ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/21787721

Transient association of the first intermediate during the refolding of bovine carbonic anhydrase B

ARTICLE *in* BIOTECHNOLOGY PROGRESS · MARCH 1992

Impact Factor: 2.15 · DOI: 10.1021/bp00014a002 · Source: PubMed

CITATIONS
28

READS
11

2 AUTHORS, INCLUDING:

Daniel I Wang
Massachusetts Institute of Technology
201 PUBLICATIONS **9,352** CITATIONS

SEE PROFILE

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.

Available from: Daniel I Wang
Retrieved on: 01 February 2016

Transient Association of the First Intermediate during the Refolding of Bovine Carbonic Anhydrase B

Jeffrey L. Cleland† and Daniel I. C. Wang*,‡

Biotechnology Process Engineering Center, Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Pharmaceutical Research & Development, Genentech, Inc., South San Francisco, California 94080

Many proteins which aggregate during refolding may form transiently populated aggregated states which do not reduce the final recovery of active species. However, the transient association of a folding intermediate will result in reduced refolding rates if the dissociation process occurs slowly. Previous studies on the refolding and aggregation of bovine carbonic anhydrase B (CAB) have shown that the molten globule first intermediate on the CAB folding pathway will form dimers and trimers prior to the formation of large aggregates (Cleland, J. L.; Wang, D. I. C. *Biochemistry* **1990**, *29*, 11072–11078; Cleland, J. L.; Wang, D. I. C. In *Protein Refolding*; Georgiou, G., DeBernardez-Clark, E., Eds.; ACS Symposium Series 470; American Chemical Society: Washington, DC, 1991; pp 169–179). Refolding of CAB from 5 M guanidine hydrochloride (GuHCl) was achieved at conditions ($[CAB]_f$ = 10–33 μM, $[GuHCl]_f$ = 1.0 M) which allowed complete recovery of active protein as well as the formation of a transiently populated dimer of the molten globule intermediate on the refolding pathway. A kinetic analysis of CAB refolding provided insight into the mechanism of the association phenomenon. Using the kinetic results, a model of the refolding with transient association was constructed. By adjusting a single variable, the dimer dissociation rate constant, the model prediction fit both the experimentally determined active protein and dimer concentrations. The model developed in this analysis should also be applicable to the refolding of proteins which have been observed to form aggregates during refolding. In particular, the transient association of hydrophobic folding intermediates may also occur during the refolding of other proteins. This analysis could, therefore, be applied to assess the reversible formation of aggregates during refolding.

Introduction

During the refolding of denatured proteins, aggregation often occurs, resulting in a decreased recovery of the native protein. Refolding at low denaturant concentrations [<1 M guanidine hydrochloride (GuHCl) or <4 M urea] and high protein concentrations (milligrams per milliliter) results in aggregation and often complete precipitation of the protein. Protein aggregation during refolding can be attributed to the association of partially folded protein species. Several studies have shown that these partially folded protein structures are hydrophobic folding intermediates (Brems et al., 1988; Mitraki and King, 1989; Cleland and Wang, 1990). In addition, transient association of a hydrophobic intermediate was observed in the refolding of bovine growth hormone, and yet, the protein eventually recovers complete activity (Brems et al., 1987). Therefore, other proteins which have been observed to aggregate at low denaturant concentrations (<1 M GuHCl) and high protein concentrations (milligrams per milliliter) may also transiently form associated species which reduce the rate of refolding.

To examine the phenomenon of transient protein association for other proteins which have been shown to aggregate during refolding, bovine carbonic anhydrase B (CAB) was studied under conditions (1 M GuHCl) which result in complete recovery of active protein. Previous studies on the refolding and aggregation of CAB revealed that the molten globule first intermediate in the refolding pathway will aggregate to form dimers, trimers, and larger aggregates (Cleland and Wang, 1990). This intermediate also reversibly associated to form multimers in 2 M GuHCl (Cleland and Wang, 1991). In addition, the molten globule first intermediate has been well characterized (Dolgikh et al., 1984; Semisotnov et al., 1987; Rodionova et al., 1989). The refolding pathway of CAB has been extensively studied (Semisotnov et al., 1990, 1987; Stein and Henkens, 1978). When diluted to 1 M GuHCl, denatured CAB in 5 M GuHCl will rapidly fold to form a hydrophobic molten globule intermediate (Dolgikh et al., 1984; Semisotnov et al., 1987). Hydrophobic clusters on the surface of the first intermediate slowly collapse to form a hydrophobic core in the second intermediate, which then folds to form the native protein structure. The slow refolding steps are related to the formation of both the second intermediate and the native protein, which require proline isomerization to obtain the correct conformation (Semisotnov et al., 1990). The stability of hydrophobic molten globule intermediates such as the first CAB folding intermediate has been postulated to be the result of the slow proline isomerization process (Ptitsyn et al., 1990). Since this hydrophobic molten globule intermediate will exist during refolding for long periods of time (minutes), it is possible that the intermediate could achieve an equilibrium with an associated state. The refolding and association of CAB was therefore studied to provide additional insight into the phenomenon of protein aggregation during refolding. A kinetic model of the re-

* To whom correspondence should be addressed.
† Genentech, Inc.
‡ MIT.

8756-7938/92/3008-0097$03.00/0 © 1992 American Chemical Society and American Institute of Chemical Engineers

folding process was also developed to understand the mechanisms of transient association during refolding.

