# Multiparameter Kinetic Study on the Unfolding and Refolding of Bovine Carbonic Anhydrase B†

Leslie F. McCoy, Jr., Elizabeth S. Rowe, and Kin-Ping Wong*

ABSTRACT: The kinetics of unfolding and refolding of bovine carbonic anhydrase B by guanidinium chloride have been studied by simultaneously monitoring several spectroscopic parameters, each of which reflects certain unique conformational features of the protein molecule. In the present report, far-UV circular dichroism (CD) was used to follow the secondary structural change, UV difference absorption was used to follow the exposure or burying of aromatic amino acid residues, and near-UV CD was used to follow tertiary structural changes during unfolding and refolding. The unfolding is described by two unimolecular rate processes, and refolding is described by three unimolecular rate processes. The minimum number of conformational species involved in the mechanism is five. The refolding of the protein followed by the above three parameters indicates that the process consists of an initial rapid phase in which the random-coiled protein is converted to an intermediate state(s) having secondary structure comparable to that of the native protein. This is followed by the burying of the aromatic amino acid residues to form the interior of the protein molecule. Subsequently, the protein molecule acquires its tertiary structure and folds into a unique conformation with the formation of aromatic clusters.

The last step in protein synthesis is the folding of the polypeptide chain into its unique functional conformation. Many studies have been initiated to decipher the molecular mechanism of this process [for reviews see Tanford (1968, 1970), Baldwin (1975), Wetlaufer & Ristow (1973), and Anfinsen & Scheraga (1975)]. Bovine carbonic anhydrase B (EC 4.2.1.1) has been shown to be ideally suited for investigation of the mechanism of folding of an unfolded polypeptide chain (Wong & Tanford, 1970, 1973; Yazgan & Henkens, 1972; Wong et al., 1972; Wong & Hamlin, 1974, 1975; Ko et al., 1977; Ikai et al., 1979; McCoy & Wong, 1979). It is a single-chain protein containing no cysteine or cystine, thus allowing folding studies to focus exclusively on the successive conformational changes without complications from sulfhydryl oxidation and disulfide formation [e.g., Lindskog et al. (1971); Wong & Tanford (1973)]. Furthermore, the denaturation and renaturation have been shown to be reversible processes separable into at least three stages by equilibrium measurements (Wong & Tanford, 1970, 1973).

The acid-denatured state of bovine carbonic anhydrase B (Wong & Hamlin, 1974) has been shown to have a structure which is intermediate between the native and random-coiled conformations. Further studies by L. F. McCoy and K.-P. Wong (unpublished experiments) have demonstrated that the protein exists in conformations which are unique to the denaturants. The equilibrium transitions in the absence of the enzymatically essential Zn(II) cofactor have also been studied by Yazgan & Henkens (1972), Wong et al. (1972), and Wong & Hamlin (1975) and have shown an increased complexity.

Several investigations have also examined the kinetics of the denaturation and renaturation of bovine carbonic anhydrase B (Yazgan & Henkens, 1972; Wong & Tanford, 1970, 1973; Wong & Hamlin, 1975; Ko et al., 1977; Stein & Henkens, 1978; Ikai et al., 1979). Such studies, however, have been preliminary in nature or, in general, followed the kinetics in only one direction, thus limiting interpretation.

The purpose of the present investigation is to examine the kinetics of denaturation and renaturation by employing a multiparameter approach in order to determine the sequence in which various conformational features of the protein are formed and the conformational nature of intermediate states.

## Materials and Methods

*Materials.* Bovine carbonic anhydrase B was purified from the crude enzyme obtained from Sigma Chemical Co. on a DEAE-Sephadex column by using a stepwise elution essentially according to the procedure described earlier (Wong & Tanford, 1973). These protein preparations have been shown to be pure by sedimentation analyses and disc gel electrophoresis. They are stored either as lyophilized powder or in saturated $(NH_4)_2SO_4$ solution in the freezer (−20 °C).

Ultrapure GdmCl[1] was obtained from Schwarz/Mann and found to meet the specifications of purity (Wong et al., 1971). All other chemicals used were analytical or reagent grade.

*Methods. UV Difference Absorption and CD.* UV difference absorption spectroscopy was performed on a Cary 118-CX double-beam spectrophotometer using 1-cm quartz matched cells and thermostated with a Lauda K-2/R water bath. CD measurements were made with a Jasco J-20 spectropolarimeter also thermostated with a Lauda K-2/R water bath. The CD spectropolarimeter was calibrated with 1 mg/mL *d*-camphor-10-sulfonic acid (Cassim & Yang, 1969).

