator sequences (Longstaff et al., 1990). The *E. coli* strain TG2 containing pINIIIompA3 was grown in 2×TY medium (Maniatis et al., 1982) containing ampicillin (50 μg mL^{-1}) and IPTG (0.5 mM) for induction of expression. Cultures were grown overnight at 37 °C with vigorous shaking to an absorbance of 6–8 at 600 nm. Cells were harvested and concentrated by centrifugation (6500*g*), and the cell paste was resuspended in 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA, then sonicated three times for 0.5 min on maximum power at 4 °C with a Heats Systems ultrasonic sonicator. Cell debris was removed by centrifugation (25000*g*) before the pH of the supernatant was lowered to 4.4 by addition of acetic acid. The precipitate was removed by centrifugation (25000*g*). Ammonium sulfate was added to 40% (w/v) to precipitate unwanted proteins, which were removed by centrifugation (25000*g*). CI2 was precipitated from this supernatant by the addition of ammonium sulfate to 65% (w/v) and collected by centrifugation (25000*g*). Protein pellets were resuspended in 10 mM sodium acetate, pH 4.4, and then separated on a Sephadex G-75 superfine column. Fractions were collected, and those containing inhibitory activity were pooled before cation-exchange chromatography on a Pharmacia FPLC Mono S column. CI2 was bound to the column in 10 mM acetate, pH 4.4, and eluted around 200 mM salt with a linear NaCl gradient. Purified CI2 was dialyzed extensively against water to remove salt, flash frozen, and stored at −70 °C. The purified protein was homogeneous as judged by $NaDodSO_4$–polyacrylamide gel electrophoresis.

Methods

Spectroscopy. The intrinsic fluorescence of CI2 increases on denaturation, which allows unfolding and refolding to be monitored by fluorescence spectroscopy. The fluorescence yield of unfolded and folded forms of the protein is constant with respect to guanidinium chloride concentration (up to 6.5 M). This gives flat baselines, and the fraction of unfolded protein at a given guanidinium chloride concentration is simply the ratio of the observed fluorescence change to the maximal fluorescence change. Fluorescence spectroscopy was used in guanidinium chloride and thermal denaturation studies as well as in the kinetic studies. The maximal change in fluorescence upon denaturation is obtained with an excitation wavelength of 280 nm and an emission wavelength of 356 nm. A Perkin-Elmer LS5B Luminescence spectrometer was used for the equilibrium studies, with slit widths of 2 mm.

Chemical Denaturation Experiments. Guanidinium chloride solutions were prepared gravimetrically in volumetric flasks. The guanidinium chloride solutions were divided into 800-μL aliquots with an SMI (American Hospital Supply) positive displacement pipetter with repetitive pipetting attachment and stored at −20 °C. For each data point in the denaturation experiment, 100 μL of CI2 stock solution in 450 mM MES was diluted into 800 μL of the appropriate denaturant concentration, with a SMI positive displacement pipetter. The final CI2 concentration was approximately 2.5 μM. The protein/denaturant solutions were preequilibrated at 25 °C for approximately 1 h. Spectroscopic measurements were carried out in a thermostated cuvette holder at 25 °C, the temperature being monitored throughout the experiment by a thermocouple immersed in a neighboring cuvette in the cell holder. The unfolding was found to be completely reversible. The activity of a sample of CI2 denatured in 6.5 M GdnHCl was found to be completely restored after dilution 100-fold into 10 mM MES, pH 6.3. Activity was measured by the inhibition by native CI2 of the serine protease subtilisin BPN′ (Longstaff et al., 1990).

Thermal Denaturation Experiments. Thermally induced unfolding was also monitored by fluorescence spectroscopy using 2.5 μM CI2 in 50 mM MES, pH 6.3. The pK_a of this buffer is temperature dependent, and the pH varies from 6.3 to 5.6 over the temperature range used. It was found that the stability of CI2 does not change significantly over this pH range. Jacketed cuvettes were heated by a circulating water bath. The temperature was monitored by a thermocouple immersed in the cuvette under observation above the light beam. Reversibility was shown by the regain in initial fluorescence after cooling.

Calorimetry. Measurements were carried out in 10 mM glycine-HCl solution at pH 2.2, 2.5, 2.8, 3.2, and 3.5. Under these conditions, the protein is soluble and no aggregation is observed on heating. Solutions were carefully dialyzed, before measurements, against solvent with the solvent being replaced several times to ensure complete equilibration. Protein concentrations in solution were determined spectrophotometrically with correction for light scattering (Winder & Gent, 1971). The concentration of protein varied from 0.5 to 1.5 mg mL^{-1}. The extinction coefficient of CI2 was determined in a separate experiment using the method of Gill and von Hippel (1989). The extinction coefficient of native CI2 is based on a comparison of the spectra for native and denatured (6 M GdnHCl) protein measured at identical protein concentrations. The concentration of denatured protein can be calculated from the known amino acid composition and extinction coefficients of tyrosine, tryptophan, and cysteine obtained from model compound studies (Edelhoch, 1967). The extinction coefficient for the native protein can then be calculated by using this concentration. The experiment was performed in triplicate, and an extinction coefficient of 7418 M^{-1} cm^{-1} at 282 nm was calculated.

Calorimetric measurements were performed with a DASM-4A capillary scanning microcalorimeter equipped with gold cells of 1.0-mL volume (Bureau of Biological Instruments of the Academy of Sciences of the USSR). Measurements were performed under an excess pressure of 2 bar. The heating rate was 1 K min^{-1}. A baseline scan with dialysis buffer filled cells was performed prior to measurements. The calorimeter was calibrated by applying an electrical current at 20 μW for 5 min. Repeated measurements with the same protein sample demonstrated reversibility of ≥90%. The areas under the transition curves were determined with a planimeter.