Experimental Procedures

Materials. Bovine carbonic anhydrase B (CAB), bovine serum albumin, guanidine hydrochloride (GuHCl), tris-(hydroxymethyl)aminomethane (Tris) sulfate, ethylenediaminetetraacetic acid (EDTA), and *p*-nitrophenyl acetate (pNPA) were purchased from Sigma Chemical Co. (St. Louis, MO) as molecular biology grade reagents. To check the purity of CAB (p*I* = 5.9), gel electrophoresis analysis with silver staining was performed on each lot of protein. HPLC-grade acetonitrile was obtained from J. T. Baker (Phillipsburg, NJ). Distilled deionized water from a MilliQ water purification system (Millipore Corp., Bedford, MA) was used to prepare all buffers and samples.

Methods. ***Protein Concentration.*** To determine the protein concentration of native CAB in a buffer of 50 mM Tris-sulfate and 5 mM EDTA at pH 7.5, absorbance at 280 nm was measured with a Model 8452 diode array spectrophotometer (Hewlett-Packard, Mountain View, CA). The protein concentration was then calculated by using an extinction coefficient of 1.83 $(mg/mL\ of\ protein)^{-1}\ cm^{-1}$ (Wong and Tanford, 1973) and a molecular weight of 30 000. A colorimetric dye-binding assay (Bio-Rad Laboratories, Richmond, CA) with bovine serum albumin denatured in 5 M GuHCl as the standard was used to assess the concentration of CAB denatured in 5 M GuHCl.

Esterase Activity. The activity of each protein solution was determined by using the esterase activity assay as described previously (Pocker and Stone, 1967; Cleland and Wang, 1990). Briefly, CAB in 5 M GuHCl was rapidly diluted to 1.0 M GuHCl and to the desired final protein concentration. The solution was then analyzed for enzymatic activity at various times after dilution. Prior to addition of substrate, pNPA, each assay sample was diluted 10-fold by 50 mM Tris-sulfate and 5 mM EDTA, pH 7.5. The formation of the product from the enzyme action, *p*-nitrophenol (pNP), and decrease in the substrate, pNPA, were measured by absorbance at 348 and 400 nm, respectively, as described previously (Cleland and Wang, 1990). The ester hydrolysis rate constant of the native protein at the same concentration in the dilution buffer (50 mM Tris-sulfate and 5 mM EDTA, pH 7.5) was used to calculate the recovery of activity in each refolding experiment.

High-Performance Liquid Chromatography (HPLC) Analysis. Size-exclusion chromatography was performed with a Protein PAK 3000 SW column (Waters, Bedford, MA) attached to a Model HP1090 analytical HPLC which was equipped with a diode array detector (Hewlett-Packard, Mountain View, CA). Before each analysis, the column was equilibrated with 10 column volumes of elution buffer (1.0 M GuHCl, 50 mM Tris-sulfate, and 5 mM EDTA, pH 7.5). Refolding of denatured CAB in 5 M GuHCl was achieved by rapid dilution to 1 M GuHCl and different protein concentrations. Aliquots (25 μL) were taken from the refolding solution at various times after dilution. Each aliquot was immediately applied to the size-exclusion column and eluted at a flow rate of 1.0 mL/min to facilitate rapid separation.

By using the method previously developed, the concentration of each species was calculated from the area of the absorbance (280 nm) peak (Cleland and Wang, 1991). The elution times for monomer and dimer were 10.1 and 9.3 min, respectively. Since previous studies indicated that the extinction coefficient at 280 nm was the same for each species, the extinction coefficients of the monomer and dimer were assumed to be equal for refolding in 1 M GuHCl (Cleland and Wang, 1991). In addition, previous studies have shown that the unfolded protein only exists in solution for 200 ms at the refolding conditions of 1 M GuHCl (Semisotnov et al., 1987; Stein and Henkens, 1978), and therefore, the peaks which were observed by size-exclusion HPLC analysis must result from the compact protein structure or its associated states.

Absorbance Measurements of Refolding Kinetics. To measure the rate of formation of the second intermediate and native protein, the change in the aromatic amino acid residues was measured at 280 nm as described previously (Semisotnov et al., 1990; Wong and Tanford, 1973). Denatured CAB in 5 M GuHCl was rapidly diluted to 1.0 M GuHCl and different protein concentrations. After dilution, the sample was immediately (dead time = 10 s) placed in a Model 8452 diode array spectrophotometer and the absorbance at 280 nm was measured at 1-s intervals with an on-line computer system (HP Vectra, Hewlett-Packard, Mountain View, CA). After 700 s, the data collection was terminated and the final data were loaded into a spreadsheet program (Lotus 123, Lotus Corp., Cambridge, MA). A semilogarithmic plot of the absorbance versus time was performed. The observed refolding rates for the two phases were then calculated from the slopes of the lines generated from a least-squares fit of the data from the semilogarithmic plot (Wong and Tanford, 1973). The change in absorbance was measured in duplicate for each final condition. The results from each experiment at the same conditions were repeatable and the rate constants were accurately (±5%) fit to the absorbance data.

Results and Discussion

Determination of Transient Association. If a transient association process occurs during refolding at 1.0 M GuHCl, the rate of refolding at this final condition should decrease with increasing protein concentration. When refolded from 5 M GuHCl to 1 M GuHCl, CAB was observed to refold and was able to completely recover its biological activity (Stein and Henkens, 1978; Cleland and Wang, 1990). The formation rate for each intermediate species during refolding has also been well studied (Stein and Henkens, 1978; Dolgikh et al., 1984; Semisotnov et al., 1987, 1990). However, the rate of refolding at 1.0 M GuHCl was not previously determined for different final protein concentrations. To assure complete unfolding, CAB was denatured in 5 M GuHCl for greater than 12 h prior to the refolding experiments (Semisotnov et al., 1990). Several refolding experiments were performed by rapid dilution of the unfolded protein in 5 M GuHCl to 1 M GuHCl at different final protein concentrations. As shown in Figure 1, the concentration of active protein was then measured as a function of time after dilution. The recovery of active protein occurred more slowly with increasing protein concentration. Therefore, the rate of refolding was dependent on the final protein concentration and the refolding pathway at high protein concentrations (≥0.20 mg/mL, 6.7 μM) may involve the formation of an associated species.