*Kinetic Experiments.* These measurements were made following rapid manual mixing which required 15–20 s. All solutions, cells, and pipets were incubated at the appropriate temperature for a minimum of 30 min prior to use. The temperature in the cuvettes was maintained at ±0.2 °C of the desired temperature throughout the experiment. Base-line drift in the Cary 118-CX spectrophotometer has a maximum of $4.5 \times 10^{-4}$ OD units/h, making corrections unnecessary. Base-line drift in the Jasco J-20 spectropolarimeter was on the order

† From the Department of Biochemistry, University of Kansas Medical Center, Kansas City, Kansas 66103 (L.F.M., E.S.R., and K.-P.W.), and the Veterans Administration Medical Center, Kansas City, Missouri 64128 (E.S.R.). *Received February 5, 1980.* This study has been supported by National Institutes of Health Grants HL 18905 and GM 22962 and by the Veterans Administration.

* Correspondence should be addressed to this author. K.-P.W. is a Research Career Development Awardee of the National Institute of General Medical Sciences (GM 70628).

[1] Abbreviations used: GdmCl, guanidinium chloride; CD, circular dichroism.

0006-2960/80/0419-4738$01.00/0 © 1980 American Chemical Society

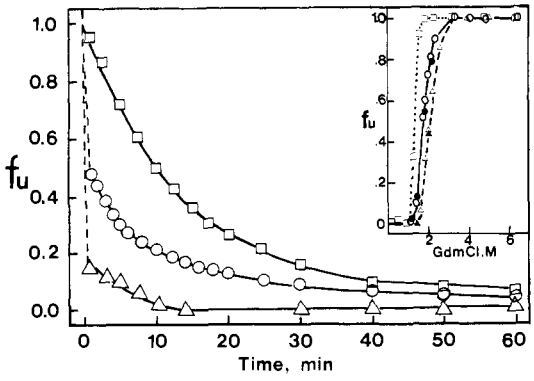


FIGURE 1: Refolding of bovine carbonic anhydrase B from 6 M GdmCl to 0.86 M as determined by far-UV CD at 222 nm (△), near-UV CD at 270 nm (□), and UV difference absorption at 291.5 nm (○). The figure shows the fraction of unfolding ($f_u$) as a function of time. The samples are maintained at 25 ± 0.1 °C at a pH of 6.5 ± 0.2, protein concentration 0.33 mg/mL. The inset shows equilibrium transition curves for the same protein as a function of [GdmCl] [taken from Wong & Tanford (1973)]: (□) near-UV CD at 269 nm; (○) difference UV absorption change at 291.5 nm; (△) optical rotation at 300 nm; filled symbols represent refolding experiments.

of 0.5 mdeg/h, and base-line corrections relative to a standard were made.

Studies by Ikai et al. (1979) and Ko et al. (1977) have shown that the kinetics of renaturation are independent of protein concentration over a 10-fold range and that the addition of excess Zn(II) decreases the rate of renaturation slightly. In the present investigation, the kinetics have been followed over a very narrow range of protein concentrations (0.29–0.34 mg/mL), and no excess Zn(II) has been added. The rates are also not appreciably affected by pH in the range of 6.5–7.5.

Resolution of kinetic phases was made by extrapolation of the semilogarithmic time-dependent plots to determine residual values and was followed to over 90% completion.

*Miscellaneous.* The concentration of bovine carbonic anhydrase B was routinely determined by spectrophotometry using an extinction coefficient of $A_{280nm}^{1mg/mL} = 1.83$ (Wong & Tanford, 1973).

## Results

*Multiparameter Kinetics of Refolding.* Refolding kinetics were studied by diluting the random-coiled protein in 6 M GdmCl to its native state in 0.86 M GdmCl. Both states have been well characterized by spectroscopic and hydrodynamic methods (Wong & Tanford, 1973). Figure 1 shows the complex refolding kinetic curves monitored by three different parameters, i.e., UV difference absorption at 291.5 nm, near-UV CD at 269 nm, and far-UV CD at 222 nm. Each of these parameters has been plotted as the fraction of unfolding, $f_u$, which is defined by the equation

$$f_u = (Y - Y_N)/(Y_D - Y_N) \quad (1)$$

where $Y$ is the observed value for the parameters and $Y_N$ and $Y_D$ are those for the parameter for the native and denatured states, respectively.

The change in ellipticity at 222 nm, which is represented by the triangles in Figure 1, reflects predominantly the change in secondary structure. The first experimental point which was observed 1 min after beginning refolding indicates that ~85% of the reaction has already occurred. The same result is obtained if the refolding is followed by an ellipticity change at ~216 nm, the wavelength at which the native protein has a minimum and also reflects secondary structural changes. Based on this result, the formation of secondary structure upon refolding was mostly completed in <1 min.

The conformational change which results in the burying of aromatic amino acid residues from the solvent during refolding was monitored by the UV difference absorption at 291.5 nm and is represented by the circles in Figure 1. The results show that about half of the total change has taken place before the first experimental point was obtained (~1 min). The same results are obtained if the refolding kinetics are followed at 286 nm, a wavelength at which both tryptophan and tyrosine contribute to the UV difference absorption change.