Kinetics. Reactions were followed with a Perkin-Elmer MPF-44B fluorescence spectrophotometer equipped with a rapid mixing head. The mixing device contained a T-jet

mixing chamber followed by a 30-ms delay loop ensuring complete mixing of the solutions before observation. A Hellma flow through compact cell, 10 mm × 3 mm × 3 mm, was used. The solutions were driven through the mixing chamber manually from two Hamilton syringes resulting in a mixing ratio of 1:10. Data were acquired with a Tandon Target microcomputer, a DT2801 data translation board, and the Bio-kine data acquisition software package by Bio-logic and analyzed with the program ENZFITTER (Elsevier Biosoft, Cambridge) by R. J. Leatherbarrow.

Unfolding was initiated by diluting the aqueous protein solution (~45 μM) into guanidinium chloride solutions of different concentrations with final concentrations between 4.0 and 7.35 M. Unfolding data were fitted to a single-exponential rate equation.

Refolding was initiated, after preequilibration of the protein in 6.5 M GdnHCl, by diluting it into low concentrations of GdnHCl. Final guanidinium chloride concentrations were between 0.59 and 4.0 M. Points between 0 and 0.59 M GdnHCl were obtained by a pH-jump experiment. The protein was initially acid denatured by lowering the pH of the solution to 1.7 by addition of 5 M HCl. At this pH, it was found that the protein is completely denatured as determined by fluorescence spectroscopy. The subsequent addition of GdnHCl makes no further difference to the fluorescence spectrum. The protein was refolded by rapid mixing (1:1) with a strongly buffered solution at high pH. The final pH of the solution was pH 6.3. An Applied Photophysics stopped-flow spectrophotometer model SF 17MV was used to monitor the fast reaction. Refolding is a triphasic process with a very fast major phase and two slow minor phases. The split time base facility of the stopped-flow was used to monitor the three phases simultaneously. Typically split time bases of 0–500 ms and 1–500 s were used to monitor the fast and slow phases, respectively. The slow phases were fitted to a double exponential with floating end point by using the SF 17MV software package by P. J. King, Applied Photophysics. The fast phase was then fitted to a double exponential with floating end point with the second rate fixed to the value obtained for the faster of the two slow phases.

A Radiometer PHM64 research pH meter was used for pH measurements. The pH-jump experiment was performed in the absence and in the presence of low concentrations of denaturant. All kinetic experiments were performed at 25 °C.

Data Analysis

Guanidinium Chloride Denaturation. This analysis is for a two-state model of denaturation where only the native and the denatured states are populated. The equilibrium constant for unfolding, K_U, and free energy of unfolding, ΔG_U, in the presence of a denaturant may be calculated from

$$K_U = (F_N - F)/(F - F_U) = \exp(-\Delta G_U/RT) \quad (1)$$

where F is the observed fluorescence and F_N and F_U are the values of the fluorescence of the native and denatured forms of the protein, respectively, and R is the gas constant, 8.314 J mol^{-1} K^{-1}. The values of F_N and F_U are independent of [GdnHCl] up to 6.5 M so eq 1 can be applied directly. It has been found experimentally that the free energy of unfolding of proteins in the presence of guanidinium chloride is linearly related to the concentration of denaturant (Tanford, 1968; Pace, 1986):

$$\Delta G_U = \Delta G_{H_2O} - m[\text{denaturant}] \quad (2)$$

where ΔG_U is the free energy of unfolding at a particular denaturant concentration, ΔG_{H_2O} is the free energy of unfolding in water, and m is a constant that is proportional to the increase in degree of exposure of the protein on denaturation. The entire data set from the fluorescence-monitored guanidinium denaturation was fitted with the nonlinear regression analysis program ENZFITTER by using eq 3, which is derived from eqs 1 and 2.

$$F = F_N - (F_N - F_U)\frac{\exp(m[\text{GdnHCl}] - \Delta G_{H_2O})/RT}{1 + \exp(m[\text{GdnHCl}] - \Delta G_{H_2O})/RT} \quad (3)$$

Thermal Denaturation. K_U as a function of temperature was calculated in a similar way to the guanidinium chloride mediated denaturation, assuming a two-state transition. Outside the transition region, the fluorescence of the native and denatured forms varies linearly with temperature, and so the values of F_N and F_U are estimated at each temperature by extrapolation [see Pace (1986) for a review of this method]. The extrapolation lowers the accuracy of the data compared with that of the guanidinium chloride mediated denaturation.

Calculation of ΔG from Thermal Denaturation. The enthalpy, $\Delta H_U(T)$, and the entropy, $\Delta S_U(T)$, of unfolding at temperature T are given by classical thermodynamics by

$$\Delta H_U(T) = -R[\partial \ln K_U/\partial(1/T)] = RT^2(\partial \ln K_U/\partial T) \quad (4)$$

$$\Delta S_U(T_m) = \Delta H_U(T_m)/T_m \quad (5)$$

where T_m is the midpoint of thermal denaturation. van't Hoff plots ($\ln K_U$ vs $1/T$) of thermal denaturation are approximately linear through the T_m region, allowing an estimation of the enthalpy and entropy of unfolding at T_m. Plots over a wider range of temperatures are expected to be curved because $\Delta H_U(T)$ and $\Delta S_U(T)$ vary with temperature according to (Privalov, 1979)

$$\Delta C_p = [\partial \Delta H_U(T)/\partial T]_p = T[\partial \Delta S_U(T)/\partial T]_p \quad (6)$$

where ΔC_p is the change in heat capacity at constant pressure that accompanies the unfolding of a protein. Thus the value of $\Delta G_U(T)$ at a given temperature cannot be simply calculated from $\Delta H_U(T)$ and $\Delta S_U(T)$ without allowing for their variation because of ΔC_p. There is evidence that ΔC_p is independent of temperature between 20 and 80 °C (Privalov, 1979). Consequently, $\Delta H_U(T)$, the enthalpy of unfolding at any temperature T, $\Delta S_U(T)$, the entropy of unfolding at T, and $\Delta G_U(T)$, the free energy of unfolding at T, can be calculated at temperatures other than T_m from (Baldwin, 1986)