Although the rate of refolding was reduced with increasing protein concentration, refolding at 1.0 M GuHCl and at all of the final protein concentrations shown in Figure 1 did result in the complete recovery of active protein after 1 h (data not shown). Thus, if an associated species did form during refolding at high protein concentrations, it must be populated only transiently. To

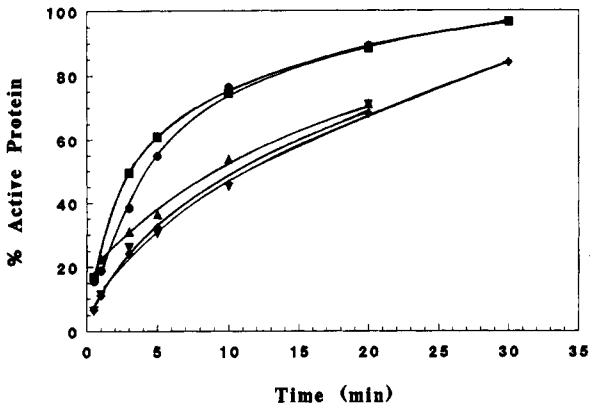

Figure 1. Refolding rate dependence on final protein concentration. CAB was refolded by rapid dilution from 5 M GuHCl to 1 M GuHCl and several different final protein concentrations. The recovery of activity was then measured as a function of time after dilution for each protein concentration: 0.10 (●), 0.20 (■), 0.35 (▲), 0.50 (◆), and 1.0 (▼) mg/mL.

determine if association was occurring during refolding at high protein concentrations, the protein was refolded by dilution from 5 M GuHCl to 1.0 M GuHCl and 1.0 mg/mL (33.3 μM) CAB. At different times after dilution, samples from the refolding solutions were analyzed by size-exclusion HPLC to measure the formation of associated protein (Figure 2). At these final conditions, size-exclusion HPLC results revealed the presence of a dimer species which decreased in concentration with time after dilution, as shown in Figure 2. The association rate process was more rapid than the time constant of the size-exclusion HPLC analysis, and therefore, only the dissociation could be measured. Since the concentration of the dimer species decreased with time, the formation of this associated state was a transient process which only reduced the rate of refolding and not the final concentration of active protein. As a control, refolding was performed at 0.10 mg/mL (3.33 μM) CAB and 1.0 M GuHCl. The dimer species was not observed by size-exclusion HPLC at these conditions. The formation of a transient dimer species is therefore dependent on the final protein concentration.

The formation of a transient dimer for refolding at 1.0 M GuHCl was comparable to the aggregation dependence on the final protein concentration which was observed in the previous refolding and aggregation studies (Cleland and Wang, 1990, 1991). The rate of dimer formation was shown to be directly dependent on the final protein concentration and inversely dependent on the final GuHCl concentration (Cleland and Wang, 1990). The dimer formation rate may also be rapid for refolding at 1.0 M GuHCl. However, since the association process was reversible at 1.0 M GuHCl, the dissociation rate may also be dependent on the final GuHCl concentration. Therefore, it can be hypothesized that as the final GuHCl concentration was increased, the rate of dissociation increased. The rate of dissociation in 1.0 M GuHCl must then be sufficiently rapid to allow the protein to refold within 1 h. In addition, since the recovery of active protein occurred at a slower rate at high protein concentrations, the transient dimer species must be formed from inactive protein.

Models for Assessment of Transient Association during Refolding. Previous kinetic and thermodynamic analyses of aggregation and refolding established that the first intermediate in the refolding pathway associates to form multimers and the rate and extent of association were dependent on the final solution conditions (Cleland and Wang, 1990, 1991). Furthermore, kinetic constants

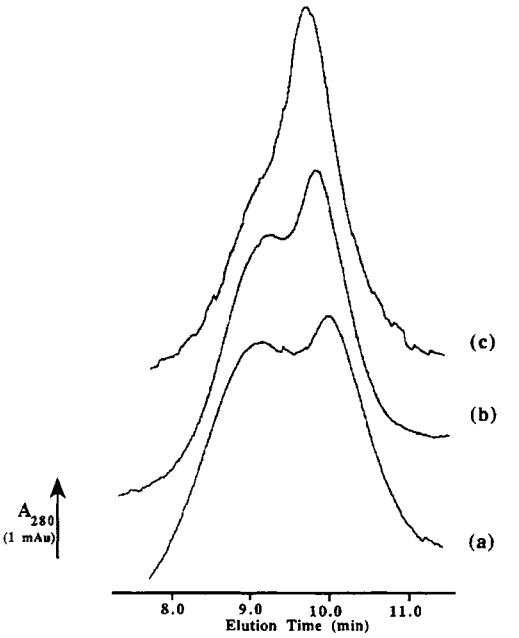

Figure 2. Measurement of transient association during refolding. Refolding was performed by rapid dilution of CAB in 5 M GuHCl to 1.0 M GuHCl and 1.0 mg/mL (33.3 μM) protein. Aliquots of the refolding solution were injected onto a size-exclusion HPLC column at a given time after dilution (see Experimental Procedures). The monomer eluted at 10.1 min and the dimer eluted at 9.3 min. The chromatographs for samples which were injected at 4.25 (a), 10 (b), and 12 (c) min after dilution are shown.

