Folding of Carp Parvalbumin Studied by Equilibrium and Kinetic Circular Dichroism Spectra†

Kunihiro Kuwajima,* Atsushi Sakuraoka, Shoichi Fueki, Michio Yoneyama, and Shintaro Sugai

Department of Polymer Science, Faculty of Science, Hokkaido University, Kita-ku, Sapporo, Hokkaido 060, Japan

Received March 21, 1988; Revised Manuscript Received May 20, 1988

ABSTRACT: The reversible equilibrium unfolding of carp parvalbumin III (p*I* = 4.25) when treated with guanidine hydrochloride and the kinetics of folding and unfolding induced by concentration jump of the denaturant have been studied by the peptide circular dichroism spectra at pH 7.0 and 4.5 °C. In the kinetic refolding reaction, a transient folding intermediate was found to be rapidly accumulated within the dead time of the stopped-flow circular dichroism (18 ms). The preequilibrium unfolding curve corresponding to the unfolding curve of the transient intermediate was obtained by measuring the refolding kinetics at various concentrations of the denaturant. Comparison of the equilibrium and preequilibrium unfolding curves has demonstrated that the total equilibrium reaction is well interpreted in terms of a three-state mechanism in which the intermediate, identical with the transient intermediate, is populated at low concentrations of the denaturant. The intermediate has been shown to have the following properties: (i) It has 60–80% of the α-helix in the native protein but does not have the specific structure responsible for strong Ca^{2+} binding; (ii) it is similar to the partially unfolded state observed at acid pH. Effects of Ca^{2+} on the unfolding equilibrium and on the refolding kinetics have also been investigated. The results have demonstrated the stepwise organization of the substructures of parvalbumin during its refolding. As a conclusion, the folding process of this protein can be divided into the three stages: (i) rapid formation of the secondary structure framework; (ii) organization of a part of the specific tertiary structure including one of the two Ca^{2+}-binding domains, with this step leading to the activated state of folding; and (iii) the final stabilization associated with organization of the rest of the molecule including the other Ca^{2+}-binding domain.

The mechanism by which the amino acid sequence of a protein directs the folding to the native conformation is a fundamental problem that remains unsolved. Two kinds of information are of special interest in studies of the folding problem: (i) the character of a transient intermediate of parrtial folding and (ii) the kinetics of the folding pathway. Kinetic circular dichroism (CD)[1] measurements provide an effective approach in both of these respects (Labhardt, 1986; Kuwajima et al., 1987). Our previous studies on the folding of α-lactalbumin and lysozyme have shown that the intermediate having native-like secondary structure accumulated rapidly at an early stage of folding can be detected and characterized by the kinetic CD spectra (Kuwajima et al., 1985; Ikeguchi et al., 1986a). Although it has been suggested that such rapid formation of the secondary structure framework in protein folding is a general phenomenon in small globular proteins, this statement may still be controversial, and the characteristics of the transient intermediate have been examined only in a limited number of proteins (Kim & Baldwin, 1982; Ptitsyn, 1987; Kuwajima et al., 1987). On the other hand, a late stage of folding is known to be a highly cooperative process associated with formation of specific tertiary structure, and this stage may, in general, involve a rate-limiting step of the folding reaction (Kuwajima, 1977; Creighton, 1980; Craig et al., 1985; Kuwajima et al., 1985; Ikeguchi et al., 1986a; Ptitsyn, 1987). The kinetic CD method is also useful for describing the reaction pathway of the late stage of folding. However, in proteins that contain proline residues, the kinetics are complicated by the presence of multiple unfolded species (Kim & Baldwin, 1982).

Carp parvalbumin is a well-known Ca^{2+}-binding protein having a molecular weight of 11 500 (Kretsinger, 1980). Because of its relatively simple architecture of the secondary and tertiary structures, this protein may serve as one of ideal proteins for the studies of protein folding. The X-ray structure of isotype III of parvalbumin has been determined by Kretsinger and Nockolds (1973). The protein is classified into an all-α-type protein and contains no disulfide bond. Because of the absence of proline in isotype III, relatively simple kinetics of folding will be expected (Lin & Brandts, 1978). The purpose of the present study is to examine the folding reaction of parvalbumin III by employing the kinetic CD measurements including stopped-flow CD. The transient intermediate observed will be compared with the intermediate characterized in the previous studies of α-lactalbumin. The kinetic pathway of folding will be discussed in relation to acquirement of its Ca^{2+}-binding capacity during the folding reaction.

Materials and Methods

Materials. Carp parvalbumin III (p*I* = 4.25) was isolated from white muscle of common carp (*Cyprinus carpio*) by the methods of Pechère et al. (1971) and Yagi et al. (1982). Gdn-HCl was of specially prepared reagent grade from Nakarai Chemicals, Ltd. All other chemicals were commercially available and of guaranteed reagent grade. The concentration

† This work was supported in part by Grants-in-Aid for Scientific Research (General Research 60580217 and 61420050 and Special Project Research 6113001) from the Ministry of Education, Science and Culture of Japan.

[1] Abbreviations: CD, circular dichroism; ORD, optical rotatory dispersion; NMR, nuclear magnetic resonance; UV, ultraviolet; Gdn-HCl, guanidine hydrochloride; EGTA, [ethylenebis(oxyethylenenitrilo)tetraacetic acid; HEDTA, hydroxyethylethylenediaminetriacetic acid; NTA, nitrilotriacetic acid.

0006-2960/88/0427-7419$01.50/0

of Gdn-HCl was determined from the refractive index at 589 nm (Nozaki, 1972).

CD Measurements and Preparation of Protein Solutions. CD spectra were taken on a Jasco J-500A spectropolarimeter (Kuwajima et al., 1985). The path length of the optical cuvette was 1.0 mm in most cases, but a 10-mm or a 10-cm path length was also employed, depending on the spectral region and concentration of the protein. The CD data were represented as mean residue ellipticity by taking 106 as the mean residue weight (Coffee & Bradshaw, 1973). All measurements were carried out at 4.5 °C by circulating thermostated water through an optical-cell compartment (a NESLAB Model RTE-9 bath circulator).

The protein concentration was estimated spectrophotometrically by using a value of the extinction coefficient, $E_{1cm}^{1\%}$, at 259 nm of 1.8 (Konosu et al., 1965). Because of the absence of tryptophan residues, however, the absorption spectrum of parvalbumin is unusually small and was found to be easily affected by a difference in light scattering between sample and reference beams; a correction for the light scattering was made by the method of Kronman et al. (1972) for the purpose of estimating the concentration. Since the CD spectrum is less sensitive to the light-scattering effect, most of the CD data reported were restandardized by using the ellipticity at 222.5 nm of the native protein, 14 000 deg·cm^2/dmol at pH 7.0 in the absence of Gdn-HCl. The protein concentrations for CD measurements were typically 25 μM for measurements at neutral pH, while for measurements of the peptide spectra below 250 nm in the acid state, 6–10-fold more dilute solutions were employed for preventing aggregation of the protein molecule. Because parvalbumin binds Ca^{2+} at neutral pH, the concentration of free Ca^{2+} in the protein solution was regulated either by adding a large excess of $CaCl_2$ over the protein concentration or by using Ca^{2+}-chelating buffer agents, e.g., 10 mM [ethylenebis(oxyethylenenitrilo)]tetraacetic acid (EGTA) plus 1 mM $CaCl_2$, which results in a Ca^{2+} concentration of an order of 10^{-7} M (pH 7.0). For measurements in the acid state, the pH was adjusted to 2.5–2.8 by a dilute HCl solution. The pH values reported were pH-meter readings at room temperature.

Kinetic Measurements. Refolding and unfolding kinetics of parvalbumin were investigated by the stopped-flow method or by a method that utilized a magnetic stirring mixer (the magnetic stirring method), depending on the time interval where the reactions occurred (pH 7.0 and 4.5 °C). In both the methods, the refolding or unfolding reaction was initiated by concentration jump of Gdn-HCl. For the refolding experiment, the solution of parvalbumin unfolded in concentrated Gdn-HCl, typically of 4 M, was rapidly diluted with a buffer solution, and the subsequent changes in the CD spectra were observed. For the unfolding experiment, the inverse concentration jump of Gdn-HCl was performed by use of the protein solution in the absence of the denaturant and a diluent solution of concentrated Gdn-HCl. The path length of the optical cuvette used in the magnetic stirring method was 1.0 cm. The dead time of the mixing was reevaluated in the present study and estimated to be about 0.5 s. The other details of the magnetic stirring method have been described in previous papers (Kuwajima et al., 1985; Ikeguchi et al., 1986a). For the stopped-flow CD measurements, a mixing device comprising two driving syringes, a slit-type mixer, and a flat observation cell was attached to the spectropolarimeter (Kuwajima et al., 1987). The driving syringes worked pneumatically at 4 kg/cm^2 by the aid of a double-acting air cylinder connected with magnetic control valves. The mixing volume ratio of two solutions in the syringes was 1:10, the optical path of the cell was 1.0 mm, and the dead time of mixing was estimated to be 18 ms.