$$\Delta H_U(T) = \Delta H_U(T_m) + \Delta C_p(T - T_m) \quad (7)$$

$$\Delta S_U(T) = \Delta S_U(T_m) + \Delta C_p \ln (T/T_m) \quad (8)$$

$$\Delta G_U(T) = \Delta H_U(T) - T\Delta S_U(T) \quad (9)$$

RESULTS

Guanidinium Chloride Denaturation. Figure 1A shows the change of fluorescence of CI2 upon titration with guanidinium chloride. There is a single sharp transition between the initial and final states. Data from the transition region (3.2–4.6 M) were transformed to give a linear plot of ΔG_U against [GdnHCl] shown in Figure 1B. The entire data set can be fitted directly to eq 3, from which the value of ΔG_{H_2O} and m can be calculated. For wild-type CI2, ΔG_{H_2O} = 7.03 ± 0.16 kcal mol^{-1} and m = 1.79 ± 0.04 kcal mol^{-1} M^{-1}. The midpoint of unfolding (K_U = 1) is at 3.92 ± 0.18 M GdnHCl. Figure 1A,B shows the best fits to a two-state model of protein folding.

In order to ascertain that the transition monitored fluorimetrically is due to the global denaturation of the protein, and not caused by some local unfolding event in the proximity of the tryptophan, a series of 1-D ^{1}H NMR experiments were

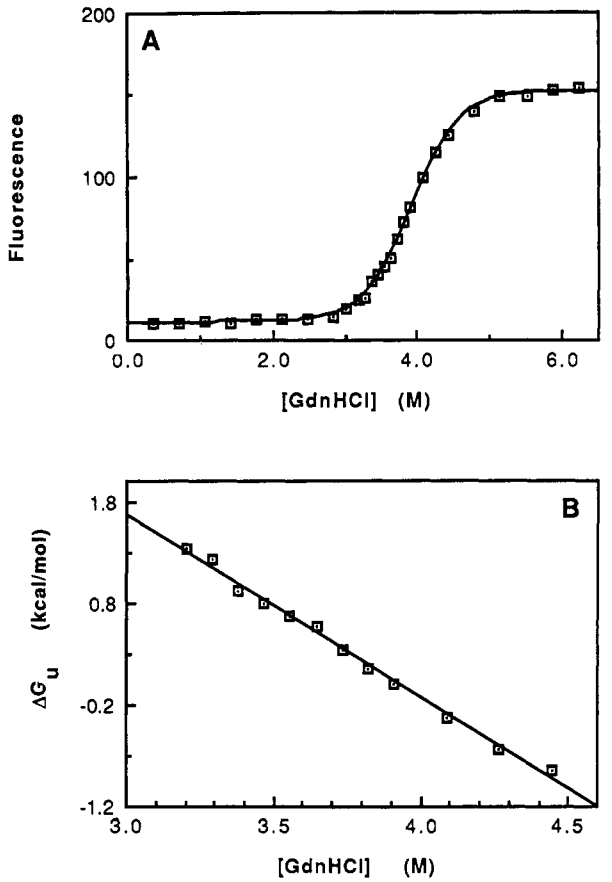

FIGURE 1: GdnHCl-induced denaturation of CI2 at 25 °C, in 50 mM MES, pH 6.3. (A) Intrinsic fluorescence monitored at 356 nm, with an excitation wavelength of 280 nm, as a function of [GdnHCl]. The solid curve shows the best fit of the data to eq 3. (B) Transformation of the fluorescence data to give a linear plot of change in free energy versus [GdnHCl] around the midpoint of the unfolding transition. The solid line indicates the best fit of the data to eq 3. Linear extrapolation of this plot to 0 M GdnHCl results in the free energy change of denaturation in the absence of denaturant.

conducted at various GdnHCl concentrations, at 25 °C, using a Bruker AM500. From such a series, it was possible to follow peaks that are characteristic of the native structure of the protein as a function of guanidinium chloride concentration. All the peaks that were followed were observed to disappear/appear concomitantly in a single transition with a midpoint of unfolding at approximately 3.8 M GdnHCl (data not shown). This is consistent with the transition observed by fluorescence spectroscopy and with the transition being due to a global denaturation.

Thermal Denaturation. The midpoint for thermal denaturation in 50 mM MES, pH 6.3, in the absence of denaturant, was found to be over 80 °C, which is too high to allow an accurate direct measurement of ΔG_{H_2O} at 25 °C. The T_m was lowered by addition of GdnHCl to final concentrations of 1.2, 1.6, 2.0, and 2.4 M. Figure 2A shows the thermal denaturation curves at the four GdnHCl concentrations. The thermal denaturation curves show a single sharp transition between native and denatured forms typical of a two-state process. Figure 2B shows the corresponding van't Hoff plots from which ΔH_{T_m}, ΔS_{T_m}, and T_m can be calculated according to eqs 4 and 5. These values, in turn, can be used to calculate ΔG_U at 25 °C at the four different GdnHCl concentrations according to eqs 7, 8, and 9. In the calculations, a value of $\Delta C_p = 0.789$ kcal deg^{-1} mol^{-1} was used, based on the result from the calorimetric studies. Figure 2C shows the corresponding linear plot of the free energy of unfolding against GdnHCl concentration, from