The third method used is the change in the near-UV CD, which reflects the formation of tertiary structure of the protein molecule, since it originates from the formation of the aromatic amino acid clusters in the protein. As can be seen by the kinetic curve represented by squares in Figure 1, almost all of the kinetic reaction can be observed during the time course of the study, i.e., ~1 to 30 min.

The kinetic results obtained indicate that the sequence of events upon refolding appears to initiate with the formation of the secondary structure within ~1 min as shown by the far-UV CD results. This step is followed by a general folding of the polypeptide chain with the secondary structural domains, resulting in the burying of aromatic amino acids into the interior of the protein as shown by the difference UV absorption measurements. Subsequently, the tertiary structure is formed by the clustering of the aromatic amino acid residues into the compact molecule.

These conclusions based on multiparameter kinetic studies are in complete agreement with the results of the equilibrium studies on the GdmCl denaturation of bovine carbonic anhydrase B (Wong & Tanford, 1970, 1973; McCoy & Wong, 1979). Their data are shown by the inset in Figure 1. The most stable region, reflected in the transition curve detected by optical rotation, is the secondary backbone folding, which has the fastest kinetics of folding. The intermediate transition curve, observed by UV difference absorption, contains several rate processes, and the curve obtained at the lowest GdmCl concentration reflects the tertiary structural changes which are also seen in the aromatic CD and have the slowest kinetics.

The kinetic data have been obtained by each of these parameters; the major portion of the data has been obtained by UV difference absorption at 235 nm because it contains contributions from each of the refolding processes.

*Kinetics as Followed by Far-UV CD.* The formation of secondary structure during folding can be monitored by changes in the far-UV CD. Because the changes between 245 and 205 nm in the CD spectra are complex, the refolding kinetics were monitored in two ways: (1) the time-dependent changes at 214 and 222 nm were followed continuously; (2) scans were made from 240 to 205 nm (8.8-min scan time) beginning at specific time intervals.

The solid curve in Figure 2 represents the far-UV CD spectrum obtained from a scan beginning 5 min after mixing. A comparison of this spectrum with the far-UV CD spectra of the native and random-coiled proteins shows that after 5 min the CD spectrum closely resembles that of the native protein (dotted curve). Clearly, many of the secondary structural features of the native protein have been formed within this time.

The dashed curve in Figure 2 represents a scan beginning 240 min after mixing. A comparison with the native protein spectrum shows that the native conformation has almost completely formed (~300 min are required for complete renaturation under these conditions). The changes which occur between 10 and 240 min indicate that more than one de-

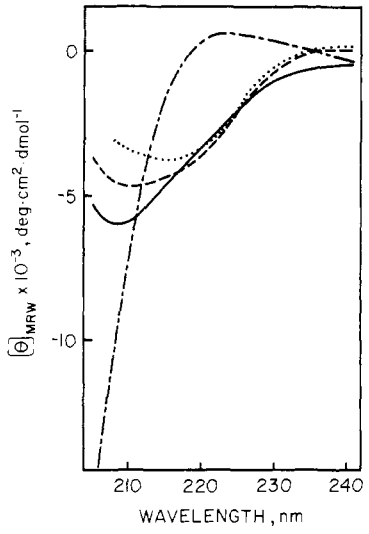


FIGURE 2: Equilibrium and kinetic composite spectra of the far-UV CD for the refolding of bovine carbonic anhydrase B. The random-coiled state in 6 M GdmCl (- - -) and the native state (···) were determined at equilibrium. The kinetic composite spectra beginning 5 min after dilution to 0.86 M GdmCl (—) and 240 min (- - -) after dilution have a scan time of 8.75 min. Protein concentration is 0.334 mg/mL, pH 6.5–6.7. All spectra are recorded at 9.0 °C. Dilution is from 6 to 0.86 M GdmCl.

Table I: Kinetics for Refolding[a] of Bovine Carbonic Anhydrase B at 9 °C

| spectroscopic methods | $\lambda_2 \times 10^4$ ($s^{-1}$) | $\lambda_3 \times 10^4$ ($s^{-1}$) |
|---|---|---|
| difference absorption changes | | |
| $\Delta A_{291.5\ nm}$ | 8.1 | 1.5 |
| $\Delta A_{280\ nm}$ | 11.2 | 1.8 |
| $\Delta A_{235\ nm}$ | 12.7 | 0.81 |
| near-UV CD changes | | |
| $\Delta[\Theta]_{269\ nm}$ | | 1.5 |
| far-UV CD changes | | |
| $\Delta[\Theta]_{222\ nm}$ | I[b] | 1.5 |
| $\Delta[\Theta]_{214\ nm}$ | I[b] | 3 |

[a] Refolding was carried out by dilution of the protein in 6–0.86 M GdmCl with a protein concentration of 0.334 mg/mL, pH 6.6. [b] I = indeterminate.

tectable event can be observed in the far-UV CD. If only one kinetic event were occurring, there would be an isosbestic point observed for the curves in Figure 2.