Calculation of Free Ca^{2+} Concentration. The equilibrium unfolding reaction and the refolding kinetics of the protein were measured at various concentrations of Ca^{2+} by use of Ca^{2+} buffer agents: EGTA, hydroxyethylethylenediaminetriacetic acid (HEDTA), and nitrilotriacetic acid (NTA). Because the present conditions, namely, 4.5 °C and high ionic strength because of the presence of Gdn-HCl, were necessarily different from the conditions usually found in the literature reporting the stability constants of the Ca^{2+}–chelator complexes, corrections were made for the stability constants to be used for evaluation of free Ca^{2+} concentration. The correction for a temperature difference could be made by use of the reported temperature dependence of the acidity constants of the chelator molecules and the temperature dependence of the stability constants of the Ca^{2+} complexes (Sillén & Martell, 1964, 1971). The correction for ionic strength, I, was made on the basis of a well-known theory of electrolyte solutions. The apparent stability constant, β_{app}, can be approximated as the equilibrium constant of the reaction

$$Ca^{2+} + H_mL^{n-} \overset{\beta_{app}}{\rightleftharpoons} CaL^{(m+n-2)-} + (mH^+) \qquad (1)$$

and

$$\beta_{app} = \frac{[CaL^{(m+n-2)-}]}{[Ca^{2+}][H_mL^{n-}]} \qquad (2)$$

where H_mL^{n-} denotes a chelator molecule that has n negative charges and m sites protonated at a given pH; m and n values for the individual chelators are $m = 2$ and $n = 2$ for EGTA and $m = 1$ and $n = 2$ for both HEDTA and NTA at pH 7.0. If it is assumed that the ionic activity coefficient, f_ν, for the ions that have a charge number ν is expressed, as a first approximation, by

$$\log f_\nu = -\nu^2\left(\frac{0.5I^{1/2}}{1 + I^{1/2}} - 0.1I\right) \qquad (3)$$

(Perrin & Dempsey, 1974), then the ionic strength dependence of β_{app} is given by the relationship

$$\log \beta_{app} = \log \beta_{app}^\circ + (m^2 + 2mn - 4m - 4n)\left(\frac{0.5I^{1/2}}{1 + I^{1/2}} - 0.1I\right) \qquad (4)$$

where β_{app}° is the apparent stability constant in an infinitely dilute solution. The validity of this relationship has been justified for the Ca^{2+} complexation of EGTA (Harafuji & Ogawa, 1980). Equation 4 was first used to estimate β_{app}° from the reported stability constants (Sillén & Martell, 1964, 1971) and then used to calculate the values of β_{app} under the present conditions. The free Ca^{2+} concentration was thus obtained by

$$-\log [Ca^{2+}] = \log \beta_{app} - \log [[CaL]/([L]_t - [CaL])] \qquad (5)$$

where [CaL] is the concentration of the Ca^{2+}–chelator complex and $[L]_t$ is the total concentration of the chelator, which was typically 5–10 mM in the present study. The values of $\log \beta_{app}^\circ$ for the above three chelators, used in eq 4, were 6.63 for EGTA, 6.17 for HEDTA, and 4.39 for NTA. The ionic strength used in investigating the Ca^{2+}-dependent refolding kinetics was 0.45–0.5 M, considering coexistent buffer ions and NaCl (0.35 M Gdn-HCl). Under this condition, eq 4 gave a good approximation; the refolding kinetics measured at the

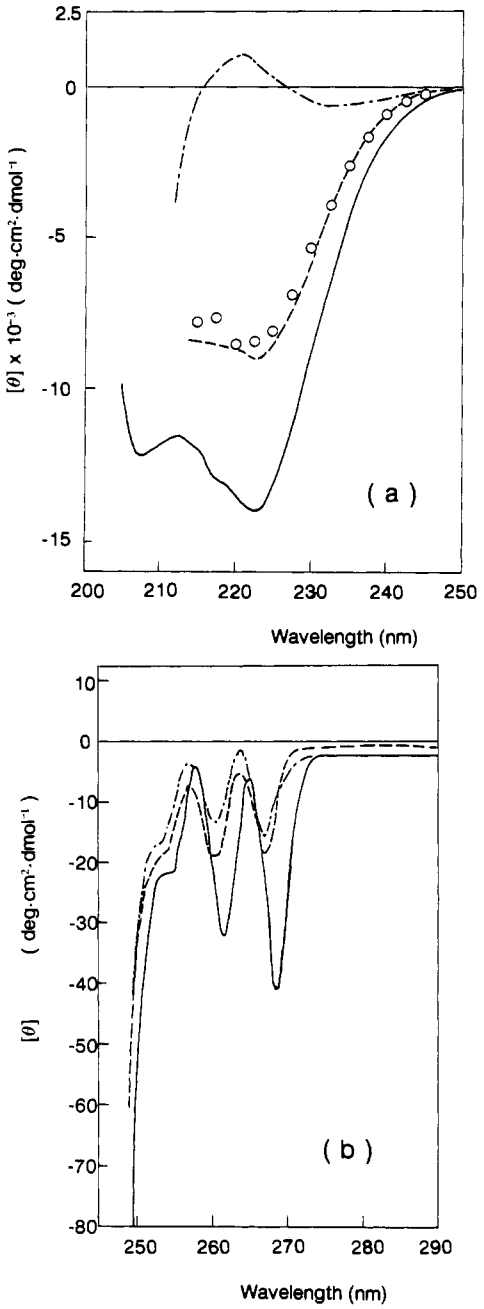


FIGURE 1: CD spectra of parvalbumin in the peptide (a) and aromatic (b) regions in the various conformational states (4.5 °C): (—) native state at pH 7.0 in 50 mM sodium cacodylate–50 mM NaCl; (- - -) acid state at pH 2.8 in 0.09 M NaCl (a) and at pH 2.5 in 50 mM sodium cacodylate–0.1 M NaCl (b); (–·–) unfolded state in 5 M Gdn-HCl and 50 mM sodium cacodylate–50 mM NaCl at pH 7.0. Open circules in (a) indicate the transient folding intermediate and correspond to the $\theta(0)$ values of refolding at 0.2 M Gdn-HCl at pH 7.0.

same calculated $[Ca^{2+}]$ adjusted with different chelators were found to be identical. The ionic strength used in investigating the Ca^{2+}-dependent equilibrium unfolding was at 1.6 M (1.5 M Gdn-HCl), and it was too high to obtain a reliable estimate of β_{app}; the error introduced in calculating log β_{app} may not exceed 1, however. Even in this case, relative concentration of free Ca^{2+} can be estimated by eq 5 as far as the concentrations adjusted with a single chelator, HEDTA in the present case, are concerned.

RESULTS

Equilibrium CD Spectra. Peptide and aromatic CD spectra of parvalbumin in the native state in the absence of Gdn-HCl

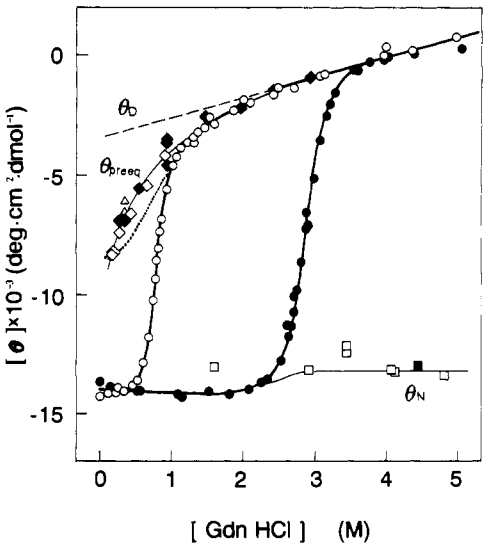


FIGURE 2: Unfolding transitions of parvalbumin measured by the ellipticity at 222.5 nm as a function of Gdn-HCl concentration at 4.5 °C. Thick solid lines show the equilibrium unfolding curves at pH 7.0 at 1 mM excess Ca^{2+} (●) and at 10 mM EGTA plus 1 mM $CaCl_2$ (○) in the presence of 50 mM sodium cacodylate and 50 mM NaCl. The dotted line is an equilibrium unfolding curve in the acid state (pH 2.5) in the presence of 50 mM sodium acetate and 50 mM NaCl. The broken line refers to the θ_D values obtained by linear extrapolation of the equilibrium unfolding curves. Thin solid lines are preequilibrium curves and show the dependence of the $\theta(0)$ on Gdn-HCl concentration (pH 7.0) (see text). The $\theta(0)$ values from the kinetic refolding curves (◇, △, ◆) give a preequilibrium unfolding curve (θ_{preeq}) of the transient folding intermediate, and the $\theta(0)$ values from the kinetic unfolding curves (□, ■) give the ellipticity values (θ_N) of the native state under the unfolding conditions. The filled symbols denote the data at 1 mM excess Ca^{2+} and the open symbols the data in the presence of chelators: (◇, □) at 10 mM EGTA plus 1 mM $CaCl_2$; (△) at 5 mM HEDTA plus 1–4 mM $CaCl_2$.