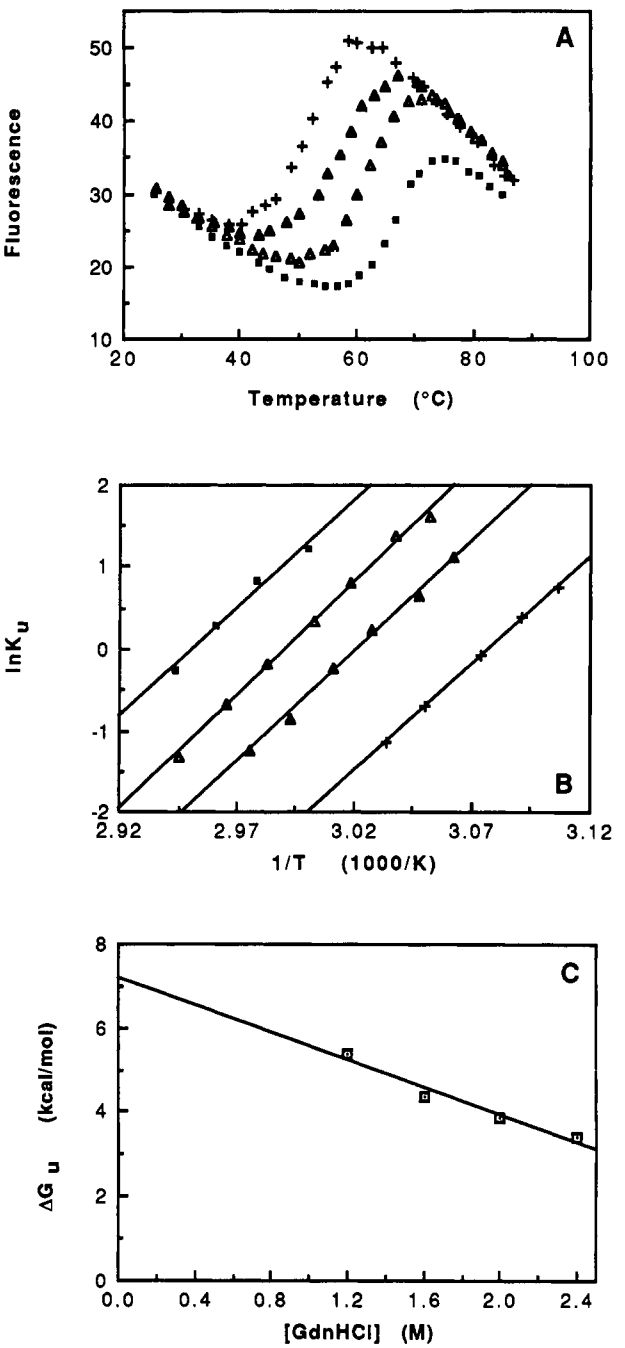

FIGURE 2: Thermal denaturation of CI2 in the presence of low concentrations of GdnHCl, in 50 mM MES, pH 6.3. (A) Intrinsic fluorescence of CI2 as a function of temperature, with (■) 1.2 M GdnHCl, (△) 1.6 M GdnHCl, (▲) 2.0 M GdnHCl, and (+) 2.4 M GdnHCl. (B) van't Hoff plots of the thermal transition curves around the midpoint of the transition, with (■) 1.2 M GdnHCl, (△) 1.6 M GdnHCl, (▲) 2.0 M GdnHCl, and (+) 2.4 M GdnHCl. The solid line indicates the best fit of the data to eq 4. (C) Plot of the free energy of denaturation at 25 °C, calculated from the van't Hoff plots, and using eqs 7–9, versus [GdnHCl]. Linear extrapolation to 0 M GdnHCl results in the free energy of denaturation in the absence of denaturant.

which the value of ΔG_{H_2O} at 25 °C can be calculated by extrapolation back to 0 M (see eq 2). This gives a value of 7.18 ± 0.43 kcal mol^{-1} for ΔG_{H_2O} and 1.62 ± 0.23 kcal mol^{-1} M^{-1} for m. These values are consistent, within experimental error, with the values obtained by guanidinium chloride denaturation.

Calorimetry. From the transition curves T_m, the transition temperature, $\Delta H_{cal}(T_m)$, the molar calorimetric enthalpy of denaturation at T_m, and $\Delta H_{vH}(T_m)$, the molar van't Hoff

Table I: Scanning Microcalorimetry Experiments

pH	T_m (K^{-1})[a]	$\Delta H_{cal}(T_m)$ (kcal mol^{-1})[a]	$\Delta H_{vH}(T_m)$ (kcal mol^{-1})[b]	$\Delta H_{vH}(T_m)/\Delta H_{cal}(T_m)$[c]	$\Delta S_{cal}(T_m)$ (kcal mol^{-1} K^{-1})[d]	ΔG_{cal}(25 °C) (kcal mol^{-1})[e]
2.2	314.6	43.5	42.3	0.97	0.138	1.944
2.5	320.2	49.2	45.5	0.93	0.154	2.787
2.8	328.2	52.8	51.1	0.97	0.161	3.728
3.2	337.0	61.0	60.8	1.00	0.181	5.201
3.5	344.0	67.5	68.8	1.02	0.196	6.475

[a] T_m, the transition temperature, and $\Delta H_{cal}(T_m)$, the molar calorimetric enthalpy of denaturation at T_m, are measured directly from the transition curves (change in specific heat with temperature). [b] $\Delta H_{vH}(T_m)$, the molar van't Hoff enthalpy of denaturation at T_m, is calculated by using eq 10. [c] $\Delta H_{vH}(T_m)/\Delta H_{cal}(T_m) = 1.0$ only for two-state transitions. [d] $\Delta S_{cal}(T_m)$ is calculated from eq 5. [e] ΔC_p can be calculated by using eq 6; ΔG_{cal} (25 °C) can then be calculated by using this value and eqs 7–9.

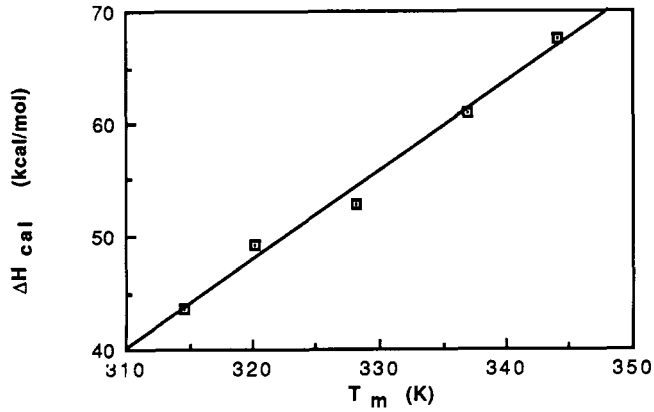

FIGURE 3: Scanning microcalorimetry of CI2. Plot of the calorimetric transition enthalpy $\Delta H_{cal}(T_m)$ versus the transition temperature T_m.