The changes taking place at 222 and 214 nm were monitored continuously as a function of time. Based upon initial CD values, ~85% of the far-UV CD change takes place before any measurements can be made (~30 s). The remaining ~15% change is too small to be measured reliably, but the process is very slow and probably corresponds to the slowest step observed by difference absorption measurement. The approximate rate constants are given in Table I.

*Kinetics as Followed by UV Difference Absorption at 235 nm.* Shown in Figure 3 are typical semilogarithmic plots of the kinetics of *unfolding* of the protein followed by UV difference absorption at 235 nm. Semilogarithmic plots of the residual difference absorption can be resolved into a second first-order process. As described by Ikai & Tanford (1971), unimolecular first-order kinetics can be described by equations of the form

$$(A_t - A_\infty)/(A_0 - A_\infty) = \sum_i (P_i e^{-\lambda_i t}) \qquad (2)$$

where the $A$'s represent the values of the observed parameter at the subscripted time. The macroscopic rate constants, $\lambda_i$,

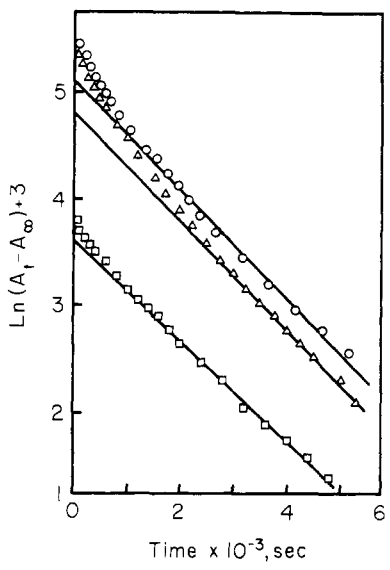


FIGURE 3: Semilogarithmic kinetic plots of unfolding monitored by difference absorption change at 235 nm. GdmCl concentrations: 1.53 (□); 1.83 (○); 1.93 M (△).

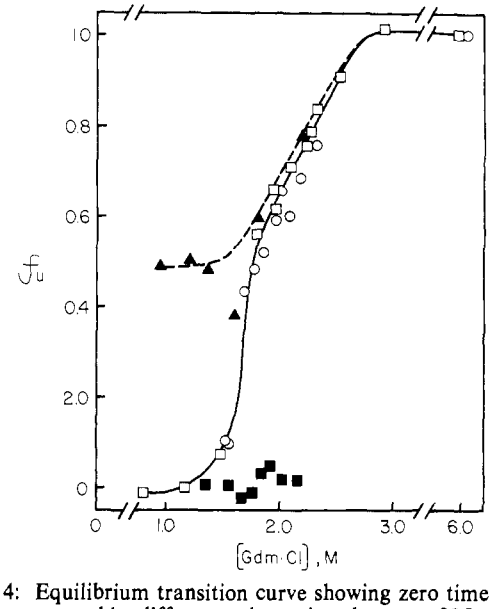


FIGURE 4: Equilibrium transition curve showing zero time kinetic points as measured by difference absorption change at 235 nm. (○) Equilibrium points from the results of Wong & Tanford (1973); (□) equilibrium points from this work; (▲) extrapolated zero time points from refolding; (■) extrapolated zero time points for unfolding.

are dependent on the microscopic rate constants only, and the coefficients $P_i$ add up to 1.0 and are dependent upon the microscopic rate constants, the parameter observed, and the initial conditions. The unfolding data at 235 nm can be described by a two-term equation:

$$(A_t - A_\infty)/(A_0 - A_\infty) = P_a e^{-\lambda_a t} + P_b e^{-\lambda_b t} \qquad (3)$$

The extrapolated zero time points are given in Figure 4. The observation that the plots in Figure 3 extrapolate to zero at zero time indicates that no additional faster process was missed. The apparent rate constants and coefficients are listed in Table II.

The *refolding* kinetics were followed by UV difference absorption at 235 nm over the same range of GdmCl concentration in the transition region. At concentrations of GdmCl above 1.75 M, the entire reaction was too fast to be followed by this technique. Below 1.75 M GdmCl, the observed data were resolved into two exponential terms; however,