and in the unfolded state in 5 M Gdn-HCl are shown in Figure 1. In order to examine whether or not the spectra under the native condition are affected by free Ca^{2+} ions, the spectra at 0 M Gdn-HCl were taken at 1 mM free Ca^{2+} and at 10 mM EGTA plus 1 mM $CaCl_2$ ($[Ca^{2+}] = 0.7 \times 10^{-7}$ M). There was no difference detected between the spectra under the two conditions, indicating that in the absence of the denaturant the native structure of parvalbumin is kept intact at a concentration of Ca^{2+} as low as 10^{-7} M (pH 7.0 and 4.5 °C). In the presence of 5 M Gdn-HCl, both the peptide and the aromatic spectra are dramatically changed, and the protein has lost most of its ordered structures. The CD spectra in the acid state (pH 2.5–2.6) in the absence of Gdn-HCl are also shown in Figure 1. The aromatic spectrum is very similar to the spectrum of the unfolded molecule in Gdn-HCl, while the peptide spectrum shows only limited changes in the CD intensity. Therefore, the acid state is not completely unfolded but contains an appreciable amount of backbone helicity.

Equilibrium Unfolding Curves. The equilibrium unfolding transition of parvalbumin was followed by monitoring the ellipticity value at 222.5 nm, where the change in ellipticity caused by the unfolding is maximal, as a function of Gdn-HCl concentration. Full reversibility of the transition was demonstrated by dilution of the protein solutions in concentrated Gdn-HCl with a buffer solution at 0 M Gdn-HCl. The results at pH 7.0 at 1 mM excess Ca^{2+} and at 10 mM EGTA plus 1 mM $CaCl_2$ are shown in Figure 2. Although the native structure at 0 M Gdn-HCl has been shown to be kept intact irrespective of the difference in Ca^{2+} concentration, the stability of the native state is remarkably enhanced by the presence of 1 mM excess Ca^{2+}. The transition midpoint is

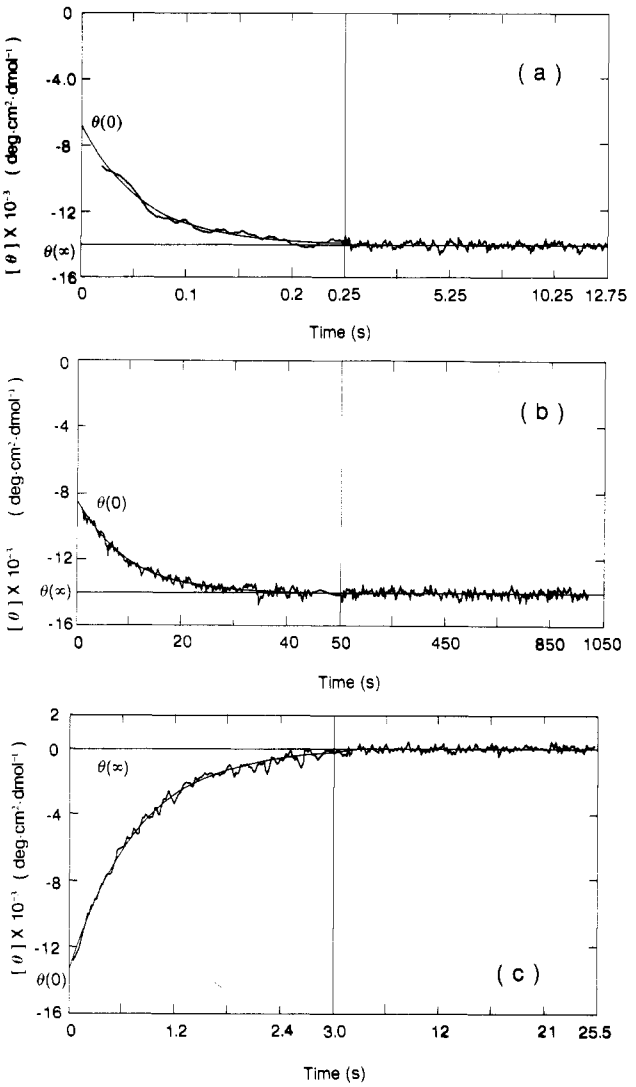


FIGURE 3: Kinetic progress curves of refolding [(a) and (b)] and unfolding (c) of parvalbumin measured by the ellipticity changes at 222.5 nm (pH 7.0 and 4.5 °C). Final conditions: (a) 0.34 M Gdn-HCl at 1 mM excess Ca^{2+} and 25 μM protein; (b) 0.19 M Gdn-HCl at 10 mM EGTA plus 1 mM $CaCl_2$ and 4.3 μM protein; (c) 4.19 M Gdn-HCl at 1 mM excess Ca^{2+} and 34 μM protein. All the solutions contained 50 mM sodium cacodylate and 50 mM NaCl. Thin solid lines in each figure show the $\theta(\infty)$ and the theoretical curve according to eq 6. The best-fit parameter values are (a) $\theta(0) = -6.8 \times 10^3$, $\theta(\infty) = -14.0 \times 10^3$ deg·cm²/dmol, and $k_{app} = 17.7\ s^{-1}$; (b) $\theta(0) = -8.4 \times 10^3$, $\theta(\infty) = -14.0 \times 10^3$ deg·cm²/dmol, and $k_{app} = 0.109\ s^{-1}$; and (c) $\theta(0) = -13.6 \times 10^3$, $\theta(\infty) = 0.0 \times 10^3$ deg·cm²/dmol, and $k_{app} = 1.32\ s^{-1}$.

located at 0.8 M Gdn-HCl when the solution contains 10 mM EGTA and at 2.9 M Gdn-HCl for the solution with 1 mM excess Ca^{2+}.

The unfolding transition of the protein in the acid state was also investigated, and the transition curve is shown by a dotted line in Figure 2. The transition is broader than that observed at pH 7.0, suggesting that some of the cooperative interactions that stabilize the native structure have been lost in the acid state.

Kinetic Refolding and Unfolding Curves. Parts a and b of Figure 3 show typical examples of kinetic progress curves for refolding of parvalbumin measured by the CD changes at 222.5 nm. The refolding reaction at 10 mM EGTA plus 1 mM $CaCl_2$ has a relaxation time longer than 9 s at a final concentration of Gdn-HCl higher than 0.17 M. The reaction was followed by the stopped-flow method and by the magnetic stirring method; the curve followed by the latter method is presented in Figure 3b. The protein concentration used in the magnetic stirring method was about 10-fold more dilute than the concentration used in the stopped-flow measurement, but the kinetic curves were the same in both measurements. In the presence of 1 mM excess Ca^{2+}, the refolding is remarkably accelerated and occurs in a time range as short as 0.1 s. Hence, it was followed only by the stopped-flow method. In all cases, the reaction kinetics are expressed by a single first-order equation

$$\theta(t) = \theta(0) + [\theta(\infty) - \theta(0)](1 - e^{-k_{app}t}) \qquad (6)$$

where $\theta(0)$, $\theta(t)$, and $\theta(\infty)$ are the ellipticity values at zero time, at time t, and at equilibrium, respectively, and k_{app} is the apparent rate constant of the reaction. The values of $\theta(0)$, $\theta(\infty)$, and k_{app} were obtained from the observed reaction curves by means of nonlinear least-squares curve fitting, and a correction for the ellipticity change that occurred within the dead time was made for estimating $\theta(0)$.

If the refolding occurs in a simple two-state mechanism, not only the kinetics is characterized by the single k_{app} but also the $\theta(0)$ must coincide with the value of the unfolded state. As shown in Figure 3, however, the $\theta(0)$ values (−6000 to −8500 deg·cm²/dmol) are much more negative than the values expected from the equilibrium unfolding curves (Figure 2). The CD intensity of $\theta(0)$ amounts to 50–60% of the intensity in the native state. Therefore, it is concluded that at an early stage of the refolding of parvalbumin a transient intermediate that has an appreciable amount of backbone secondary structure forms rapidly within the dead time of the stopped-flow mixing.