enthalpy of denaturation at T_m can be calculated as a function of pH. Calorimetric enthalpies for unfolding, $\Delta H_{cal}(T_m)$, are calculated numerically from the area under the transition curve. Assuming a two-state transition, the van't Hoff enthalpies at T_m, $\Delta H_{vH}(T_m)$, can be calculated from the calorimetric transition curves by using (Privalov, 1986; Shrake & Ross, 1990; Kitamura & Sturtevant, 1989)

$$\Delta H_{vH}(T_m) = 4RT_m^2\frac{C_p(T_m)}{\Delta H_{cal}(T_m)} \qquad (10)$$

where $C_p(T_m)$ is the molar excess heat capacity at T_m. Transition temperatures, T_m, are taken as the peak of the calorimetric transition curve. Thermodynamic parameters determined from the data are summarized in Table I. The ratio of the van't Hoff enthalpy of denaturation to the calorimetric enthalpy of denaturation is shown, and the ratio is unity over the pH range studied. This shows that the two-state assumption for the unfolding transition applies under these conditions.

The variation of the calorimetric transition enthalpies with temperature is shown in Figure 3. As expected over this temperature range, ΔC_p is invariant and $\Delta H_U(T_m)$ is linearly dependent on T_m. By use of eq 6, the calorimetric value for the change in heat capacity on denaturation, ΔC_p, can be calculated from the gradient of such a plot. For CI2 the value of ΔC_p is 0.789 kcal mol^{-1} K^{-1}.

This corresponds to a specific heat capacity change of 0.0834 cal g^{-1} K^{-1}. This is slightly less than the average value found for globular proteins of about 0.1 kcal g^{-1} K^{-1} (Privalov, 1979) and may result from the long unstructured tail of some 20 amino acids at the N-terminus of CI2, which should not contribute to a change in heat capacity on denaturation.

The unfolded state of a protein has a much larger heat capacity than the native state (Privalov et al., 1989). This is mainly due to the increased hydration of hydrophobic residues that are exposed in the unfolded state but buried in the native state. The change in heat capacity on denaturation therefore reflects the change in the degree of exposure of hydrophobic residues (Privalov & Gill, 1989). The large positive value of ΔC_p for CI2 suggests that the unfolding reaction starts from a compact native structure, in which many of the hydrophobic residues are buried, and proceeds to an unfolded state in which such hydrophobic residues are highly solvated.

Both $\Delta H_U(T_m)$ and $\Delta S_U(T_m)$, measured from scanning microcalorimetry experiments (see Table I) are large and positive as has been observed for the high-temperature melting of a number of proteins (Privalov, 1979). At 25 °C, ΔH_U is 30.5 kcal mol^{-1} and ΔS_U is 102.3 cal mol^{-1} K^{-1}, which is comparable to the values found of 35 kcal mol^{-1} and 116 cal mol^{-1} K^{-1} for T4 lysozyme at 28 °C (Chen & Schellman, 1989). Unfortunately, it was not possible to measure the thermal denaturation of CI2 by scanning microcalorimetry at pH 6.3 because of its high T_m. This means it is not possible to compare the value for ΔG_{H_2O} calculated from calorimetric data to those values obtained by other methods.

Unfolding Kinetics. The rate constant for unfolding, k_u, is generally found to increase with increasing final [GdnHCl] according to

$$\ln k_u = \ln k_u^{H_2O} + m_{k_u}[\text{GdnHCl}] \qquad (11)$$

where $k_u^{H_2O}$ is the rate constant for unfolding in water and m_{k_u} is a constant. The unfolding of CI2 is monophasic, and the data can be fitted to a single-exponential rate equation. The observed first-order rate constant is found to increase exponentially with increasing [GdnHCl] over the range 4.5–7.35 M, in accordance with eq 11 (see Figure 4).

Refolding Kinetics. The refolding of CI2, by either [GdnHCl] jump or by pH-jump experiments, is a triphasic process with the three phases well-resolved at 25 °C and at low concentrations of denaturant. The multiphasic nature of the refolding reaction results from a heterogeneous population in the denatured state due to proline isomerization. The fast phase, 77% of the amplitude measured by stopped-flow spectroscopy, corresponds to the folding of the fraction of unfolded protein that has all its prolines in the trans conformation in solution. The following discussion applies only to the fast phase; characterization of the two slower phases and evidence that these correspond to rate-limiting proline isomerization reactions are discussed in the following paper (Jackson & Fersht, 1991).

If the process is a reversible transition between just two states, native, N, and denatured, U,

$$U \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N$$

eqs 2 and 11 imply that the rate constant for folding, k_f, must follow the rate law

$$\ln k_f = \ln k_f^{H_2O} - m_{k_f}[\text{GdnHCl}] \qquad (12)$$

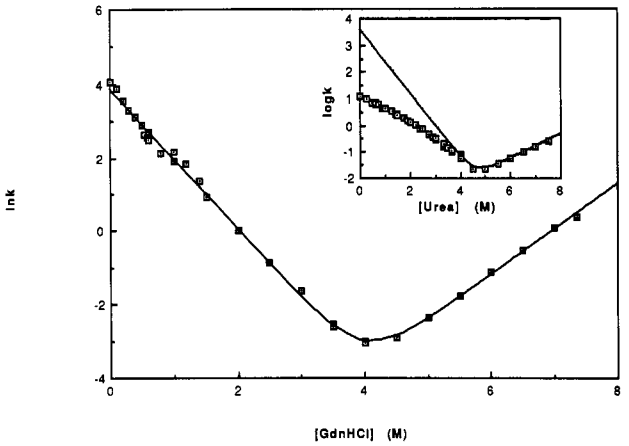

FIGURE 4: [GdnHCl] dependence of natural logarithm of the rate constants for denaturation and renaturation of CI2. Rate constants measured at 25 °C, in 50 mM MES pH 6.3. Points between 0 and 0.59 M GdnHCl were obtained by pH-jump experiments. Points between 0.59 and 7.35 M GdnHCl were obtained from [GdnHCl] jump experiments. Each point is obtained from the average of at least three separate experiments. The solid curve is the best fit of the data to a two-state model using eq 13. The insert shows the [urea] dependence of the rate constants for denaturation and renaturation for wild-type barnase. The solid curve is that calculated for a two-state system on the basis of the unfolding and equilibrium data (Matouschek et al., 1990).

where $m = m_{k_f} + m_{k_u}$ and $K_U = k_f/k_u$. These relationships are perhaps the most important criteria that show that a system follows the two-state model.

