# Folding of a Four-Helix Bundle: Studies of Acyl-Coenzyme A Binding Protein†

Birthe B. Kragelund,‡ Carol V. Robinson,§ Jens Knudsen,|| Christopher M. Dobson,*,⊥ and Flemming M. Poulsen*,‡

*Carlsberg Laboratorium, Kemisk Afdeling, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark, Oxford Centre for Molecular Sciences, Dyson Perrins Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QY, U.K., Institute of Biochemistry, Odense University, Campusvej 55, DK-5230 Odense M, Denmark, and Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QT, U.K.*

*Received January 30, 1995; Revised Manuscript Received March 17, 1995*⊗

ABSTRACT: The refolding from denaturing conditions of a small four-helix bundle, the acyl-coenzyme A binding protein, has been investigated by utilizing an array of fast-reaction techniques. Stopped-flow tryptophan fluorescence for measuring the incorporation of aromatic residues into the protein core and far- and near-ultraviolet circular dichroism to measure the formation of secondary and tertiary structure, respectively, together with the formation of persistent structure measured by hydrogen exchange pulse labeling experiments analyzed by electrospray ionisation mass spectrometry all show that 90% of the acyl-coenzyme A binding protein molecules achieve their fully folded and active, native state with a time constant of less than 5 ms at 25 °C and of ca. 30 ms at 5 °C. The kinetic parameters measured by the different techniques are closely similar, indicating that the different elements of structure form effectively concomitantly. There is no evidence for a significant population of any partially structured intermediate states, and the kinetics are identical whether refolding occurs from an unfolded state generated either by low pH or by addition of guanidine hydrochloride. The kinetics of both refolding and unfolding are monophasic processes for practically 90% of the molecules, and can be described by a two-state model. The results add to our knowledge of the folding scheme of different structural motifs and are discussed in terms of current views of the mechanisms of protein folding.

Our knowledge of the manner by which proteins fold to their native states has advanced considerably in recent years, in large part as a result of the development and utilization of fast-reaction techniques capable of monitoring discrete structural changes taking place along the folding pathways (Baldwin, 1993; Matthews, 1993; Evans & Radford, 1994). A number of distinct steps in folding is commonly observed in such studies, including the rapid initial formation of secondary structure and the slower development of stabilizing interactions prior to the formation of the fully folded, native state (Matthews, 1993). The proteins for which the folding behavior hitherto has been analyzed are diverse both in their native three-dimensional structures and in the complexity of their folding processes. Some proteins such as chymotrypsin inhibitor 2 (CI2)[1] and ubiquitin appear to be able to fold rapidly and highly cooperatively (Jackson & Fersht, 1991a,b; Briggs & Roder, 1992), while others such as barnase, lysozyme, myoglobin, and interleukin 1$\beta$ appear to fold more slowly and to involve more defined intermediates and even multiple pathways (Bycroft *et al.*, 1990; Radford *et al.*, 1992; Jennings & Wright, 1993; Varley *et al.*, 1993). The divergence between these proteins in the structural progression during the folding process is not yet understood, and the extent to which the sequence of events during folding is determined by factors such as the type of secondary structure, the overall topology, and the thermodynamic stability remains to be elucidated. From a study of the folding behavior of proteins that vary in size, topology, and complexity, one may gain information on the influence on the folding process of these and other determinants. In this paper, we describe the refolding from denaturing conditions of acyl-coenzyme A binding protein (ACBP). This small, 86-residue, single-chain protein is folded into an up−down−down−up four-helix bundle structure (Andersen & Poulsen, 1992, 1993; Kragelund *et al.*, 1993), an example of a well-defined supersecondary structure motif (Figure 1). The arrangement of the four helices in the native state is maintained mainly by side chain interactions primarily established in three hydrophobic junctions where the helices interact either pairwise or in threes (B. B. Kragelund, P. Højrup, M. S. Jensen, C. Wogensen, E. Juul, J. Knudsen, and F. M. Poulsen, unpublished). ACBP has no disulfide bonds, bound metal ions, cofactors, or prosthetic groups and is therefore a relatively simple system for studying fundamental steps in the folding process. The folding has been probed by a range of independent physical techniques designed to monitor accurately the formation of distinct features of the protein structure ranging from the development of secondary structure to the ability to bind a specific ligand.