A reaction curve for unfolding of the protein observed at 1 mM excess Ca^{2+} is shown in Figure 3c, and a similar result was obtained at 10 mM EGTA plus 1 mM $CaCl_2$ (4.2 M Gdn-HCl). The relaxation time of the reaction is 0.8 s, and the reaction was followed by the stopped-flow method. Also in this case, the total reaction curve is expressed by single first-order kinetics. In contrast to the results of refolding, the $\theta(0)$ value is close to the value in the native state. More than 90% of the difference in ellipticity between the native and the unfolded states was observed kinetically.

It was also found that when the refolding and unfolding reactions were induced in the unfolding transition zone with the same final concentration of Gdn-HCl, the k_{app} values in both reactions were identical.

Structure and Stability of Transient Folding Intermediate. In order to investigate the secondary structure fractions in the transient intermediate of refolding and to compare them with the fractions in the equilibrium states, the refolding reaction curve at 0.2 M Gdn-HCl was measured at various wavelengths. The refolding rate was found to be independent of wavelength, and the wavelength dependence of $\theta(0)$ gives a CD spectrum of the folding intermediate. The resultant CD spectrum is compared with the equilibrium spectra in Figure 1. The secondary structure fractions for the equilibrium states and for the transient folding intermediate were estimated from these CD spectra by the methods of Greenfield and Fasman (1969), Saxena and Wetlaufer (1971), and Chen et al. (1974). When the secondary structure fractions of the different states are compared in this way, it is particularly important whether the same wavelength region is chosen in the calculations (Kuwajima et al., 1985), and hence the calculations were made by using the data in a wavelength region from 215 to 250 nm. The structural fractions in the intermediate state were found to agree well with the fractions in the acid state, and the α-helix contents in these states are 60–80% of the value in the native state; the helix content in the native state was estimated

to be 33–46%, depending on the method employed.

The stability of the secondary structure formed in the intermediate was investigated by measuring the kinetic refolding curves at various concentrations of Gdn-HCl. Although the k_{app} of refolding was found to decrease sharply with increasing Gdn-HCl concentration (Figure 5), the single kinetics were observed at any concentration of the denaturant. The preequilibrium between the unfolded state and the intermediate is attained within the dead time under all conditions investigated. The dependence of $\theta(0)$ on the concentration of Gdn-HCl, therefore, represents a preequilibrium unfolding curve of the transient intermediate [see Ikeguchi et al. (1986a)]. Such a transition curve is also shown in Figure 2 and compared with the equilibrium unfolding curves. The following results are manifested. (i) The $\theta(0)$ values obtained from the kinetic refolding curves at 1 mM excess Ca^{2+} fall on the same transition curve as drawn from the $\theta(0)$ values at 10 mM EGTA plus 1 mM $CaCl_2$. The preequilibrium between the unfolded state and the intermediate is not affected by the presence of excess Ca^{2+}. (ii) The transition curve of the folding intermediate is less cooperative than the equilibrium unfolding transitions between the native and the unfolded states, and there is similarity between the intermediate and the acid state, although the latter state is a little more stable than the intermediate.

Analysis of Equilibrium Unfolding. In a traditional analysis of the equilibrium unfolding curves, it is often assumed that the unfolding transition is a simple two-state reaction between the native (N) and the fully unfolded (D) states and that the ellipticity values of each pure state in the base-line region before or after the transition can be extrapolated linearly to the transition zone. However, the assumptions fail to interpret the unfolding transition of parvalbumin at 10 mM EGTA plus 1 mM $CaCl_2$. The equilibrium unfolding curve between 1 and 2 M Gdn-HCl shows a deviation from the linear extrapolation in ellipticity of the D state, although the major cooperative transition does not take place in this region (Figure 2). The preequilibrium unfolding curve of the transient intermediate coincides with the equilibrium unfolding curve above 1 M Gdn-HCl. These demonstrate that the equilibrium curve is represented as a composite of at least two transitions between the N and the intermediate and between the intermediate and the D states. Therefore, the total unfolding equilibrium involves at least three species

$$D \rightleftharpoons I \rightleftharpoons N \tag{7}$$

where I denotes the intermediate that appears as the transient folding intermediate in the kinetic refolding, but its equilibrium population is not negligible when the N state is not sufficiently stable compared with the I state. The transition between N and I is more cooperative than the I $\rightleftharpoons$ D transition, only the N $\rightleftharpoons$ I transition is observed kinetically and the I $\rightleftharpoons$ D reaction is too rapid to measure by the stopped-flow method. The apparent equilibrium constant is thus associated with the first-order process observed in the kinetic experiments and can be calculated from the equilibrium unfolding curve (from N to D) and the preequilibrium unfolding curve (from I to D)

$$K_{app} = \frac{[I] + [D]}{[N]} = \frac{f_I + f_D}{f_N} \tag{8}$$

where f_N, f_I, and f_D are the fractions of the three states at equilibrium and $f_N + f_I + f_D = 1$. The ellipticity value at equilibrium, θ_{equil}, relates to the ellipticity value of each pure state as

$$\theta_{equil} = \theta_N f_N + \theta_I f_I + \theta_D f_D \tag{9}$$

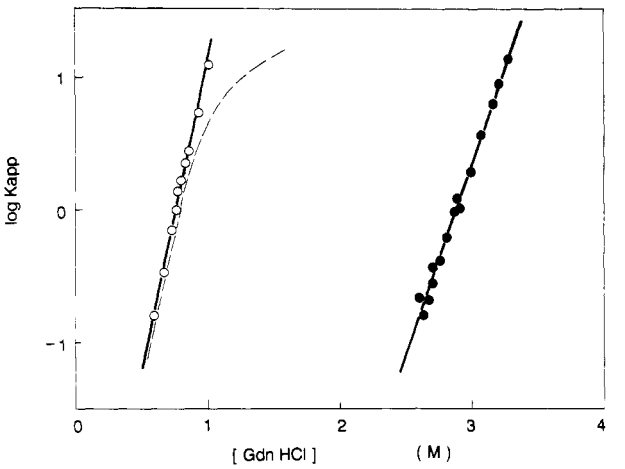


FIGURE 4: Dependence of the logarithmic K_{app} of parvalbumin on Gdn-HCl concentration at 1 mM excess Ca^{2+} (●) and at 10 mM EGTA plus 1 mM $CaCl_2$ (○) (pH 7.0 and 4.5 °C). A broken line indicates the values obtained by assuming a two-state model between N and D and a linear extrapolation of θ_D (Figure 2).

and the ellipticity value at preequilibrium between I and D is represented by

$$\theta_{preeq} = \theta_I\left(\frac{f_I}{f_I + f_D}\right) + \theta_D\left(\frac{f_D}{f_I + f_D}\right) \tag{10}$$

Therefore

$$K_{app} = \frac{\theta_N - \theta_{equil}}{\theta_{equil} - \theta_{preeq}} \tag{11}$$

In order to calculate the K_{app} values by this equation, the θ_N values in the transition zone have also to be estimated. Thin solid lines in Figure 2 show the θ_N and θ_{preeq} values used for the calculation. The former values are assumed to be the $\theta(0)$ values of the kinetic unfolding curves. The $\theta(0)$ values observed were always a little larger than the values expected from the linear extrapolation in ellipticity of the N state, but this difference was less than 10% of the ellipticity difference between N and D. Thus, whichever values are used as θ_N, the results are not affected significantly. The K_{app} values are shown in Figure 4 as a function of Gdn-HCl concentration. The logarithmic K_{app} is found to linearly depend on the denaturant concentration, and this is a common feature observed in cooperative protein unfolding induced by the denaturant (Schellman, 1978; Pace, 1986). If the two-state reaction between N and D and the linear extrapolation of θ_D are assumed, the plot of log K_{app} vs Gdn-HCl concentration shows a significant deviation from the linearity. By assuming the linear relationship, the Gibbs energy changes of unfolding at 0 M Gdn-HCl could be estimated, and they are 20 kJ/mol at 10 mM EGTA plus 1 mM $CaCl_2$ and 44 kJ/mol at 1 mM excess Ca^{2+}.