Table II: Kinetics of Unfolding and Refolding of Bovine Carbonic Anhydrase B as Followed by Difference Absorption at 235 nm ($\Delta A_{235\ nm}$)

| [GdmCl] (M) | (A) Unfolding app rate const × $10^4$ ($s^{-1}$) $\lambda_a$ | $\lambda_b$ | coeff $P_a$ | $P_b$ |
|---|---|---|---|---|
| 1.51 | 14.3 | 4.46 | 0.145 | 0.608 |
| 1.53 | 12.0 | 5.15 | 0.103 | 0.449 |
| 1.66 | 11.6 | 3.83 | 0.428 | 0.727 |
| 1.73 | 17.3 | 5.10 | 0.492 | 0.675 |
| 1.83 | 17.3 | 4.64 | 0.398 | 0.464 |
| 1.93 | 43.8 | 14.6 | 0.474 | 0.446 |
| 1.98 | 70.6 | 24.2 | 0.421 | 0.414 |
| 2.06 | 92.7 | 92.7 | 0.404 | 0.404 |
| 2.16 | 228 | 93.6 | 0.451 | 0.432 |
| 2.27 | 456 | 91.6 | 0.736 | 0.156 |

| [GdmCl] (M) | (B) Refolding app rate const × $10^4$ ($s^{-1}$) $\lambda_2$ | $\lambda_3$ | coeff $P_1$ | $P_2$ | $P_3$ |
|---|---|---|---|---|---|
| 1.74 | 353 | 25.4 | 0.798 | 0.067 | 0.135 |
| 1.59 | 185 | 20.1 | 0.778 | 0.071 | 0.151 |
| 1.40 | 279 | 13.2 | 0.543 | 0.137 | 0.320 |
| 1.30 | 124 | 13.4 | 0.514 | 0.081 | 0.405 |
| 0.86 | 42.3 | 7.4 | 0.521 | 0.202 | 0.277 |

the observed data did not extrapolate to the appropriate point at zero time, indicating that at least one more exponential term in eq 2 is required to describe the data. The apparent zero time points for the refolding experiments are shown in Figure 4. As seen here and in Table I, the rapid initial reaction accounts for about 58% of the total UV difference absorption change in refolding. The refolding data is thus described by the equation

$$(A_t - A_\infty)/(A_0 - A_\infty) = P_1e^{-\lambda_1 t} + P_2e^{-\lambda_2 t} + P_3e^{-\lambda_3 t} \quad (4)$$

The apparent rate constants and coefficients are listed in Table II (B).

*Kinetics as Followed by Near-UV CD.* The near-UV CD is the parameter which reflects the least stable structural elements of the protein and originates from the tertiary structure of the protein molecule, particularly the aromatic amino acid clusters. The kinetics of unfolding and refolding were followed at 269 nm; this wavelength was chosen because it is the largest trough in the near-UV CD spectrum and because it can reasonably be assigned to one or more tyrosyl residues in the protein.

Some typical *unfolding* data are shown in Figure 5. As seen here, the plots are nearly linear, with some evidence of slight curvature. The rate constants and coefficients are given in Table III. In the unfolding experiments the process is dominated by one term in eq 2, and the rate constant is comparable to the slower of the rates observed by different UV absorption change. In *refolding*, only one rate is observed, and, when extrapolated back to zero time, it appears to go beyond the expected zero time value. These refolding results indicate that the change in near-UV CD is associated with the slower unfolding process. The coefficients of the other terms of eq 2 are nearly zero for this parameter. In the refolding experiments, the apparent value of $P$ for the observed process is >1, indicating that some other terms in eq 2 have negative values of these coefficients. This would be consistent with the interpretation that there is no change in near-UV CD during the faster refolding processes and that these processes precede the process giving rise to the change in near-UV CD, thus leading to an apparent lag in the near-UV CD change.

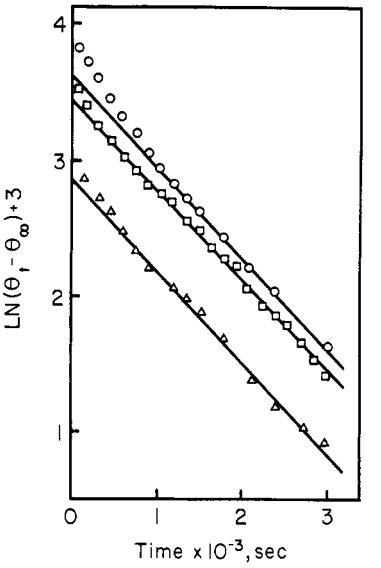


FIGURE 5: Semilogarithmic kinetic plot of unfolding monitored by near-UV CD change at 269 nm. The approach to equilibrium kinetics for carbonic anhydrase at 1.53 (Δ), 1.69 (□), and 1.81 M (○) GdmCl. Conditions are as described in Figure 1.

Table III: Kinetic Parameters for the Unfolding and Refolding of Bovine Carbonic Anhydrase B[a] as Followed by Near-UV CD ($\Delta[\Theta]$) at 269 nm

| [GdmCl] (M) | app rate const $\lambda_a$ × $10^4$ ($s^{-1}$) | coeff $P_a$ |
|---|---|---|
| | (A) Unfolding | |
| 1.53 | 4.0 | 0.30 |
| 1.60 | 6.7 | 0.43 |
| 1.69 | 5.8 | 0.73 |
| 1.81 | 7.7 | 0.98 |
| 2.05 | 82.0 | 0.96 |
| | (B) Refolding | |
| 1.88 | | |
| 1.40 | 8.5 | 1.16 |
| 1.30 | 8.4 | 1.23 |
| 0.86 | 9.9 | 1.0( |
| 0.86 | 9.5 | 0.93 |

[a] Final protein concentrations range from 0.29 to 0.34 mg/mL in the pH range 6.5–7.0.