† B.B.K. was supported by a Short Term EMBO Fellowship. This is a contribution from the Danish Protein Engineering Research Centre, which is supported by the Biotechnology Initiative of the Danish Research Councils, and from the Oxford Centre for Molecular Sciences, which is supported jointly by BBSRC, EPSRC, and MRC. The research of C.M.D. is supported in part by an International Research Scholars award from the Howard Hughes Medical Institute.

* To whom correspondence should be addressed.
‡ Carlsberg Laboratorium, Kemisk Afdeling.
§ Dyson Perrins Laboratory, University of Oxford.
|| Odense University.
⊥ New Chemistry Laboratory, University of Oxford.

⊗ Abstract published in *Advance ACS Abstracts,* May 1, 1995.

[1] Abbreviations: ANS, 8-anilino-1-naphthalenosulfonate; ACBP, acyl-coenzyme A binding protein; CD, circular dichroism; CI2, chymotrypsin inhibitor 2; CoA, coenzyme A; ESI-MS, electrospray ionization mass spectrometry; GuHCl, guanidine hydrochloride; HSQC, heteronuclear single quantum coherence; NMR, nuclear magnetic resonance; UV, ultraviolet; SF, stopped-flow.

0006-2960/95/0434-7217$09.00/0 

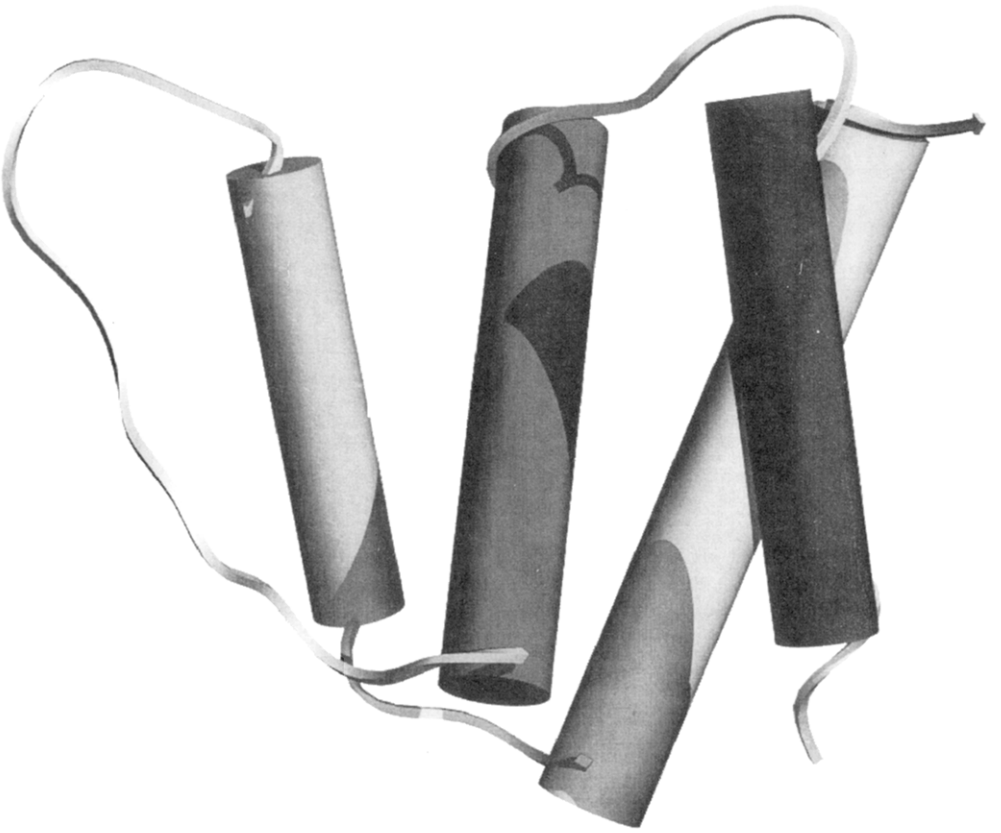

FIGURE 1: Ribbon diagram of recombinant bovine ACBP drawn through the backbone atoms C, $C^{\alpha}$, and N in the lowest energy NMR structure (Andersen *et al.*, 1993). The four helices are shown as cylinders in individual shadings: helix A1, darkest (residues 4−15); helix A2 (residues 21−36); helix A3 (residues 51−62); and helix A4, brightest (residues 65−84). Residues not involved in helical secondary structure are shown by the ribbon only. The arrowhead indicates the C-terminus.

## MATERIALS AND METHODS

*Equilibrium Folding.* Recombinant bovine ACBP was produced and purified as described by Mandrup *et al.* (1991). The purity was checked by ESI-MS and revealed a minor fraction of around 5% of N-terminal acetylated molecules. Fluorescence and CD spectra were recorded of recombinant bovine ACBP at 5 °C, in 0.02 M sodium acetate, pH 5.3 (native state), in 6 M GuHCl, 0.02 M sodium acetate, pH 5.3 (GuHCl-unfolded state), and in HCl, pH 2.25 (acid-unfolded state). CD spectra in the far- and near-UV were recorded at 5 °C at a protein concentration of 0.18 mg/mL and 1.8 mg/mL, respectively, with a path length of 0.1 cm using a Jobin-Yvon CD6 circular dichrograph. Fluorescence spectra at 5 °C for the individual states were recorded on a Perkin-Elmer LS 50B fluorometer; $\lambda_{max}$ is reported with the