Analysis of Refolding and Unfolding Rate Constants. The rate constant k_{app} has been calculated from the kinetic refolding and unfolding curves at various concentrations of Gdn-HCl. Because the rapid preequilibrium of the I $\rightleftharpoons$ D reaction is attained under any conditions of the refolding, the rate constants for the folding and unfolding directions (k_{fold} and k_{unf}, respectively) can be defined and relate to k_{app} and K_{app} as

$$k_{fold} = \frac{k_{app}}{1 + K_{app}} \tag{12}$$

and

$$k_{unf} = k_{fold}K_{app} \tag{13}$$

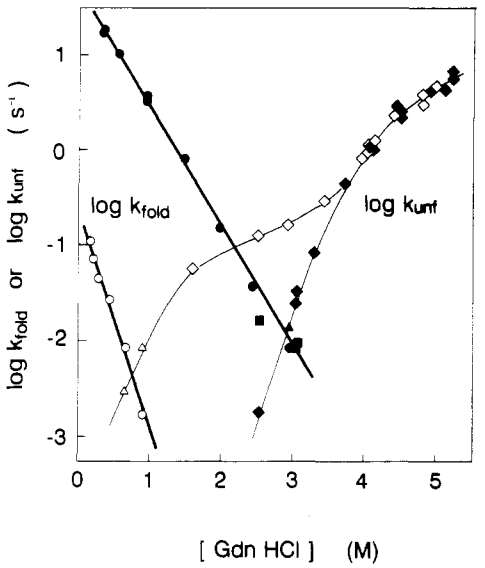


FIGURE 5: Dependence of the logarithmic k_{fold} (○, ●, ■) and k_{unf} (△, ▲, ◇, ◆) of parvalbumin on Gdn-HCl concentration (pH 7.0 and 4.5 °C): (filled symbols) data at 1 mM excess Ca^{2+}; (open symbols) data at 10 mM EGTA plus 1 mM $CaCl_2$. Circles and triangles denote the data obtained from the kinetic refolding curves, and squares and diamonds are those from the unfolding curves.

On the basis of the three-state mechanism, the simplest reaction scheme is written by

$$D \underset{\text{rapid}}{\overset{K_{DI}}{\rightleftharpoons}} I \overset{k_+}{\underset{k_-}{\rightleftharpoons}} N \qquad (14)$$

preequil

Thus, k_{fold} and k_{unf} can also be expressed in terms of the rate constants k_+ and k_- and the equilibrium constant of the D ⇌ I reaction, K_{DI}, as

$$k_{fold} = \frac{K_{DI}k_+}{1 + K_{DI}} \qquad (15)$$

and

$$k_{unf} = k_- \qquad (16)$$

Although a more complicated reaction scheme may also be required for a complete description of the real folding process [see Ikeguchi et al. (1986a)], the simplified treatment does not harm any conclusions drawn in the present study.

The k_{fold} and k_{unf} were calculated by eq 12 and 13 from the observed k_{app} and from the K_{app} values evaluated by assuming the linear relationship between log K_{app} and Gdn-HCl concentration (Figure 4). The resulting k_{fold} and k_{unf} values are shown in Figure 5. The logarithmic k_{fold} decreases linearly with an increase in Gdn-HCl concentration. The logarithmic k_{unf} increases linearly with an increase in the denaturant concentration in the transition zones of the equilibrium unfolding. However, in the base-line unfolded regions, above 1.5 M Gdn-HCl in 10 mM EGTA plus 1 mM $CaCl_2$ and above 3.6 M Gdn-HCl in the presence of 1 mM excess Ca^{2+} (Figure 2), the linear relationships are broken, and at a high concentration of Gdn-HCl (>4 M), the k_{unf} values at 1 mM excess Ca^{2+} are the same as the values in the presence of EGTA. Thus, the unfolding reaction under this condition might occur through a pathway different from the folding ⇌ unfolding pathway at a low concentration of Gdn-HCl.

Effect of Ca^{2+} on the Unfolding Equilibrium. At 1.5 M Gdn-HCl, parvalbumin is in the N state at 1 mM excess Ca^{2+}, and it is unfolded at 10 mM EGTA plus 1 mM $CaCl_2$ (Figure 2). The effect of Ca^{2+} on the unfolding equilibrium at 1.5 M Gdn-HCl was investigated by measuring the ellipticity values at 222.5 nm at various concentrations of free Ca^{2+} regulated by HEDTA. The K_{app} values have been calculated by eq 11, in which only θ_{equil} is dependent on Ca^{2+} concentration. The results are shown in Figure 6. The logarithmic K_{app} is found to linearly depend on the logarithm of Ca^{2+} concentration, and the slope of the plot gives the number of Ca^{2+} ions that have been released from the protein during the equilibrium unfolding (see below). The result is consistent with the fact that native parvalbumin has two Ca^{2+}-binding sites, both of which have similar affinities for Ca^{2+} (Kretsinger, 1980; Moeschler et al., 1980). Because the slope of the above plot is close to −2, both of the binding sites are disrupted during the cooperative unfolding of the protein molecule. The Ca^{2+}-dependent unfolding of the protein that has two independent Ca^{2+}-binding sites with equal affinities in the N state can be expressed by the equation (Schellman, 1975)

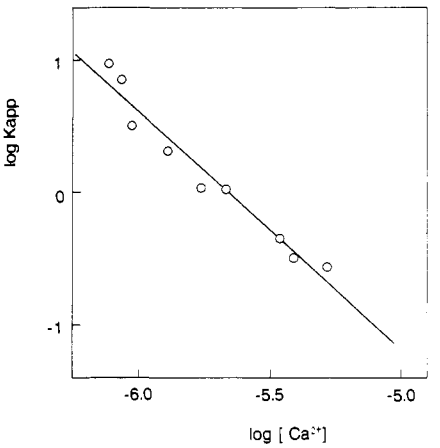
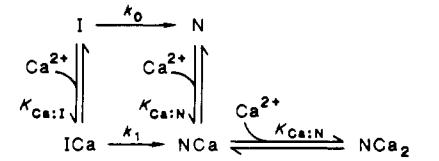


FIGURE 6: Effect of free Ca^{2+} concentration on the unfolding equilibrium of parvalbumin at 1.5 M Gdn-HCl (pH 7.0 and 4.5 °C). The Ca^{2+} concentrations were adjusted with 5 mM HEDTA and various concentrations of $CaCl_2$.

$$K_{app} = \frac{K_{app}^{0}}{(1 + K_{Ca:N}[Ca^{2+}])^2} \qquad (17)$$

where K_{app}^{0} refers to the K_{app} in the absence of Ca^{2+} and $K_{Ca:N}$ is the binding constant in the N state. In eq 17, it is also assumed that the D and I states have negligible affinity for Ca^{2+} in the Ca^{2+} concentration range studied ($[Ca^{2+}] < 10^{-5}$ M).

Effect of Ca^{2+} on the Refolding Kinetics. The kinetic refolding curves of parvalbumin at 0.35 M Gdn-HCl were measured at various concentrations of free Ca^{2+}. Under these conditions, the refolding goes to completion and only single-phase kinetics are observed at any concentration of Ca^{2+} employed. Thus, the observed k_{app} is practically identical with k_{fold}. The logarithmic k_{fold} is plotted as a function of the logarithm of Ca^{2+} concentration in Figure 7. The plot shows a sigmoidal dependence of log k_{fold} on log $[Ca^{2+}]$ and gives a slope of 1 in the middle. This suggests that the rate-limiting step of the refolding is associated with the binding of one Ca^{2+} ion to the protein. Because the preequilibrium D ⇌ I reaction is not affected by Ca^{2+} as inferred from the preequilibrium unfolding curve in Figure 2, only the k_+ part of k_{fold} (eq 15) is dependent on Ca^{2+} concentration, and only the folding process from the I state will be considered. The results of Figure 7 are interpreted by the reaction scheme



(18)

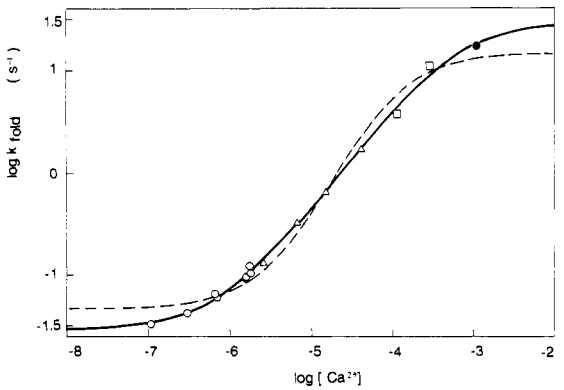


FIGURE 7: Effect of free Ca^{2+} concentration on the rate (k_{fold}) of the refolding of parvalbumin induced by a concentration jump of Gdn-HCl from 4 to 0.35 M (pH 7.0 and 4.5 °C). The free Ca^{2+} concentration was adjusted by the chelators EGTA (○), HEDTA (△), and NTA (□). A filled circle denotes the value at 1 mM excess Ca^{2+} without chelators. A solid line is the theoretical curve according to eq 20; $K_{Ca:I} = 1.50 \times 10^3$ M^{-1} and $K_{Ca:A} = 1.44 \times 10^6$ M^{-1}. A broken line is the curve according to eq 21; $n = 2$, $K_{Ca:I} = 1.32 \times 10^4$ M^{-1}, and $K_{Ca:A} = 2.34 \times 10^5$ M^{-1}.