previously derived for the formation of each folding intermediate indicate that it is kinetically impossible for the second intermediate or native protein to significantly contribute to the observed association phenomenon (Semisotnov et al., 1990, 1987; Stein and Henkens, 1978). To address the effect of the final protein concentration on refolding in 1 M GuHCl, two different models were developed to describe the folding process. First of all, refolding at low protein concentrations (<10 μM) and 1.0 M GuHCl resulted in a more rapid recovery of activity than previously reported, and therefore, a detailed analysis was required to obtain the model parameters for these final conditions. If refolding was performed at higher protein concentrations (>10 μM) and 1.0 M GuHCl, the rate of refolding was greatly reduced. To describe the decreased rate of refolding, it was necessary to develop a more complex model involving a reversible bimolecular process. Refolding at both low and high protein concentrations in 1.0 M GuHCl resulted in pathways which, unlike the refolding and aggregation model (Cleland and Wang, 1990), required a detailed mathematical analysis to fully describe the folding process.

Model of Refolding at Low Protein Concentrations (<10 μM). Refolding of CAB at low protein concentrations (<10 μM) in 1.0 M GuHCl resulted in a more rapid recovery of active protein than refolding at higher protein concentrations (Figure 1). In addition, equilibrium refolding studies of CAB in 2.0 M GuHCl revealed that the first intermediate did not associate at low protein concentrations (<10 μM; Cleland and Wang, 1991). Protein aggregation during refolding at 1.0 M GuHCl and at low protein concentrations was not observed by size-exclusion HPLC studies. Therefore, the pathway for refolding of CAB at low protein concentrations in 1.0 M GuHCl can be described by a series of unimolecular reactions as shown in Figure 3, where the reverse reaction rates were assumed to be small relative to the forward rates (Semisotnov et al., 1987). When diluted from 5.0 M GuHCl to 1.0 M

$$U \xrightarrow{k_{i1}} I_1 \xrightarrow{k_{i2}} I_2 \xrightarrow{k_n} N$$

Figure 3. Refolding at low protein concentrations (<10 μM) and 1.0 M GuHCl results in the rapid formation of the first intermediate (I_1), which will fold to form the second intermediate (I_2) with the rate constant k_{i2}. The second intermediate folds to form the native protein (N) with a rate constant k_n. The rate constants and the assumptions made in deriving the equations are presented in the text.

GuHCl, the unfolded protein, U, will fold to form the first intermediate, I_1, with a rate constant of k_{i1}. The first intermediate will continue to fold to form the second intermediate, I_2. Finally, the native protein, N, is formed from the second intermediate (Semisotnov et al., 1990, 1987; Stein and Henkens, 1978).

The pathway depicted in Figure 3 can be modeled by a series of rate expressions for each species. The rate of decrease in the unfolded protein ($d[U]/dt$) may be expressed as

$$d[U]/dt = -k_{i1}[U] \quad (1)$$

The initial concentration of unfolded protein is equivalent to the total protein concentration ($[CAB]_f$), and the concentration profile of the unfolded protein can be determined from integration of eq 1:

$$[U] = [CAB]_f \exp(-k_{i1}t) \quad (2)$$

The kinetics for the first intermediate can be described by the rate expression

$$d[I_1]/dt = k_{i1}[U] - k_{i2}[I_1] \quad (3)$$

By using eqs 2 and 3, the first intermediate concentration as a function of time may be written as

$$[I_1] = \left(\frac{k_{i1}[CAB]_f}{k_{i2} - k_{i1}}\right)[\exp(-k_{i1}t) - \exp(-k_{i2}t)] \quad (4)$$

Previous refolding studies indicated that the first intermediate was formed very rapidly and that the rate was not dependent on the initial protein concentration (k_{i1} = 23.1 s^{-1}, $t_{1/2}$ = 0.03 s; Semisotnov et al., 1987). The rate of formation of the first intermediate was much greater than the rate of formation of the second intermediate or the native protein ($k_{i1} \gg k_{i2}$ or k_n) as reported previously (Semisotnov et al., 1990, 1987; Stein and Henkens, 1978). Therefore, eq 4 can be reduced to a simple exponential decay dependent on the rate constant for conversion to the second intermediate:

$$[I_1]=[CAB]_f \exp(-k_{i2}t) \quad (5)$$

The formation of the second intermediate followed by its conversion to the native protein can be described by the rate expression

$$d[I_2]/dt = k_{i2}[I_1] - k_n[I_2] \quad (6)$$

Using this rate expression and eq 5, the concentration profile for the second intermediate can then be expressed as

$$[I_2] = \left(\frac{k_{i2}[CAB]_f}{k_n - k_{i2}}\right)[\exp(-k_{i2}t) - \exp(-k_n t)] \quad (7)$$

Finally, the second intermediate folds to form the native protein as shown in Figure 3, and the rate of native protein formation may be represented by

$$d[N]/dt = k_n[I_2] \quad (8)$$

By combination of eqs 7 and 8, the concentration profile for native protein can be expressed as

$$[N] = [CAB]_f\left\{\left(\frac{k_{i2}}{k_n - k_{i2}}\right)\exp(-k_n t) - \left(\frac{k_n}{k_n - k_{i2}}\right) \times \exp(-k_{i2}t) + 1\right\} \quad (9)$$

Equations 5, 7, and 9 were used to completely describe the pathway depicted in Figure 3.