The rate constant for the fast phase, the rate of refolding, decreases exponentially with [GdnHCl] over the range 0–3.5 M (Figure 4). The complete kinetics of folding and unfolding can be fitted to eq 13, which is derived from eqs 11 and 12,

$$\ln k = \ln [k_f^{H_2O} \exp(-m_{k_f}[\text{GdnHCl}]) + k_u^{H_2O} \exp(m_{k_u}[\text{GdnHCl}])] \quad (13)$$

and based on a two-state transition, where k is the rate of unfolding or refolding at a particular GdnHCl concentration. The kinetic data for unfolding and refolding of CI2 fit to this two-state model. The calculated fit of the data to eq 13 is shown by the solid curve in Figure 4. The values of $k_f^{H_2O}$, $k_u^{H_2O}$, m_{k_f}, and m_{k_u}, calculated from the kinetic data according to eq 13, are 47.8 s^{-1}, 1.81×10^{-4} s^{-1}, -1.90 M^{-1}, and 1.24 M^{-1}, respectively. From these results, the equilibrium values of ΔG_{H_2O} and m can be calculated, taking into account the equilibria due to proline isomerization

$$U_c \underset{23\%}{\overset{K_{iso}}{\rightleftharpoons}} U_t \underset{77\%}{\overset{k_f^{H_2O}}{\underset{k_u^{H_2O}}{\rightleftharpoons}}} N$$

where U_t is the denatured protein with all its prolines in a trans conformation, U_c is the denatured protein with a proline in a cis conformation, and N is the native state, which has all its prolines in a trans conformation. Defining the equilibrium constant for isomerization, K_{iso}, as $[U_c]/[U_t]$, and the apparent equilibrium constant for unfolding calculated from the kinetic experiments, K, as $k_f^{H_2O}/k_u^{H_2O}$, it follows that the equilibrium constant for unfolding, K_U, is $K/(1 + K_{iso})$. From the relative amplitudes of the slow and fast phases, 23% and 77%, respectively, $K_{iso} = 0.299$ and $K_U = 2.03 \times 10^5$ M. This gives a value for ΔG_{H_2O} of 7.24 ± 0.22 kcal mol^{-1}. Similarly m can be calculated from the kinetic data by using the relationship $m = m_{k_f} + m_{k_u}$. This gives an m value of 1.86 ± 0.05 kcal mol^{-1} M^{-1} and a midpoint of unfolding of 3.89 M GdnHCl. Within experimental error, these values agree with the values of 7.03 ± 0.16 kcal mol^{-1} and 1.79 ± 0.04 kcal mol^{-1} M^{-1} for ΔG_{H_2O} and m and a midpoint of unfolding at 3.92 ± 0.18 M GdnHCl, calculated from equilibrium experiments.

In addition to showing that the kinetics of unfolding and refolding of CI2 can be fitted a two-state transition, we have shown that $m = m_{k_f} + m_{k_u}$, and $K_U = k_f/k_u$. This is very strong evidence that CI2 behaves kinetically, as well as thermodynamically, as a two-state system of folding.

Discussion

Criteria for a Two-State Transition. For the equilibrium denaturation of a protein to be described as two state, (i) the unfolding data must fit to a single transition curve as described by eqs 1–3, (ii) this transition must be independent of the probe used to determine the state of the protein, i.e., fluorescence, absorbance, circular dichroism, and NMR, etc, and (iii) the ratio of the van't Hoff enthalpy of denaturation to the calorimetric enthalpy of denaturation must be unity.

For the kinetics of unfolding and refolding to be described by a two-state model, (i) both unfolding and refolding must be monophasic processes (allowing for the possible heterogeneity of the unfolded form due to proline isomerization), (ii) if the logarithm of the equilibrium constant for unfolding and that for the rate constant for unfolding is linearly dependent on the concentration of denaturant, then the logarithm of the rate constant for refolding should also be linearly dependent on denaturant concentration, i.e., the data should fit to a two-state transition as described by eq 13, and (iii), most importantly, the values for ΔG_{H_2O} and m calculated from kinetic experiments must be the same as the values calculated from the equilibrium experiments, i.e., $m = m_{k_f} + m_{k_u}$, and $K_U = k_f/k_u$. This is perhaps the strongest evidence for a two-state transition.

Kinetic experiments are more sensitive than equilibrium experiments for observing intermediate states. Intermediate, which may not be stable under the conditions used in equilibrium experiments, may be metastable in the native conditions used in refolding experiments and, therefore, be detectable by kinetics. As a consequence, although the equilibrium properties of many proteins show two-state behavior, the kinetic properties of much fewer proteins also show two-state behavior. CI2 is a rare example of a protein in which both equilibria and kinetics are described by a two-state transition. The evidence for this, from both equilibrium and kinetic experiments, is discussed below.

Evidence from Equilibrium Experiments. In all cases, the equilibrium denaturation of CI2 appears to be a two-state process where only the native and the denatured states are significantly populated. Both the GdnHCl-induced denaturation and the thermal denaturation curves show single sharp transitions between native and denatured forms that can be fitted to a two-state model. In addition, scanning microcalorimetry, which does not assume a mechanistical model for the transition, also shows that thermal denaturation is a two-state process. The ratio of the van't Hoff enthalpy of denaturation, ΔH_{vH}, to the calorimetric enthalpy of denturation, ΔH_{cal}, measured from such experiments is unity over the pH range 2.2–3.5. The calorimetric enthalpy of denaturation is the same as the van't Hoff enthalpy of denaturation only for two-state systems (Privalov, 1979, 1986; Shrake & Ross, 1990; Kitamura & Sturtevant, 1989). The average ratio of $\Delta H_{vH}/\Delta H_{cal}$ for CI2 is 0.98 ± 0.03, which is within the range observed for the two-state denaturation of several globular proteins of 1.00 ± 0.05 (Privalov, 1979).