where $K_{Ca:I}$ is the binding constant of Ca^{2+} in the I state and k_1 and k_0 are the rate constants of refolding from the I state with and without the bound Ca^{2+}, respectively. According to eq 18, the refolding is accelerated with increasing Ca^{2+} concentration when k_1 is larger than k_0. Since in the above scheme the Ca^{2+} binding to each pure state occurs rapidly compared with the folding process, k_{fold} can be written by

$$k_{fold} = k_0'\left(\frac{1 + (k_1/k_0)K_{Ca:I}[Ca^{2+}]}{1 + K_{Ca:I}[Ca^{2+}]}\right) \quad (19)$$

where k_0' is a constant under the present condition and corresponds to $K_{DI}k_0/(1 + K_{DI})$ (eq 15); K_{DI} does not depend on Ca^{2+} concentration. Because we can assume that the refolding reaction occurs through the critical activated state that exists on the reaction pathway from I to N, the acceleration of refolding by the bound Ca^{2+}, expressed by eq 19, can be interpreted in terms of stabilization of the activated state by the bound Ca^{2+}. The activated state can thus also bind Ca^{2+}. Considering the Ca^{2+}-dependent equilibrium between I and the activated state and using Schellman's treatment (Schellman, 1975), k_{fold} is related to the binding constant $K_{Ca:I}$ and the binding constant in the activated state $K_{Ca:A}$ as

$$k_{fold} = k_0'\left(\frac{1 + K_{Ca:A}[Ca^{2+}]}{1 + K_{Ca:I}[Ca^{2+}]}\right) \quad (20)$$

Therefore, $(k_1/k_0)K_{Ca:I}$ in eq 19 corresponds to $K_{Ca:A}$. The observed dependence of k_{fold} on Ca^{2+} concentration has been fitted to eq 20 by the nonlinear least-squares method, and the best-fit values of $K_{Ca:I}$ and $K_{Ca:A}$ are $K_{Ca:I} = 1.50 \times 10^3$ M^{-1} and $K_{Ca:A} = 1.44 \times 10^6$ M^{-1}. A solid line in Figure 7 is the theoretical curve drawn on the basis of these values of the parameters. In order to test the uniqueness of the fit to the present model, the observed dependence of k_{fold} has also been fitted to a more general equation that allows binding of n Ca^{2+} ions:

$$k_{fold} = k_0'\left(\frac{1 + K_{Ca:A}[Ca^{2+}]}{1 + K_{Ca:I}[Ca^{2+}]}\right)^n \quad (21)$$

The best-fit theoretical curve with $n = 2$ in eq 21 is shown by a broken line in Figure 7, and the agreement with the experimental data is clearly less satisfactory. The results demonstrate that only one of the two Ca^{2+}-binding sites in parvalbumin is restored during the activation process of refolding and has a binding constant as high as 10^6 M^{-1}. Because of such a high value of $K_{Ca:A}$, some specific structure responsible for the Ca^{2+} binding must be organized in the activated state, but no such structure is stable as yet in the I state. Although the I state also has affinity for Ca^{2+}, the binding constant is less than $K_{Ca:A}$ by a factor of 1000. Because the I state is not stabilized against the D state at a concentration of Ca^{2+} up to 1 mM (Figure 2), a similar order of magnitude in the binding constant must also be attributed to the D state. Derancourt et al. (1978) have reported that a 33-residue fragment (residues 76–108) of one Ca^{2+}-binding domain of parvalbumin has a binding constant of 300 M^{-1}, although the fragment may be mostly unfolded in the absence of sufficient Ca^{2+} and hence the Ca^{2+} binding in the I or D state might not be surprising.

DISCUSSION

In the present study, the equilibrium and kinetics of the Gdn-HCl-induced unfolding of parvalbumin can be fully interpreted in terms of a three-state reaction scheme in which a highly structured intermediate, I, is populated at a low concentration of Gdn-HCl, except unusual behavior of k_{unf} at a high concentration of Gdn-HCl (Figure 5). The population of the intermediate has been demonstrated from the preequilibrium unfolding curve obtained from kinetic refolding curves, and the I state acts as a transient folding intermediate under a native condition. The kinetics of refolding from the I to the N state has been characterized in terms of acquirement of the Ca^{2+}-binding capacity. Because the reaction between I and D is too rapid to measure, the present reaction scheme can be reduced to a two-state reaction by assuming that both I and D belong to the same species and by regarding the preequilibrium unfolding as a change due to a solvent effect. However, the markedly ordered secondary structure of the intermediate evidenced by the transient CD spectra (Figure 1a) makes the two-state scheme unlikely. Moreover, close analogy of the present results with the previous results of α-lactalbumin, for which the three-state unfolding when treated with Gdn-HCl has been well established, demonstrates that the three-state scheme is much more plausible. The mechanism of folding of parvalbumin will be discussed on the basis of these results, and comparison with the studies of α-lactalbumin will thus be made.

Intermediate in Parvalbumin. The I state has the following properties: (i) It has 60–80% of the α-helix in the native protein (0.2 M Gdn-HCl, pH 7.0 and 4.5 °C). (ii) It does not have the specific structure responsible for the strong Ca^{2+} binding in the N state. (iii) The I ⇌ D transition is less cooperative than the transition from the N state. (iv) The I state is similar to the acid state in the CD spectra (Figure 1) and in the Gdn-HCl-induced transitions (Figure 2).

Recently, similar transient folding intermediates have been reported in other small globular proteins, and these include α-lactalbumin, lysozyme, carbonic anhydrase, ribonuclease, β-lactamase, ferricytochrome *c*, β-lactoglobulin, tryptophan synthetase α-subunit, and growth hormone (Kuwajima, 1977; Kuwajima et al., 1985, 1987; Ptitsyn, 1987; McCoy et al., 1980; Dolgikh et al., 1984; Schmid & Baldwin, 1979; Kim, 1986; Brems & Baldwin, 1985; Mitchinson & Pain, 1985; Beasty & Matthews, 1985; Brems et al., 1987). Because most of these proteins are different from each other in both structure and function in the native state, accumulation of the transient intermediate is not restricted to a particular group of proteins.

Among the above proteins, the transient intermediate of α-lactalbumin has been well characterized (Kuwajima et al., 1985; Ikeguchi et al., 1986a,b; Gilmanshin & Ptitsyn, 1987). α-Lactalbumin is also known to be a Ca^{2+}-binding protein, having a single Ca^{2+}-binding site with an apparent binding constant of an order of 10^7–10^9 M^{-1} (Hiraoka et al., 1980; Segawa & Sugai, 1983; Hamano et al., 1986; Mitani et al., 1986; Permyakov et al., 1987a). The reported properties of the intermediate in α-lactalbumin are closely analogous to the properties in parvalbumin. The intermediate of α-lactalbumin assumes native-like secondary structure without specific tertiary structure and without the strong affinity for Ca^{2+}. The intermediate has been shown to be essentially the same as the partially unfolded species observed at acid and alkaline pH's or at an increased temperature (Kuwajima et al., 1981, 1985; Ikeguchi et al., 1986a). The relatively high stability of the intermediate in α-lactalbumin allowed us to detect the same state as a stable equilibrium species in Gdn-HCl-induced unfolding, and the three-state scheme of the equilibrium unfolding in α-lactalbumin has been well documented (Kuwajima et al., 1976, 1981; Ikeguchi et al., 1986a,b).

It has been demonstrated that the secondary structure in the transient intermediate is stabilized primarily by local interactions and that long-range interactions are concerned with the late stages of folding from the I to the N state (Kuwajima, 1977; Kuwajima et al., 1985; Ikeguchi et al., 1986a; Ptitsyn, 1987). The native structure of a protein is only marginally stable and maintained by a cooperative network of both types of interactions. Disruption of integrity of the cooperative network, by changing pH or by increasing temperature, thus leads to a conformational state that is essentially the same as the transient intermediate under a condition where the local interactions are still operative. Therefore, similarity between the transient intermediate and the partially unfolded species at equilibrium may have a general significance as a characteristic of globular proteins.

Removal of the two tightly bound Ca^{2+} ions from parvalbumin is known to bring about a partially unfolded species (Parello et al., 1974; Donato & Martin, 1974; Cox et al., 1979; Cavé et al., 1979; Williams et al., 1986), and this species may also fall under the same category of the intermediate. The conformational change induced by the Ca^{2+} removal occurs at room temperature in some cases or at a higher temperature (e.g., 45 °C) in other cases, depending on a species difference of parvalbumin or a difference in the solution condition. The partially unfolded apoparvalbumin contains 50–80% of the native helicity as indicated by the CD and ORD spectra (Donato & Martin, 1974; Cox et al., 1979), but the specific tertiary structure measured by the NMR and the aromatic CD spectra are almost completely disrupted (Cavé et al., 1979; Williams et al., 1986).