excitation wavelength at 280 nm. Unfolding by GuHCl at 5 °C was monitored by tryptophan fluorescence, excitation wavelength at 295 nm and emission wavelength at 356 nm, at a protein concentration of 4 $\mu$M. The concentrations of GuHCl were determined by refractive index (Nozaki, 1972), and samples were equilibrated for at least 1 h before measurement. Amide hydrogen to deuterium exchange at pH 2.25, 25 °C, was measured by two-dimensional $^{15}N-^{1}H$ heteronuclear single quantum coherence (HSQC) NMR spectra as described previously (B. B. Kragelund, J. Knudsen, and F. M. Poulsen, submitted for publication) and compared to intrinsic rate constants calculated by using recently improved parameters (Bai *et al.*, 1993) and pH and temperature parameters for deuterium oxide dissociation constant (Covington *et al.*, 1966).

*Stopped-Flow Experiments.* Refolding was initiated by mixing protein denatured in 6 M GuHCl, 0.02 M sodium acetate, pH 5.3, with 10 volumes of 0.02 M sodium acetate, pH 5.3 at 5 °C, to give a residual denaturant concentration of 0.545 M GuHCl. Final protein concentrations were 1.82 mg/mL (near-UV CD), 0.09 mg/mL (fluorescence), and 0.18 mg/mL (far-UV CD). When refolding from the acid-denatured state, the dilution buffer contained 0.02 M sodium acetate and 0.6 M GuHCl, pH 5.6, resulting in an identical refolding mixture (i.e., 0.545 M GuHCl, $pH_{final} = 5.3$). Refolding kinetics in the presence of ligand were obtained with *S*-hexadecyl-CoA, a stable thio-ether analog of palmitoyl-CoA, present in a 1:1 molar ratio in the refolding buffer. *S*-Hexadecyl-CoA was synthesized and purified as previously described (Rosendal *et al.*, 1993). The CD experiments were done using a Jobin-Yvon CD6 circular dichrograph equipped with a Biologic SFM3 stopped-flow module, and fluorescence experiments were performed on an Applied Photophysics SX 17MV stopped-flow fluorometer with excitation at either 280 nm or 295 nm and emission at 357 nm. The influence of denaturant on refolding and unfolding rates was examined at 5 °C by intrinsic fluorescence using a Biologic SFM4 module connected to a Perkin-Elmer Lambda9 spectrophotometer, excitation wavelength at 280 nm. Total emission above 320 nm was collected. Various concentrations of GuHCl were included in the refolding/unfolding buffer, and the reaction was initiated by diluting either unfolded ACBP (6 M GuHCl, 0.02 M sodium acetate, pH 5.3) or folded ACBP (0.02 M sodium acetate, pH 5.3) into these.