In addition, the transition monitored by fluorescence spectroscopy is the same as that measured by NMR spec-

troscopy. All the NMR peaks that were followed as a function of [GdnHCl] were observed to disappear/appear concomitantly in a single transition which followed that observed by fluorescence spectroscopy. This also shows that the transition monitored by fluorescence spectroscopy corresponds to a global denaturation of the protein and not a local unfolding event.

Evidence from Kinetic Experiments. Unfolding is a first-order reaction whose rate increases exponentially with increasing GdnHCl concentration. Refolding, however, is more complex and involves at least three well-resolved phases. The three phases result from heterogeneity in the unfolded form due to proline isomerizations [see following paper (Jackson & Fersht, 1991)]. The fast phase corresponds to the refolding of the fraction of denatured protein that has all its prolines in a native trans conformation. It represents, therefore, a rate-limiting step involving folding of the protein and not proline isomerization. This phase is also first order, and the rate decreases exponentially with increasing GdnHCl concentration. This results in linear plots of both ln k_f and ln k_u against [GdnHCl]. The data for unfolding and refolding can be fitted to a scheme, described by eq 13, in which only two states are populated. This is a marked difference from the results found for the folding of other small monomeric proteins, such as barnase, where there is significant nonlinearity in the plot of ln k_f against urea concentration (see the insert of Figure 4) (Matouschek et al., 1990) and also parvalbumin (Kuwajima et al., 1988) and α-lactalbumin and lysozyme (Ikeguchi et al., 1986).

It has not yet been established where on the folding pathway the change in fluorescence that is used to monitor the folding state of the protein occurs. If the change occurs before the rate-limiting step, it is possible that the only way of detecting deviation from a two-state model may be calculating the equilibrium constant from the kinetics and comparing it to the equilibrium data, rather than by curvature in the plots of rate constants. Because of the inaccuracies inherent in kinetic measurements, this approach leaves intermediates with stabilities of less than 10% of the folded form undetected.

For barnase, in addition to the curvature in the plot of ln k_f against denaturant concentration, there is significant deviation between the values of $k_f^{H_2O}$ based on the two-state model and calculated from the equilibrium and unfolding data and that measured experimentally. The measured value is three orders of magnitude lower than the predicted value. Both the curvature in the ln k_f plot and the deviation between the measured and calculated value for $k_f^{H_2O}$ are the result of the accumulation of an intermediate on the folding pathway at low concentrations of denaturant where the intermediate is stable. In contrast, for CI2 not only is there no curvature in the plot of ln k_f against [GdnHCl] but the values for ΔG_{H_2O} and m calculated from the kinetic experiments agree with those obtained from equilibrium experiments. This shows that only two states, the native and denatured, are significantly populated on either the unfolding or refolding pathways. There is no evidence from the kinetics of refolding of CI2 (Figure 4) for the existence of an intermediate state on the folding pathway.

It is now generally believed that proteins must fold through intermediate states (Ptitsyn, 1987), and there is increasing evidence for the existence of transient intermediates on the folding pathway [Röder et al. (1988), Udgaonkar and Baldwin (1988), Bycroft et al. (1990), Matouschek et al. (1990), and see Kim and Baldwin (1990) for a review on folding intermediates]. Since intermediates states are not detected in either the equilibrium experiments or the kinetic experiments for CI2, this suggests that they are high in energy and only transiently populated.

How Unfolded Is the Denatured State? There is some evidence suggesting that the nature of the denatured state depends upon the denaturation conditions (Creighton, 1990). It is thought that the GdnHCl denatured state is the most extended and has the least order of all denatured states, including the thermally denatured state (Tanford, 1968). For CI2, the values of ΔG_{H_2O} calculated from both GdnHCl denaturation experiments and thermal denaturation experiments are 7.03 ± 0.16 and 7.18 ± 0.43 kcal mol^{-1}, respectively, at 25 °C. This shows that the denatured states formed by GdnHCl and thermal denaturation are, if not structurally, thermodynamically equivalent.

Values for ΔG_{H_2O} and m, measured by GdnHCl-induced denaturation, are 7.03 ± 0.16 kcal mol^{-1} and 1.79 ± 0.04 kcal mol^{-1} M^{-1}, respectively. Whereas the value of ΔG_{H_2O} lies within the range of values that have been measured for small globular proteins (Crighton, 1990), the m value seems small in comparison to many others. Barnase, for example, has an m value from GdnHCl denaturation studies of 4.51 kcal mol^{-1} M^{-1} (Kellis et al., 1989), a mutant of T4 lysozyme has an m value of 4.5 kcal mol^{-1} M^{-1} (Chen & Schellman, 1989), RNase A and RNase T_1 have values of 3.0 and 2.6, respectively, at pH 7 (Pace et al., 1990), and α-chymotrypsin and β-lactoglobulin have values of 4.1 and 3.9, respectively (Greene & Pace, 1974). The m value is proportional to $\Delta\alpha$, the average fractional change in degree of exposure of residues on denaturation. CI2 contains a large reactive site loop, with many residues accessible to solvent and, in addition, an unstructured N-terminus of some 20 amino acids that are completely exposed to solvent even in the native state. Therefore, the average change in degree of exposure on unfolding will be less than for other, more globular, proteins.