Stepwise Organization of Ca^{2+}-Binding Domains in the Kinetic Folding of Parvalbumin. According to the X-ray structure, native parvalbumin consists of three helix–loop–helix substructures called AB, CD, and EF domains, and two Ca^{2+} ions are bound by the loops in the CD and EF domains (Kretsinger & Nockolds, 1973). The three substructures are tightly packed together, forming a single hydrophobic core, and the overall structure of the protein is better represented as a single cooperative unit. In accord with these features, equilibrium unfolding of the protein is known to exhibit a cooperative transition without an indication of independent domains (Filimonov et al., 1978; Lin & Brandts, 1978; Iio & Hoshihara, 1984; Williams et al., 1986). At this point, however, it is an important question whether or not the refolding kinetics from the I state is also represented by simultaneous organization of the substructures. If the two Ca^{2+}-binding domains are organized simultaneously, the exponent factor n of eq 21 takes a value of 2 or 0, depending on location of the activated state along the reaction pathway. If the two domains are organized in a stepwise manner, in which the activated state has restored only one of the two Ca^{2+}-binding sites, the n value of eq 21 equals 1, and both k_{fold} and k_{unf} should depend on Ca^{2+} concentration. The results in the present study clearly show that the latter case is true in parvalbumin. The Ca^{2+} concentration dependence of k_{unf} under a condition where the refolding occurs is expected from Figure 5, in which k_{unf} at 1 mM excess Ca^{2+} is estimated to be 3 orders of magnitude smaller than the value at 10 mM EGTA plus 1 mM $CaCl_2$.

The results in Figure 6 demonstrate that the release of two bound Ca^{2+} ions from parvalbumin occurs cooperatively during equilibrium unfolding, but this does not necessarily mean that the Ca^{2+}-binding process itself is cooperative. Under the condition of Figure 6, the native state is sufficiently stabilized only with saturation of the two binding sites, and no partially filled intermediate is populated. The cooperativity of Ca^{2+} binding in this sense has been reported previously (Parello et al., 1974; Cavé et al., 1979). In contrast to this apparent cooperativity, the metal-ion binding to native apoparvalbumin, which may take place with no global conformational rearrangement, is known to show no significant cooperativity; the apoprotein under a strongly native condition (e.g., at 4.5 °C in the presence of a monovalent metal cation, as in the present study) assumes the native structure (Williams et al., 1986). The kinetic studies of metal-ion dissociation in native parvalbumin (Corson et al., 1983; Breen et al., 1985; Permyakov et al., 1987b) and the ^1H and ^{113}Cd NMR studies of the protein (Williams et al., 1986; Svärd & Drakenberg, 1986) have shown that the two metal-ion sites are independent of each other and that partially metal-filled intermediates are present at a metal ion to protein ratio less than 2. The noncooperative Ca^{2+} binding in native parvalbumin may be consistent with the stepwise organization of the Ca^{2+}-binding domains during the kinetic refolding. Because the Ca^{2+} binding to one of the two sites is independent of the state of the other site, the organization of only one of the two domains can lead to such a high affinity of the site for Ca^{2+} as to be compared with the affinity of the corresponding site in the native molecule. The partially organized structure is thus remarkably stabilized by Ca^{2+} binding and can act as a trigger of the refolding process from the I state. Iio and Hoshihara (1984) have studied the kinetics of conformational transitions of dansylated parvalbumin induced by Ca^{2+} binding and removal by means of stopped-flow fluorescence measurements. Their results are fully consistent with the above view and have shown that both refolding and unfolding rate constants depend on Ca^{2+} concentration, although the equilibrium transition is accompanied by cooperative uptake of Ca^{2+} ions. Because their measurements were made in the absence of Gdn-HCl, their observed conformational transition corresponds to the $N \rightleftharpoons I$ transition in the present study.

Conclusions

From the results above, the total folding reaction of parvalbumin can be divided into the three stages: (i) the formation of the transient intermediate that has local helicity; (ii) the organization of a part of the specific tertiary structure including one of the two Ca^{2+}-binding domains, with this step leading to the activated state of folding; and (iii) the final stabilization associated with the organization of the rest of the molecule including the other Ca^{2+}-binding domain.

There are two working models of protein folding (Kim & Baldwin, 1982): (i) the *framework model*, in which the local secondary structure is formed early in folding, and (ii) the *modular assembly model*, in which essentially complete folding of any part of a protein occurs at one time, although different parts of the protein fold at different times (folding by parts). These two models are not mutually exclusive, and the real folding may be represented by their combination. The present results show that the I state at the first stage of folding in parvalbumin is consistent with the framework model and that the late stages of folding may provide an example of the modular assembly.

Finally, we should mention a model proposed by Lin and Brandts (1978), who have demonstrated the presence of two different D states in parvalbumin caused by an isomerization reaction, other than proline isomerization, of the unfolded molecule. Their model was based on the observations of biphasic refolding kinetics under a native condition but single-phase unfolding kinetics in a base-line unfolded region when the reactions were followed by fluorescence spectra of phenylalanyl residues, and it was assumed that both the reactions occurred along the same pathway. However, as shown in the present study, the pathway taken by the unfolding at a high concentration of Gdn-HCl might be different from the pathway of refolding. The refolding kinetics observed in the present study corresponds to the fast phase observed by Lin and Brandts when a correction is made for a temperature difference. Thus, the slow phase observed by these authors is ascribed to a minor change in environment of some phenylalanyl residue(s), which is detectable by fluorescence but silent in the peptide CD kinetics.

ACKNOWLEDGMENTS

We are grateful to Dr. K. Nitta and M. Ikeguchi for their critical readings of the manuscript.

Registry No. Ca, 7440-70-2.

REFERENCES

Beasty, A. M., & Matthews, C. R. (1985) *Biochemistry 24*, 3547–3553.

Breen, P. J., Johnson, K. A., & Horrocks, W. D., Jr. (1985) *Biochemistry 24*, 4997–5004.

Brems, D. N., & Baldwin, R. L. (1985) *Biochemistry 24*, 1689–1693.

Brems, D. N., Plaisted, S. M., Dougherty, J. J., Jr., & Holzman, T. F. (1987) *J. Biol. Chem. 262*, 2590–2596.

Cavé, A., Pages, M., Morin, Ph., & Dobson, C. M. (1979) *Biochimie 61*, 607–613.

Chen, Y. H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350–3359.

Coffee, C. J., & Bradshaw, R. A. (1973) *J. Biol. Chem. 248*, 3305–3312.

Corson, D. C., Lee, L., McQuaid, G. A., & Sykes, B. D. (1983) *Can. J. Biochem. Cell Biol. 61*, 860–867.

Cox, J. A., Winge, D. R., & Stein, E. A. (1979) *Biochimie 61*, 601–605.

Craig, S., Hollecker, M., Creighton, T. E., & Pain, R. H. (1985) *J. Mol. Biol. 185*, 681–687.

Creighton, T. E. (1980) *J. Mol. Biol. 137*, 61–80.

Derancourt, J., Haiech, J., & Pechère, J.-F. (1978) *Biochim. Biophys. Acta 532*, 373–375.

Dolgikh, D. A., Kolomiets, A. P., Bolotina, I. A., & Ptitsyn, O. B. (1984) *FEBS Lett. 165*, 88–92.

Donato, H., Jr., & Martin, R. B. (1974) *Biochemistry 13*, 4575–4579.

Filimonov, V. V., Pfeil, W., Tsalkova, T. N., & Privalov, P. L. (1978) *Biophys. Chem. 8*, 117–122.

Gilmanshin, R. I., & Ptitsyn, O. B. (1987) *FEBS Lett. 223*, 327–329.

Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4116.

Hamano, M., Nitta, K., Kuwajima, K., & Sugai, S. (1986) *J. Biochem. (Tokyo) 100*, 1617–1622.

Harafuji, H., & Ogawa, Y. (1980) *J. Biochem. (Tokyo) 87*, 1305–1312.

Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S., & Murai, N. (1980) *Biochem. Biophys. Res. Commun. 95*, 1098–1104.

Iio, T., & Hoshihara, Y. (1984) *J. Biochem. (Tokyo) 96*, 321–328.

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

Ikeguchi, M., Kuwajima, K., & Sugai, S. (1986b) *J. Biochem. (Tokyo) 99*, 1191–1201.