*Hydrogen Exchange Labeling and ESI-MS.* Samples for mass spectrometry to measure protection of amides from exchange at different times following the initiation of refolding were prepared using a Biologic QFM5 rapid mixing quench flow apparatus. The scheme described by Radford *et al.* (1992) was applied under the following conditions: Lyophilized protein was dissolved in 99.9% $D_2O$, 6 M GuDCl, and 0.02 M deuterated sodium acetate, pH 5.3 to a concentration of 10 mg/mL. Refolding was initiated at pH 5.3, 5 °C, by an 11-fold dilution into refolding buffer (0.02 M sodium acetate, pH 5.3), and the protein was allowed to refold for various defined times up to 5 s. Exchange of protein deuterons during refolding was neglected, except for long refolding times ($t$ > 1s). Deuterons not trapped in persistent structure during refolding were exchanged for protons by applying an 8.5 ms high-pH pulse, in 0.2 M boric acid, pH 11.0, to give a resulting exchange pH of 10.2. Deuterons with a protection factor less than ca. 10 [$1/k_{\text{intrinsic,avg. ACBP}}$(5 °C, pH 10.2) = 0.76 ms] would be substantially exchanged under these conditions. The pH of the solution was then lowered to pH 5.0 by addition of acetic acid, and folding was allowed to go to completion at a final protein concentration of 0.48 mg/mL. The samples were washed at 5 °C with 5 volumes of formic acid in water, pH 5, and electrospray ionization mass spectra obtained essentially as described by Robinson *et al.* (1994), from $H_2O$ at pH 5.0 and low temperature. Hydrogen exchange labeling experiments of the complex were performed with *S*-hexadecyl-CoA present in the deuterated unfolded protein/6 M GuDCl solution. The time-dependent populations of folded and unfolded ACBP were extracted from the spectra after base-line correction and normalization by a least-squares analysis of the contributions from the $t$ = 0 ms spectrum and $t$ = 5004 ms spectrum to each spectrum. Line broadening due to exchange during refolding occurred at long refolding times ($t$ > 1s). The dead time of the quench-flow apparatus was around 4 ms.

*Data Analysis.* Data were fitted using ACE/gr (Turner, 1992) or in-house programs. The uncertainties of the exponential fits have been calculated by the fitting procedure under the assumption that the error of all data points has the same normal distribution, and represents 95% confidence intervals. The ESI-MS data were fitted to a single exponential as no slow phase could be revealed under the conditions of the experiment.