Since previous refolding kinetics were studied at high protein concentrations where the protein has been shown to form transient dimer species, refolding at low protein concentrations will result in a greater refolding rate than that reported previously [0.40 mg/mL (13.3 μM) CAB, 1.0 M GuHCl; Semisotnov et al., 1990]. One study conducted at low protein concentrations (5–10 μM) and 0.70 M GuHCl indicated a more rapid recovery of activity than at high protein concentrations (Henkens et al., 1982). The increase in refolding rate should only be reflected by a change in the rate of conversion from the first intermediate to the second intermediate (k_{i2}) since the equilibrium association occurred through the first intermediate. Previous studies on the refolding of CAB have not investigated the enzymatic activity of each intermediate. However, since the molten globule first intermediate is stable and fully populated in 2 M GuHCl and the protein is completely inactive in 2 M GuHCl, the first intermediate is assumed to be inactive (Wong and Tanford, 1973; Rodionova et al., 1989). The second intermediate should exist at equilibrium between 1.0 M (native) and 2.0 M (first intermediate) GuHCl. Denaturation experiments by Wong and Tanford revealed a slight transition in the enzymatic activity at 1.4 M GuHCl and the protein was 25% active at this condition (Wong and Tanford, 1973). Therefore, as a first approximation, the second intermediate was assumed to be 25% active (Wong and Tanford, 1973; Rodionova et al., 1989). With this approximation, eqs 5, 7, and 9 were used with the previously reported rate constant for the conversion of the second intermediate to the native protein ($k_n = 6.93 \times 10^{-2}$ min^{-1}, $t_{1/2}$ = 10 min; Semisotnov et al., 1990) to determine the rate constant for the formation of the second intermediate, k_{i2}. The value of k_{i2} was adjusted to fit the model to the activity data obtained for refolding at 0.10 mg/mL (3.33 μM) and 1.0 M GuHCl as shown in Figure 1. The rate constant for conversion from the first intermediate to the second intermediate, k_{i2}, at low protein concentrations and 1.0 M GuHCl was thereby calculated to be 1.39 min^{-1} ($t_{1/2}$ = 30 s), which was much greater than that reported previously (k_{i2} = 0.297 min^{-1}, $t_{1/2}$ = 120 s; Semisotnov et al., 1990; Stein and Henkens, 1978). The difference in these rates was the result of the side reaction of association, which was not considered in the previous studies.

Model of Refolding at High Protein Concentrations (>10 μM). In contrast to the pathway shown in Figure 3, refolding at high protein concentrations (>10 μM) and 1.0 M GuHCl resulted in slower refolding of the protein (Figure 1) and the observed formation of a transient dimer species (Figure 2). The dimer has been shown to occur from an association of the first intermediate in the refolding pathway (Cleland and Weng, 1990, 1991). Therefore, the first intermediate will reversibly associate to form a dimer under refolding conditions at high protein concentrations (>10 μM). The model for refolding at high protein concentrations in 1.0 M GuHCl could then be described by the pathway shown in Figure 4. In addition, the dissociation rate should be slower than the association rate, since the association of the first intermediate has

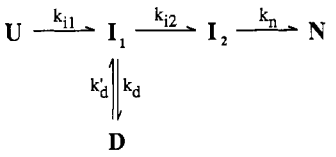

Figure 4. Refolding at high protein concentrations (>10 μM) and 1.0 M GuHCl resulted in the formation of a transient dimer species (D). The rate constants and the assumptions used to derive the equations are discussed in the text.

been shown to occur with little reversibility at low GuHCl concentrations (<1.0 M) and several protein concentrations (0.10–1.0 mg/mL) (Cleland and Wang, 1990).

With the values for the rate constants at refolding in the absence of association and the knowledge of transient dimer formation during refolding, the model and parameters for refolding with the association pathway were determined as shown in Figure 4. For this model, the rate of decrease in the unfolded protein was represented by eq 1. Since the rate of formation of the first intermediate was very rapid relative to the other folding steps and independent of the initial protein concentration (see derivation of eq 5), the initial concentration of the first intermediate was assumed to be equal to the total protein concentration ($[I_1]_{t=0} = [CAB]_f$). With this assumption, the rate expression for the formation of first intermediate becomes

$$d[I_1]/dt = k'_d[D] - k_{i2}[I_1] - k_d[I_1]^2 \qquad (10)$$

To compare the results obtained for refolding at high protein concentrations with those reported in the literature, eq 10 can be simplified by grouping the parameters into an apparent rate constant, k_{app}:

$$k_{app} = -k'_d[D]/[I_1] + k_{i2} + k_d[I_1] \qquad (11)$$

As an approximation, the apparent rate constant can be used to determine the concentration as a function of time for each species if the dependence of k_{app} on the first intermediate and the dimer is neglected. This assumption resulted in the same equations for each species as before, where k_{i2} would be replaced by k_{app} (eqs 5, 7, and 9). Using this apparent rate constant, the data were fitted to the activity data from refolding at 1.0 mg/mL (33.3 μM) CAB and 1.0 M GuHCl. At these conditions, the apparent rate constant was 0.337 min^{-1} ($t_{1/2}$ = 124 s) which was very similar to the previously reported rate constant (k_{i2} = 0.297 min^{-1}; Semisotnov et al., 1990, 1987; Stein and Henkens, 1978). To confirm these rates, refolding was performed by dilution of CAB in 5 M GuHCl to 1.0 M GuHCl and 0.50 mg/mL (16.7 μM) protein. After dilution to the final conditions, the refolding was measured by the change in absorbance at 280 nm as shown in Figure 5. The absorbance results shown in Figure 5 can be modeled by two single-exponential functions, which has been previously observed (Semisotnov et al., 1990). The two exponential rate constants were calculated to be 0.30 min^{-1} ($t_{1/2}$ = 120 s) and 7.24 × 10^{-2} min^{-1} ($t_{1/2}$ = 9.5 min). These values were comparable to the previously reported kinetics (k_{i2} = 0.297 min^{-1} and k_n = 6.93 × 10^{-2} min^{-1}; Semisotnov et al., 1990, 1987; Stein and Henkens, 1978).