A value of $\Delta\alpha$ can be calculated both from the crystal structure and from the experimental m value. $\Delta\alpha$ is calculated from the experimental m value by using the amino acid composition of the protein and the known free energies of transfer of amino acids from water into various concentrations of denaturant, δg_{tr}, (Pace, 1986). The total free energy of transfer, ΔG, can be calculated as a function of [GdnHCl] by summation of the free energy of transfer over all residues (see eq 14). The m value can then be related directly to $\Delta\alpha$ using

$$\Delta G = \Delta G_{H_2O} + \Delta\alpha_i \sum n_i \delta g_{tr,i} \qquad (14)$$

eqs 2 and 14. The value for $\Delta\alpha$ for CI2 determined in this way is 0.38, compared with a value for barnase of 0.53 calculated by using the same method (Pace et al., 1990).

$\Delta\alpha$ can also be calculated from the crystal structure by using solvent accessible surface area calculations. The solvent accessible surface area of all the residues in native CI2 was calculated relative to an extended chain conformation that represents the unfolded form (Miller et al., 1987). This gives a value for $\Delta\alpha$ of 0.40. Within the accuracy of the calculations, the two values for $\Delta\alpha$ agree. The low experimental m value, therefore, seems to be a consequence of the less globular nature of CI2 relative to other small proteins and not a result of significant residual structure in the unfolded form.

Conclusions. In conclusion, CI2 is representative of the simplest possible system for studying protein folding. It is a rare example of a classical two-state system where only the native and denatured states are significantly populated, either under equilibrium conditions or, on the unfolding or refolding pathways, under nonequilibrium conditions. In addition, there is no evidence for significant residual structure in the denatured state. CI2 is, therefore, an ideal system for a simple analysis

of both the equilibrium and kinetics of protein folding.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Yuri Griko and Andreas Matouschek for performing the calorimetric experiments in the laboratory of Professor P. L. Privalov. We are also grateful to Dr. Tim Hubbard and Dr. Cyrus Chothia from the IRC, Cambridge, for calculations on the solvent accessible surface area of CI2.

REFERENCES

Baldwin, R. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8069–8072.

Bycroft, M., Matouschek, A., Kellis, J. T., Serrano, L., & Fersht, A. R. (1990) *Nature 346*, 488–490.

Chen, B., & Schellman, J. A. (1989) *Biochemistry 28*, 685–691.

Clore, G. M., Gronenborn, A. M., Kjaer, M., & Poulsen, F. M. (1987a) *Protein Eng. 1*, 305–311.

Clore, G. M., Gronenborn, A. M., James, M. N. G., Kjaer, M., McPhalen, C. A., & Poulsen, F. M. (1987b) *Protein Eng. 1*, 313–318.

Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231–297.

Creighton, T. E. (1988) *Proc. Natl. Acad. Sci. U.S.A 85*, 5082–5086.

Creighton, T. E. (1990) *Biochem. J. 270*, 1–16.

Edelhoch, H. (1967) *Biochemistry 6*, 1948–1954.

Edelhoch, H., & Osbourne, J. (1976) *Adv. Protein Chem. 30*, 183–245.

Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319–326.

Greene, R. F., Jr., & Pace, C. N. (1974) *J. Biol. Chem. 249*, 5388–5393.

Ikeguchi, M., Kuwajima, K., Mitani, M., & Sugai, S. (1986) *Biochemistry 25*, 6965–6972.

Jackson, S. E., & Fersht, A. R. (1991) *Biochemistry* (following paper in this issue).

Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.

Jonassen, I. (1980) *Carlsberg Res. Commun. 45*, 47–48.

Kellis, J. T., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914–4922.

Kiefhaber, T., Quaas, R., Hahn, U., & Schmid, F. X. (1990a) *Biochemistry 29*, 3053–3061.

Kiefhaber, T., Quaas, R., Hahn, U., & Schmid, F. X. (1990b) *Biochemistry 29*, 3061–3070.

Kiefhaber, T., Grunert, H. P., Hahn, U., & Schmid, F. X. (1990c) *Biochemistry 29*, 6475–6480.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Kitamura, S., & Sturtevant, J. M. (1989) *Biochemistry 28*, 3788–3792.

Kjaer, M., & Poulsen, F. M. (1987) *Carlsberg Res. Commun. 52*, 355–362.

Kjaer, M., Ludvigsen, S., Sorensen, O. W., Denys, L. A., Kindtler, J., & Poulsen, F. M. (1987) *Carlsberg Res. Commun. 52*, 327–354.

Kuwajima, K., Sakuraoka, A., Fueki, S., Yoneyama, M., & Sugai, S. (1988) *Biochemistry 27*, 7419–7428.

Longstaff, C., Campbell, A., & Fersht, A. R. (1990) *Biochemistry 29*, 7339–7347.

McPhalen, C. A., & James, M. N. G. (1987) *Biochemistry 26*, 261–269.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Matouschek, A., Kellis, J. T., Serrano, L., Bycroft, M., & Fersht, A. R. (1990) *Nature 346*, 440–445.

Miller, S., Janin, J., Lesk, A. M., & Chothia, C. (1987) *J. Mol. Biol. 196*, 641–656.

Pace, C. N. (1986) *Methods Enzymol. 131*, 266–279.

Pace, C. N., Laurents, D. V., & Thomson, J. A. (1990) *Biochemistry 29*, 2564–2572.

Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167–241.

Privalov, P. L. (1986) *Methods Enzymol. 131*, 4–51.

Privalov, P. L., & Gill, S. J. (1989) *Adv. Protein Chem. 39*, 191–234.

Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 272–293.

Röder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700–704.

Shortle, D. (1989) *J. Biol. Chem. 264*, 5215–5318.

Shrake, A., & Ross, P. D. (1990) *J. Biol. Chem. 265*, 5055–5059.

Svendsen, I., Jonassen, I., Hejgaard, J., & Boisen, S. (1980) *Carlsberg Res. Commun. 45*, 389–395.

Tanford, C. (1968) *Adv. Protein Chem. 23*, 121–282.

Thomas, P. G., Russell, A. J., & Fersht, A. R. (1985) *Nature 318*, 375–376.

Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694–699.

Winder, A. F., & Gent, W. L. G. (1971) *Biopolymers 10*, 1243–1251.