## RESULTS

*Equilibrium Folding of ACBP.* Folded and unfolded states of ACBP denatured either by GuHCl or by low pH were measured by CD in the far- and near-UV, by tryptophan fluorescence and through measurement of amide hydrogen exchange by NMR spectroscopy. Both denaturants produce spectroscopically distinct, highly unfolded denatured states of ACBP (Figure 2a,b). A cooperative, reversible two-state process with only the native and the unfolded form was detectably populated throughout the transition for both denaturants and shown for GuHCl only, in Figure 2c. No indications of the existence of any stable partly folded intermediate have been found under these or any other conditions explored. The free energy of unfolding of ACBP in water, $\Delta G_{U-F}^{H_2O}$, and the $m$-value, a constant proportional to the degree of exposure of the protein on denaturation, were determined by fitting the data (Figure 2c) by a nonlinear

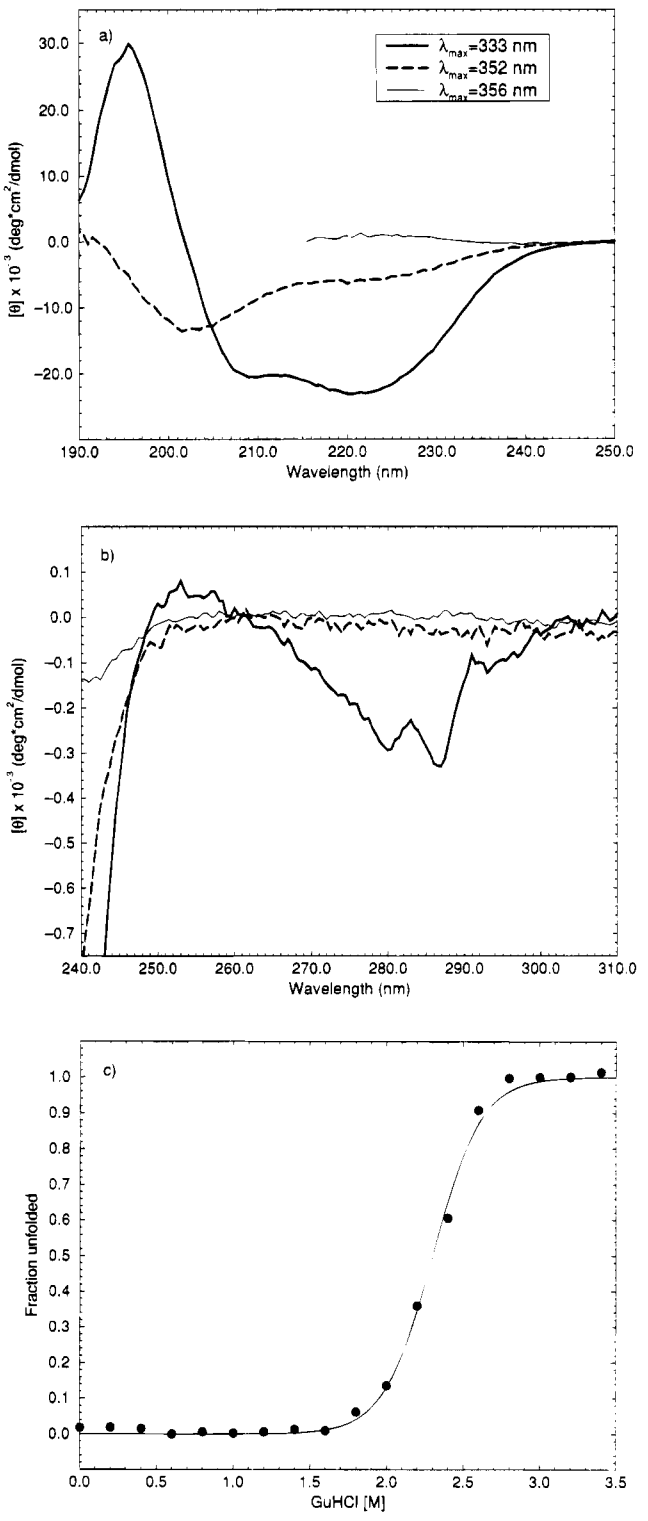


FIGURE 2: Structural characterization of the native and denatured states of recombinant bovine ACBP, 5 °C. (a) Far-UV and (b) near-UV CD spectra of ACBP: (boldface solid) native, 0.02 M sodium acetate, pH 5.3, 5 °C; (solid) GuHCl-induced denatured, 6 M GuHCl, 0.02 M sodium acetate, pH 5.3, 5 °C; (boldface dashed line) acid-induced denatured (HCl, pH 2.25, 5 °C). The mean residue elipticity, [$\theta$], is shown on the ordinate, and fluorescence emission $\lambda_{max}$ values with an excitation wavelength at 295 nm are shown in the inset of panel a. (c) Unfolding of recombinant bovine ACBP at 5 °C, 0.02 M sodium acetate, pH 5.3, followed by tryptophan fluorescence, excitation wavelength at 295 nm, emission at 356 nm. The intensity of the fluorescence has been normalized to that observed for the unfolded and folded states.