The assumption of an apparent rate constant did not consider the effect of the protein concentration on the rate of refolding. It was therefore necessary to analyze more rigorously the pathway described in Figure 4. First of all, the rate of dimer formation has been shown to occur rapidly at low GuHCl concentrations and the association rate in 2 M GuHCl is also a rapid process (k_d = 5.16 min^{-1} μM^{-1}). When refolding occurred at high protein concentrations, the forward rate of dimer formation was very rapid since the initial concentration of the first intermediate was equal to the total protein concentration. Therefore, the rate constants for the formation of dimer and the first intermediate were greater than those for the other steps on the pathway ($k_{i1} > k_d \gg k'_d, k_{i2}$, and k_n). The rate expression for dimer formation is

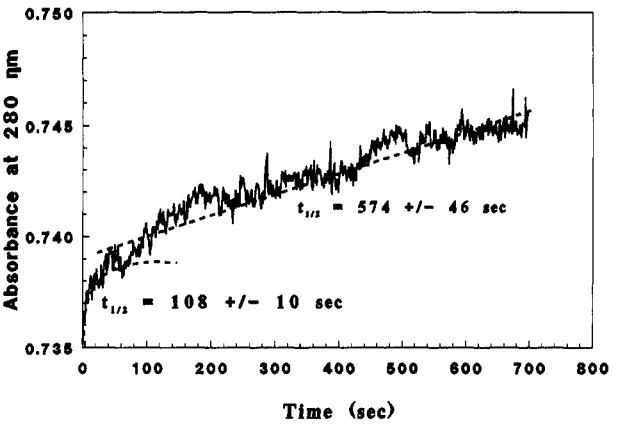

Figure 5. Refolding kinetics measured by absorbance at 280 nm. Absorbance measurements were made within 10 s of dilution from 5 M GuHCl to 1 M GuHCl and 0.50 mg/mL CAB. The curve can be modeled by two exponentials with half-times of 108 and 574 s as shown (Semisotnov et al., 1990).

$$d[D]/dt = k_d[I_1]^2 - k'_d[D] \qquad (12)$$

The equilibrium constant for the association of the first intermediate is defined as

$$K_D = [D]/[I_1]^2 \qquad (13)$$

and, therefore, eq 12 becomes

$$d[D]/dt = (k_d/K_D)[D] - k'_d[D] \qquad (14)$$

By assuming a rapid equilibrium association ($K_D \gg 1$), the rate expression for the dimer can be reduced to

$$d[D]/dt = -k'_d[D] \qquad (15)$$

and the concentration profile for the dimer can be calculated from the solution:

$$[D] = [D]|_{t=0} \exp(-k'_d t) \qquad (16)$$

The initial concentration of dimer can be calculated from the rapid equilibrium approximation at the initial condition as described by eq 13, assuming the initial concentration of the first intermediate ($[I_1]_{t=0}$) was equal to the total protein concentration ($[CAB]_f$) for the rapid folding of the unfolded protein ($k_{i1} \gg k_d$ or k_{i2}).

After determination of the concentration profile for the dimer, the concentration profile of the first intermediate was calculated assuming that the conversion from the first intermediate to the second intermediate was slow relative to the equilibrium association. Using this assumption, the rate equation for the first intermediate as shown by the model in Figure 4 can be reduced to

$$d[I_1]/dt = k'_d[D] - k_{i2}[I_1] \qquad (17)$$

The solution to this expression for the concentration of the first intermediate ($[I_1]$) yielded a relationship between the dissociation rate constant, k'_d, and the first interme-

diate:

$$[I_1] = \left(\frac{k'_d[D]_{t=0}}{k_{i2} - k'_d}\right) \exp(-k'_d t) + \left([CAB]_f + \frac{k'_d[D]_{t=0}}{k'_d - k_{i2}}\right) \exp(-k_{i2} t) \quad (18)$$

Next, the second intermediate rate expression was a simple relationship between the rate of formation from the first intermediate and the rate of folding to the native protein and can be expressed as

$$d[I_2]/dt = k_{i2}[I_1] - k_n[I_2] \quad (19)$$

Equations 18 and 19 were then used to determine the concentration profile for the second intermediate:

$$[I_2] = a \exp(-k'_d t) + b \exp(-k_{i2} t) + c \exp(-k_n t) \quad (20)$$

$$a = \frac{k'_d[D]_{t=0}}{(k_n - k'_d)(k_{i2} - k'_d)}$$

$$b = \frac{[I_1]_{t=0} - \left(\frac{k'_d[D]_{t=0}}{k_{i2} - k'_d}\right)}{k_n - k_{i2}}$$

$$c = \frac{k'_d[D]_{t=0}}{(k'_d - k_n)(k_{i2} - k'_d)} + \frac{[I_1]_{t=0} - \left(\frac{k'_d[D]_{t=0}}{k_{i2} - k'_d}\right)}{k_{i2} - k_n}$$

After calculation of the concentration of the first and second intermediates as well as the dimer concentration at each time, a simple mass balance on the total protein yielded the concentration profile for native protein:

$$[N]_f = [CAB]_f - [I_1]_t - [I_2]_t - 2[D]_t \quad (21)$$

Equations 16, 18, 20, and 21 now completely describe the pathway shown in Figure 4.