*Proposed Kinetic Scheme.* The number of terms in the exponential eq 2 is equal to $n - 1$, where $n$ is the number of species involved in the mechanism (Ikai & Tanford, 1971). The apparent rate constants, $\lambda_i$, are dependent only upon the microscopic rate constants and are independent of the initial conditions or the observed parameter; thus the $\lambda$ values are the same whether folding or unfolding is followed, and they are the same for any observable parameter. The coefficients $P_i$, in contrast, are dependent on the initial conditions and the parameter observed as well as the microscopic rate constants; thus they can be zero or negative in a particular experiment, so that the term can be undetected.

We will first consider the data obtained in the transition region where it is possible to follow the approach to equilibrium from both directions, thus obtaining a maximum amount of information. Considering the region of GdmCl concentration around 1.5–1.7 M, we see in the inset of Figure 1 that the transition observed by UV difference absorption is halfway complete, the near-UV CD transition is nearly complete, and the transition as observed by optical rotation or far-UV CD is just beginning. Consideration of the data tabulated in Table II for these concentrations indicates that the complete eq 2 for this process requires a minimum of four terms, indicating

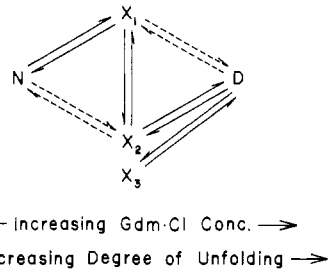


FIGURE 6: Proposed kinetic scheme showing the minimum number of conformational species. Dashed arrows represent pathways which cannot be ruled out.

the presence of at least five species. The resulting general equation is

$$[(A_t - A_\infty)/(A_0 - A_\infty)]_x = P_{1,x}e^{-\lambda_1 t} + P_{2,x}e^{-\lambda_2 t} + P_{a,x}e^{-\lambda_a t} + P_{b,x}e^{-\lambda_b t} \quad (5)$$

where $x$ designates the specific initial conditions and parameter observed. This result is based on the comparison of the apparent rate constants for refolding and unfolding as followed by UV difference absorption at 235 nm (eq 3 and 4). It is reasonable to conclude that $\lambda_3$ (from refolding experiments) is the same as $\lambda_a$ or $\lambda_b$ and also that the apparent rate constants observed in the near-UV CD at 269 nm are the same as $\lambda_a$ or $\lambda_b$ from the UV difference absorption experiments. It should be emphasized, however, that five species is a minimum estimate, and additional species are not ruled out, particularly in view of the fact that a portion of the refolding was too fast to be measured and could require more than one additional term. The single dominant term observed by following the near-UV CD change at 269 nm corresponds to one of the slower terms from the UV difference absorption data. Presumably the dominance of this term in the near-UV CD kinetics is because the CD change is due to a small region of the protein whose exposure occurs in one step. While the other steps in the unfolding mechanism affect the kinetics of this change, the $P$ values for the other processes are small or in some cases negative, as observed by near-UV CD.

Several general conclusions about the mechanism can be made by consideration of the data in the transition region, by spectral analysis, and by examination of the effect of changing GdmCl concentration on the various parameters. Figure 6 shows a schematic diagram of a kinetic scheme containing a minimum of five species that is compatible with the data and which will aid in its discussion. One feature of the proposed mechanism is that it is at least partially linear. In unfolding, only two relatively slow processes account for the whole change, whereas in refolding a rapid preequilibration occurs in which intermediates accumulate prior to the slower steps which are rate limiting in the unfolding direction. As illustrated in Figure 6, however, a portion of the reaction scheme may be cyclical. A second important feature of the proposed mechanism is the species $X_3$ which is shown as being off the direct folding pathway from D to N. This is based on some features of the dependence of the apparent rates on the GdmCl concentration discussed below. Similar wrongly folded intermediates have been detected in the folding mechanism of cytochrome $c$ and other proteins (Ikai & Tanford, 1971; Ikai et al., 1973) and previously suggested for carbonic anhydrase (Carlsson et al., 1973, 1975).