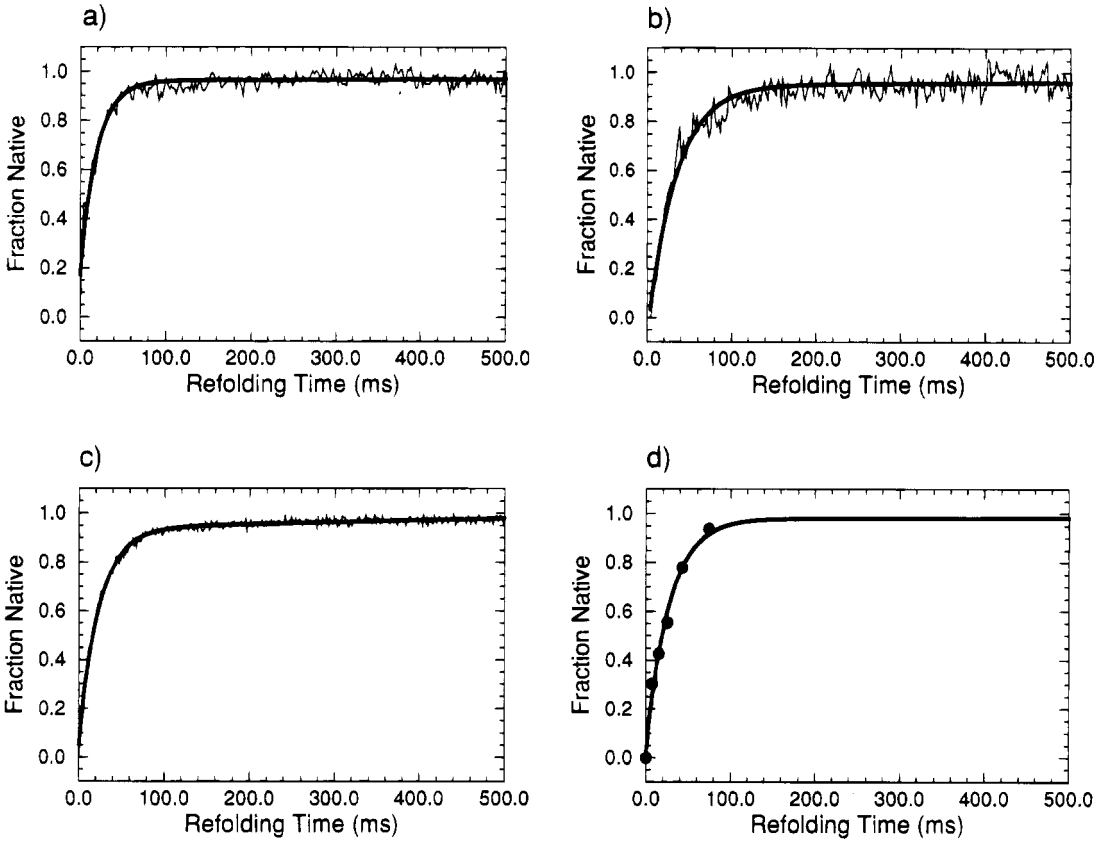


FIGURE 3: Refolding of recombinant bovine ACBP at 5 °C monitored by (a) far-UV (225 nm) CD, (b) near-UV (286 nm) CD, (c) tryptophan fluorescence with an excitation wavelength at 295 nm and an emission wavelength at 357 nm, and (d) electrospray ionization mass spectrometry combined with hydrogen exchange pulse-labeling. The boldface curves drawn represent the average of (a, b, c) a double-exponential fit and (d) a single-exponential fit to the individual time courses which were normalized to the unfolded and folded states, respectively. The traces represent the average of 55 shots (far-UV, 225 nm), 45 shots (near-UV, 286 nm), and 30 shots (fluorescence).

analysis as described by Jackson *et al.* (1993) for a two-state reaction. For ACBP, $\Delta G_{U-F}^{H_2O} = 6.1 \pm 0.9$ kcal mol$^{-1}$ and $m = 2.7 \pm 0.8$ kcal mol$^{-1}$ M$^{-1}$. The midpoint of unfolding is at 2.36 M GuHCl. The *m*-value determined for ACBP is within the range of values reported for other globular proteins (Chen & Schellman, 1989; Kellis *et al.*, 1989; Pace *et al.*, 1990; Jackson & Fersht, 1991a), although small. The structure of recombinant bovine ACBP is reminiscent of a hollow bowl (Andersen & Poulsen, 1993), and this architecture may be the origin of the smaller degree of exposure of residues on denaturation. For CI2, however, the *m*-value reported is even smaller (Jackson & Fersht, 1991a), indicating that the hydrophobic core of CI2 is smaller than for ACBP.

In the acid-induced unfolded state at pH 2.25, the far-UV CD spectrum shows 25% of residual helical elipticity at 222 nm which in the spectrum of ACBP in 6 M GuHCl is absent (Figure 2a). Amide hydrogen−deuterium exchange studies on the acid-induced unfolded state as measured by two-dimensional NMR spectroscopy showed an apparent bulk protection factor of less than 2 (data not shown). This degree of protection is similar to the average protection observed in other highly unfolded states (Arcus *et al.*, 1994; Buck *et al.