With the equations necessary to model the pathway with reversible association, the parameters for refolding were next determined. First of all, the rate constants for the formation of the second intermediate (k_{i2}) and the native protein (k_n) were assumed to be the same as those calculated from the previous model presented in Figure 3 (k_{i2} = 1.39 min^{-1}) and previous studies (k_n = 6.93 × 10^{-2} min^{-1}; Semisotnov et al., 1990, 1987; Stein and Henkens, 1978). Therefore, the only adjustable parameter in the series of equations is the rate constant for the dissociation of the dimer (k'_d). To determine this parameter, activity data from refolding at 1.0 mg/mL (33.3 μM) and 1.0 M GuHCl was used (Figure 1). As an initial estimate of the dissociation rate constant, the dissociation constant calculated from previous equilibrium studies (Cleland and Wang, 1991) was modified to reflect the different GuHCl concentration by using the relationship developed for the dimer formation rate during aggregation and refolding (Cleland and Wang, 1990):

$$k'_d = k'_d|_{eq,2MGuHCl} \left(\frac{1\ M\ GuHCl}{2\ M\ GuHCl}\right)^{-6.7} \quad (22)$$

This equation resulted in an initial dissociation rate constant, k'_d, of 0.407 min^{-1}. From this value, the initial concentrations of the first intermediate and dimer were calculated to be 3.3 μM and 15 μM, respectively. These values were obtained by using equation 13 with an equilibrium constant of 1.3 μM^{-1}, which was determined from previous equilibrium studies (Cleland and Wang, 1991).

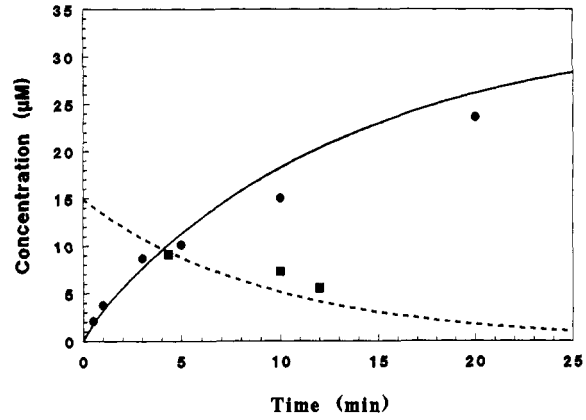

Figure 6. Reversible association pathway model comparison to refolding data. Refolding was performed at 1.0 mg/mL CAB and 1.0 M GuHCl. The recovery of active protein is plotted as a function of time (●) for comparison with the model prediction (—), which is calculated as described in the text. In addition, the concentration of dimer species was calculated from the data in Figure 2 (■). The model prediction of the dimer concentration (- - -) was also calculated as described in the text.

These values were then used to calculate the concentration profile for the active protein as a function of time. The value of k'_d was iteratively changed until a reasonable fit (±15%) to the activity data was achieved as shown in Figure 6. The dissociation rate constant, k'_d, for this fit was found to be 0.107 min^{-1}. The predicted dimer concentrations were then plotted and compared to the experimental dimer concentrations, which have been previously presented (Figure 2). The close match between the model prediction for both the experimentally determined active protein and dimer concentrations was successfully achieved by modification of only one parameter (k'_d). The deviation from the model prediction of active protein at longer time scales is the result of an overestimate of the kinetic constant for formation of the native protein from the second intermediate (k_n = 6.93 × 10^{-2} min^{-1}; Semisotnov et al., 1990, 1987; Stein and Henkens, 1978). In addition, the final association rate constant, k_d, was calculated to be 0.138 min^{-1} μM^{-1} based on the equilibrium relationship and the value of k'_d. These results also support the assumptions used to develop the model. The first assumption was the rapid equilibrium association ($K_D \gg 1$), which was true for the conditions used to fit this model (33.3 μM CAB). Since the equilibrium association was rapid at these conditions, the equilibrium prediction of the initial concentrations (eq 13) was also a valid assumption. However, the complete validation of this assumption would require the measurement of dimer concentrations at very early refolding times (<1 min). Overall, the model equations (eqs 16, 18, 20, and 21) accurately represented the reversible association pathway (Figure 4) and the observed decrease in the refolding rate at high protein concentrations (>10 μM) and 1.0 M GuHCl.

Conclusion

The successful modeling of the transient association phenomenon provided insight into the mechanism of protein aggregation during refolding. Following the initial collapse of unfolded CAB to a molten globule structure, a rapid association of the hydrophobic molten globule occurred, resulting in the formation of a transient dimer. The rate of refolding was reduced by the self-association of the molten globule first intermediate in the refolding pathway of CAB. The observed rapid association process

indicated that the rate of refolding would be reduced by the time required for dissociation of the dimer. These results also clearly showed that the formation of aggregates can occur under conditions which are normally observed to result in complete refolding. The association of a hydrophobic molten globule intermediate has also been observed during the refolding of bovine growth hormone (Brems et al., 1987). The rate of dimer formation under the conditions used in these studies (1.0 mg/mL CAB, 1 M GuHCl; $k_d = 0.138\ \mu M^{-1}\ min^{-1}$) was comparable to the second-order rate constant described by Kiefhaber et al. (1991) ($K_2 = 10^5\ M^{-1}\ s^{-1}$) for the aggregation of lactate dehydrogenase as determined by light scattering. To address these observations, additional aggregation and refolding studies of proteins which refold through a hydrophobic molten globule intermediate should be performed. The results of these studies should yield further insight into the possibility of a relationship between the protein properties and the aggregation of folding intermediates.