Thus far we have been considering only the data obtained over a narrow range of GdmCl concentrations near the center of the transition. The effect of GdmCl on each process provides additional insight into the features of the folding and unfolding processes. The species indicated in Figure 6 are

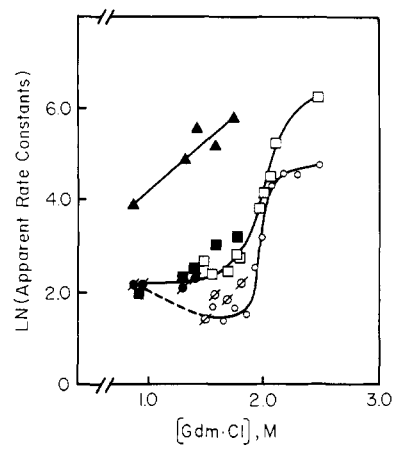


FIGURE 7: Dependence of apparent rate constants on GdmCl concentration. (▲) Refolding $\lambda_2$ ($\Delta A_{235}$); (■) refolding $\lambda_a$ or $\lambda_b$ ($\Delta A_{235}$); (◭) refolding $\lambda_a$ or $\lambda_b$ ($\Delta[\theta]_{269}$); (◮) unfolding $\lambda_a$ or $\lambda_b$ ($\Delta[\theta]_{269}$); (○) unfolding $\lambda_a$ ($\Delta A_{235}$); (□) unfolding $\lambda_b$ ($\Delta A_{235}$). Solid lines indicate dependence of individual rate constants. The dashed line indicates that in refolding both slow steps were not resolved and must be very similar.

arranged as if the GdmCl concentration and the degree of unfolding were progressing from left to right. Thus, N is the native protein, $X_1$, $X_2$, and $X_3$ are intermediate states which are present at intermediate concentrations of GdmCl, and D is the random-coiled state present at high GdmCl concentration. Consideration of the inset in Figure 1 suggests that in the center of the transition the dominant species are the intermediates $X_1$, $X_2$, and $X_3$ with very little of the native material remaining (as seen by the transition monitored by near-UV CD changes at 269 nm). The intermediate species apparently are in some ways similar to the native conformation in that they are considerably more compact than a random coil (Ikai et al., 1979), and the far-UV CD spectra are more similar to that of the native protein than to the random-coiled protein (Figure 3). As the GdmCl concentration is increased, the random-coiled protein becomes the only species present.

Figure 7 shows a plot of the apparent rate constants as a function of GdmCl concentration. It is seen here that all of the observed rates increase dramatically around 2.0 M GdmCl, as the random-coiled state D becomes significantly more stable. Each apparent rate constant is a complex function of the microscopic rate constants, depending on the exact mechanism. Qualitatively, it is concluded that the loss of secondary structure which occurs as the intermediates unfold to the random coil is extremely rapid and that these steps are responsible for the increase in all of the observed rate constants which occurs in about 2.0 M GdmCl. It is generally expected that each of the individual *microscopic* rate constants are monotonic functions of [GdmCl], with the steps toward increased exposure of hydrophobic groups increasing in rate with increasing GdmCl concentration and steps toward decreasing exposure of such buried groups decreasing with increasing denaturant concentration. For these reasons the *observed* rate constants which are functions of both refolding and unfolding microscopic constants generally pass through a minimum in the center of the transition curve. As seen in Figure 7, $\lambda_a$ and $\lambda_b$ do exhibit a relative minimum in the center of the transition. It is interesting to note, however, that $\lambda_2$, the fastest *observed* rate in refolding, continues to decrease with decreasing GdmCl concentration well below the transition region. This observation is consistent with the scheme shown in Figure 6 where in refolding there is a rapid preequilibration between D and $X_3$, and the further folding rate is then a function of the equilibrium concentration of D as well as the *microscopic* rate

constants for folding. Thus, even though the microscopic refolding rate constants are increasing as [GdmCl] decreases, the overall rate is decreased as the $D \rightarrow X_3$ equilibrium shifts toward $X_3$ and the equilibrium concentration of $D$ becomes smaller.

## Discussion

The present results show that the conformational equilibria of bovine carbonic anhydrase B as a function of [GdmCl] involve significant concentrations of a minimum of five conformational species. The native structure contains aromatic groups in asymmetric clusters which are not present in any of the partially unfolded states. The conformational change resulting in the disappearance of this spectral feature occurs first, as the GdmCl concentration increases, and it is relatively slow.

There are a minimum of three partially unfolded intermediate states whose UV difference absorption spectra and CD spectra are intermediate between the native conformation and random-coiled protein. The sequence of their appearance and disappearance could not be determined. However, it has been suggested that there is at least one which is not on the direct folding pathway. Again, the interconversion of these intermediate species is relatively slow. Ikai et al. (1979) have shown that at 2.0 M GdmCl the carbonic anhydrase has a relatively compact shape as measured by gel filtration. Our kinetic composite far-UV CD spectra also suggest that the backbone folding of the protein at this intermediate stage is similar to that of the native structure. The intermediate states differ from the native conformation primarily in spectral properties arising from aromatic groups. It thus appears that these conformations may be the result of shifts in the relationships of major folded regions of the native structure and of small shifts in the asymmetric environment of aromatic groups.

The protein secondary structure begins to unfold around 1.8 M GdmCl, giving rise to major changes in the far-UV CD. The rates of these changes in secondary structure are rapid and in our methods could not be followed quantitatively. It is the conversion of the immediate states to the random-coiled protein which give rise to the large change in hydrodynamic radius of gyration observed by Ikai et al. (1979).