Kim, P. S. (1986) *Methods Enzymol. 131*, 136–156.

Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–489.

Konosu, S., Hamoir, G., & Pechère, J.-F. (1965) *Biochem. J. 96*, 98–112.

Kretsinger, R. H. (1980) *CRC Crit. Rev. Biochem. 8*, 119–174.

Kretsinger, R. H., & Nockolds, C. E. (1973) *J. Biol. Chem. 248*, 3313–3326.

Kronman, M. J., Jeroszko, J., & Sage, G. W. (1972) *Biochim. Biophys. Acta 285*, 145–166.

Kuwajima, K. (1977) *J. Mol. Biol. 114*, 241–258.

Kuwajima, K., Nitta, K., Yoneyama, M., & Sugai, S. (1976) *J. Mol. Biol. 106*, 359–373.

Kuwajima, K., Ogawa, Y., & Sugai, S. (1981) *J. Biochem. (Tokyo) 89*, 759–770.

Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874–881.

Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S., & Nagamura, T. (1987) *FEBS Lett. 221*, 115–118.

Labhardt, A. M. (1986) *Methods Enzymol. 131*, 126–135.

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

McCoy, L. F., Jr., Rowe, E. S., & Wong, K.-P. (1980) *Biochemistry 19*, 4738–4743.

Mitani, M., Harushima, Y., Kuwajima, K., Ikeguchi, M., & Sugai, S. (1986) *J. Biol. Chem. 261*, 8824–8829.

Mitchinson, C., & Pain, R. H. (1985) *J. Mol. Biol. 184*, 331–342.

Moeschler, H. J., Schaer, J.-J., & Cox, J. A. (1980) *Eur. J. Biochem. 111*, 73–78.

Nozaki, Y. (1972) *Methods Enzymol. 26*, 43–50.

Pace, C. N. (1986) *Methods Enzymol. 131*, 266–280.

Parello, J., Cavé, A., Puigdomenech, P., Maury, C., Capony, J.-P., & Pechère, J.-F. (1974) *Biochimie 56*, 61–76.

Pechère, J.-F., Demaille, J., & Canpny, J.-P. (1971) *Biochim. Biophys. Acta 236*, 391–408.

Permyakov, E. A., Murakami, K., & Berliner, L. J. (1987a) *J. Biol. Chem. 262*, 3196–3198.

Permyakov, E. A., Ostrovsky, A. V., & Kalinichenko, L. P. (1987b) *Biophys. Chem. 28*, 225–233.

Perrin, D. D., & Dempsey, B. (1974) *Buffers for pH and Metal Ion Control*, Chapman and Hall, London.

Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273–293.

Saxena, V. P., & Wetlaufer, D. B. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 969–972.

Schellman, J. A. (1975) *Biopolymers 14*, 999–1018.

Schellman, J. A. (1978) *Biopolymers 17*, 1305-1322.
Schmid, F. X., & Baldwin, R. L. (1979) *J. Mol. Biol. 135*, 199-215.
Segawa, T., & Sugai, S. (1983) *J. Biochem. (Tokyo) 93*, 1321-1328.
Sillén, L. G., & Martell, A. E. (1964) *Stability Constants of Metal-Ion Complexes*, Special Publication No. 17, The Chemical Society, London.
Sillén, L. G., & Martell, A. E. (1971) *Stability Constants of Metal-Ion Complexes Supplement No. 1*, Special Publication No. 25, The Chemical Society, London.
Svärd, M., & Drakenberg, T. (1986) *Acta Chem. Scand., Ser. B B40*, 689-693.
Williams, T. C., Corson, D. C., Oikawa, K., McCubbin, W. D., Kay, C. M., & Sykes, B. D. (1986) *Biochemistry 25*, 1835-1846.
Yagi, K., Matsuda, S., Nagamoto, H., Mikuni, T., & Yazawa, M. (1982) in *Calmodulin and Intracellular Ca^{++} Receptors* (Kakiuchi, S., Hidaka, H., & Means, A. R., Eds.) pp 75-91, Plenum, New York.

Alterations of the Glutamine Residues of Human Apolipoprotein AI Propeptide by in Vitro Mutagenesis. Characterization of the Normal and Mutant Protein Forms†

Ali Roghani and Vassilis I. Zannis*

Section of Molecular Genetics, Cardiovascular Institute, Departments of Medicine and Biochemistry, Housman Medical Research Center, Boston University Medical Center, 80 East Concord Street, Boston, Massachusetts 02118

Received March 9, 1988; Revised Manuscript Received May 31, 1988

ABSTRACT: We have used site-directed mutagenesis to independently alter the Gln residues at positions −1 and −2 of the human apoAI propeptide to Arg residues. The normal and mutated genes were placed under the control of the mouse metallothionein 1 promoter in a bovine papilloma virus (BPV) vector which also carries a copy of the human metallothionein 1A gene. Following transfection of mouse fibroblast C127 cells with the vectors, cell clones resistant to $CdCl_2$ were selected and analyzed for production of apoAI mRNAs and protein. The RNA blotting analysis showed that the steady-state apoAI mRNA levels of cell clones expressing either the normal or the mutant apoAI gene are 3-5-fold higher than that of the liver or HepG2 cells. Two-dimensional gel electrophoresis of radiolabeled apoAI showed that the apoAI-expressing clones secreted mainly the proapoAI form. Furthermore, both mutant proapoAI's differed by one positive charge from the normal apoAI. Secretion of apoAI into the culture medium follows apparent first-order kinetics and gives similar rate constants for the normal and mutant apoAI forms. Separation of secreted apoAI by density gradient ultracentrifugation in the presence of human plasma or HDL shows identical distribution of plasma and nascent (normal and mutant) apoAI. The findings indicate that in the cell system used the modification of either of the two glutamines of the apoAI prosegment does not affect the intracellular transport and secretion of apoAI, and its ability to associate with HDL.

Apolipoprotein AI (apoAI)[1] is the major protein component of HDL and serves as a cofactor for the enzyme lecithin-cholesterol acyltransferase (LCAT) (Fielding et al., 1972; Soutar et al., 1975). In tissue culture, apoAI is a ligand for the recognition of HDL by a putative receptor (Fidge & Nestel, 1985; Hwang & Menon, 1985; Schmitz et al., 1985a,b) and promotes efflux of cholesterol from cells (Jackson et al., 1975).

Studies of normal human apoAI synthesis and secretion (Zannis et al., 1980, 1983; Gordon et al., 1983) and DNA sequence analysis of the human apoAI gene (Karathanasis et al., 1983; Shoulders et al., 1983; Cheung & Chan, 1983) have established that newly secreted apoAI has a six amino acid long NH_2-terminal extension (prosegment) with the sequence Arg-His-Phe-Trp-Gln-Gln. It is interesting that both the human (Zannis et al., 1983) and rat (Gordon et al., 1982) apoAI prosegments terminate in a pair of Gln residues. In contrast, the chicken apoAI propeptide contains Gln-His (Rajavashisth et al., 1987), and most other known propeptides contain a pair of basic amino acids at this region (Chance et al., 1968; Hamilton et al., 1974; Russell & Geller, 1975; Nakanishi et al., 1979; Kangawa & Matsuo, 1979; Patzelt et al., 1979; Steiner et al., 1980). The latter configuration allows these propeptides to be cleaved in secretory vesicles by trypsin or carboxypeptidase B type proteases, whereas the apoAI prosegment is removed in plasma by a specific calcium-dependent metalloprotease which cleaves the −1 Gln−+1 Asp bond of proapoAI (Edelstein et al., 1984; Sliwkowski & Windmueller, 1984; Bojanovski et al., 1984).

Recent studies have suggested important roles of protein pre- and prosegments for the structure, function, and intracellular processing and transport of proteins (Folz & Gordon, 1986, 1987; Wu et al., 1986; Suttie et al., 1987; Jorgensen et al., 1987; Blachly-Dyson & Stevens, 1987; Johnson et al., 1987; Valls et al., 1987; Wise et al., 1988).

This study was undertaken in order to investigate the biological significance of the two Gln residues of apoAI propeptide. The results indicate that in the cell system used alteration of either of the Gln residues of the human apoAI

† This work was supported by grants from the National Institutes of Health (HL33952 and HL26335). V.I.Z. is an Established Investigator of the American Heart Association.

[1] Abbreviations: apoAI, apolipoprotein AI; HDL, high-density lipoprotein; LDL, low-density lipoprotein; IDL, intermediate-density lipoprotein; VLDL, very low density lipoprotein; SDS, sodium dodecyl sulfate; BPV, bovine papilloma virus; mMT-1, mouse metallothionein 1; bp, base pair; kb, kilobase.

0006-2960/88/0427-7428$01.50/0