Notation

$[CAB]_f$	final protein concentration after dilution, mg/mL or μM
$[D]$	dimer concentration, μM
$[D]_t$	dimer concentration at a given time t, μM
$[D]_{t=0}$	dimer concentration immediately after dilution, μM
$[GuHCl]_f$	final GuHCl concentration after dilution, M
$[I_1]$	concentration of first intermediate in CAB refolding, μM
$[I_1]_t$	CAB first intermediate concentration at a given time t, μM
$[I_1]_{t=0}$	CAB first intermediate concentration immediately after dilution, μM
$[I_2]$	concentration of second intermediate in CAB refolding, μM
$[I_2]_t$	CAB second intermediate concentration at a given time t, μM
k_{app}	apparent rate constant for formation of second intermediate in CAB refolding pathway, 1/min
k_d	association rate constant for dimer formation, $\mu M^{-1}\ min^{-1}$
k'_d	dissociation rate constant for dimer, 1/min
k_{i1}	rate constant for formation of first intermediate in CAB refolding, 1/s
k_{i2}	rate constant for formation of second intermediate in CAB refolding, 1/min
k_n	rate constant for folding to native state, 1/min
$[N]$	native protein concentration, μM
$[N]_t$	native protein concentration at a given time t, μM
t	time after dilution to final conditions, min or s
$t_{1/2}$	half-time for refolding reaction step, min or s
$[U]$	unfolded protein concentration, μM

Acknowledgment

This work was supported by National Science Foundation Grant CDR-88-03014.

Literature Cited

Brems, D. N.; Plaisted, S. M.; Dougherty, J. J.; Holzman, T. F. The kinetics of bovine growth hormone folding are consistent with a framework model. *J. Biol. Chem.* **1987**, *262*, 2590.

Brems, D. N.; Plaisted, S. M.; Havel, H. A.; Tomich, C.-S. C. Stabilization of an associated folding intermediate of bovine growth hormone. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3367.

Cleland, J. L.; Wang, D. I. C. Refolding and aggregation of bovine carbonic anhydrase B: quasi-elastic light scattering analysis. *Biochemistry* **1990**, *29*, 11072.

Cleland, J. L.; Wang, D. I. C. Equilibrium association of a molten globule intermediate in the refolding of bovine carbonic anhydrase. In *Protein Refolding*; Georgiou, G., DeBernardez-Clark, E., Eds.; ACS Symposium Series 470; American Chemical Society: Washington, DC, 1991; pp 169–179.

Dolgikh, D. A.; Kolomiets, A. P.; Bolotina, I. A.; Ptitsyn, O. B. "Molten-globule" state accumulates in carbonic anhydrase folding. *FEBS Lett.* **1984**, *165*, 88.

Henkens, R. W.; Kitchell, B. B.; Lottich, S. C.; Stein, P. J.; Williams, T. J. Detection and characterization using circular dichroism and fluorescence spectroscopy of a stable intermediate conformation formed in the denaturation of bovine carbonic anhydrase with guanidinium chloride. *Biochemistry* **1982**, *21*, 5918.

Kiefhaber, T.; Rudolph, R.; Kohler, H.-H.; Buchner, J. Protein aggregation in vitro and in vivo quantitative model of the kinetic competition between folding and aggregation. *Bio/Technology* **1991**, *9*, 825.

Mitraki, A.; King, J. Protein folding intermediates and inclusion body formation. *Bio/Technology* **1989**, *7*, 690.

Pocker, Y.; Stone, J. T. The catalytic versatility of erythrocyte carbonic anhydrase. III. Kinetic studies of the enzyme-catalyzed hydrolysis of *p*-nitrophenyl acetate. *Biochemistry* **1967**, *6*, 668.

Ptitsyn, O. B.; Pain, R. H.; Semisotnov, G. V.; Zerovnik, E.; Razgulyaev, O. I. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **1990**, *262*, 20.

Rodionova, N. A.; Semisotnov, G. V.; Kutyshenko, V. P.; Uverskii, V. N.; Bolotina, I. A.; Bychkova, V. E.; Ptitsyn, O. B. Two-stage equilibrium unfolding of carboanhydrase B by strong denaturing agents. *Mol. Biol. (Moskow)* **1989**, *23*, 683.

Semisotnov, G. V.; Rodionova, N. A.; Kutyshenko, V. P.; Ebert, B.; Blanck, J.; Ptitsyn, O. B. Sequential mechanism of refolding of carbonic anhydrase B. *FEBS Lett.* **1987**, *224*, 9.

Semisotnov, G. V.; Uversky, V. N.; Sokolovsky, I. V.; Gutin, A. M.; Razgulyaev, O. I.; Rodionova, N. A. Two slow stages in refolding of bovine carbonic anhydrase B are due to proline isomerization. *J. Mol. Biol.* **1990**, *213*, 561.

Stein, P. J.; Henkens, R. W. Detection of intermediates in protein folding of carbonic anhydrase with fluorescence emission and polarization. *J. Biol. Chem.* **1978**, *253*, 8016.

Wong, K.-P.; Tanford, C. Denaturation of bovine carbonic anhydrase B by guanidine hydrochloride: a process involving separable sequential conformational transitions. *J. Biol. Chem.* **1973**, *248*, 8518.

Accepted January 2, 1992.

Registry No. CAB, 9001-03-0.