Although a complete multiparameter kinetic analysis must await the studies of the fast portion of the reaction by stopped-flow CD [e.g., Luchins & Beychok (1978)] and stopped-flow spectrophotometry, the kinetic mechanism obtained in this work confirms the sequential steps proposed earlier based on equilibrium studies (Wong & Tanford, 1970, 1973; Wong & Hamlin, 1974, 1975; McCoy & Wong, 1979). The multiparameter kinetic approach used here should be useful in studying the folding mechanism of other proteins.

The general kinetic scheme presented here is consistent with other studies in the literature on bovine carbonic anhydrase B. Ko et al. (1977) reported the presence of a partially folded but enzymatically inactive intermediate. This could correspond to $X_1$ and/or $X_2$ in our scheme. Ikai et al. (1979) have observed the gradual appearance over a period of 24 h of an intermediate species which does not recover enzymatic activity upon dilution. This intermediate could be accounted for by a reversible process which has very slow kinetics because it is off the main folding pathway and there is an obligatory intermediate present in very small amounts (Rowe, 1976). In our scheme, this very slowly equilibrating intermediate would be a sixth one, which we would not have detected under our conditions.

Proline isomerization as a rate-limiting step in protein folding has been investigated and convincingly demonstrated in several cases (Brandts et al., 1975, 1977; Hagerman, 1977; Lin & Brandts, 1978; Schmid & Baldwin, 1978; Creighton, 1978; Cook et al., 1979). In light of these studies, it is possible that one or more of the processes which we observe include a cis–trans proline isomerization. This interesting possibility deserves further investigation along with additional attempts to characterize the intermediate states.

## References


Anfinsen, C. B., & Scheraga, H. A. (1975) *Adv. Protein Chem. 29*, 205–300.

Baldwin, R. L. (1975) *Annu. Rev. Biochem. 44*, 453–638.

Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953–4963.

Brandts, J. F., Brennen, M., & Lin, M.-N. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4178–4181.

Carlsson, U., Henderson, L. E., & Lindskog, S. (1973) *Biochim. Biophys. Acta 310*, 376–387.

Carlsson, U., Aasa, R., Henderson, V. E., Jonsson, B. H., & Lindskog, S. (1975) *Eur. J. Biochem. 52*, 25–36.

Cassim, J. Y., & Yang, J. T. (1969) *Biochemistry 8*, 1947–1951.

Cook, K. H., Schmid, F. X., & Baldwin, R. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6157–6161.

Creighton, T. E. (1978) *J. Mol. Biol. 125*, 401–406.

Hagerman, P. J. (1977) *Biopolymers 16*, 731–739.

Ikai, A., & Tanford, C. (1971) *Nature (London) 230*, 100–102.

Ikai, A., Fish, W., & Tanford, C. (1973) *J. Mol. Biol. 73*, 165–184.

Ikai, A., Tanaka, S., & Noda, H. (1979) *Arch. Biochem. Biophys. 190*, 39–45.

Ko, B. P. N., Yazgan, A., Yeagle, P. L., Lottich, S. C., & Henkens, R. W. (1977) *Biochemistry 16*, 1720–1725.

Lin, L.-N., & Brandts, J. R. (1978) *Biochemistry 17*, 4102–4110.

Lindskog, S., Henderson, L. E., Kannan, K. K., Liljas, A., Nyman, P. O., & Strangberg, B. (1971) *Enzymes, 3rd Ed. 5*, 587–665.

Luchins, J. I., & Beychok, S. (1978) *Science (Washington, D.C.) 199*, 425–426.

McCoy, L. F., & Wong, K.-P. (1979) *Biopolymers 18*, 2893–2904.

Rowe, E. S. (1976) *Biochemistry 15*, 905–916.

Schmid, F. X., & Baldwin, R. L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4764–4768.

Stein, P. J., & Henkens, R. W. (1978) *J. Biol. Chem. 253*, 8016–8018.

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

Tanford, C. (1970) *Adv. Protein Chem. 24*, 1–95.

Wetlaufer, D. B., & Ristow, S. (1973) *Annu. Rev. Biochem. 42*, 135–158.

Wong, K. P., & Tanford, C. (1970) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 29*, 335.

Wong, K. P., & Tanford, C. (1973) *J. Biol. Chem. 248*, 8518–8523.

Wong, K. P., & Hamlin, L. M. (1974) *Biochemistry 13*, 2678–2683.

Wong, K. P., & Hamlin, L. M. (1975) *Arch. Biochem. Biophys. 170*, 12–22.

Wong, K. P., Roxby, R., & Tanford, C. (1971) *Anal. Biochem. 40*, 459–466.

Wong, K. P., Allen, S. H., & Hamlin, L. M. (1972) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 31*, 923.

Yazgan, A., & Henkens, R. W. (1972) *Biochemistry 11*, 1314–1318.