*, 1994) and strongly indicates that no elements of structure persist in these states which are capable of significantly protecting against amide hydrogen exchange. Moreover, under the conditions in which refolding occurs in the experiments reported here, all the spectroscopic data show that the protein adopts the fully native state.

*Kinetics of ACBP Folding.* The kinetic refolding of ACBP was followed by a combination of time-resolved techniques designed to probe different characteristics of the structure as these develop during the folding process (Evans & Radford, 1994). Stopped-flow CD in the far- and near-UV were monitored to probe respectively the formation of secondary structure, i.e., the four α-helices, and the development of an environment able to immobilize aromatic side chains within the protein structure, reflecting native tertiary interactions. In addition, since the native structure of ACBP has a well-defined hydrophobic core which involves the two tryptophans and three of the four tyrosine residues (Andersen & Poulsen, 1992, 1993; Kragelund *et al.*, 1993), stopped-flow fluorescence was used to monitor the burial of these residues from solvent as a consequence of their incorporation into the protein core.

Refolding from different denatured states was examined by dilution of the protein from two denaturing solvents; 6 M GuHCl, 0.02 M sodium acetate, pH 5.3, and dilute HCl, pH 2.25. The reaction mixture in the two sets of experiments following dilution was made identical (i.e., containing 0.545 M GuHCl and having a pH of 5.3) by including GuHCl in the refolding buffer in the latter case. Preliminary experiments at 25 °C showed that the major events in the folding process of ACBP occur in less than 5 ms and are too fast to follow accurately with present methodology. Detailed kinetic refolding experiments were therefore carried out at 5 °C.

Measurements by all three stopped-flow techniques resulted in kinetic refolding profiles that are very closely

Table 1: Kinetic Data on the Refolding of ACBP at 5 °C from (a) 6 M GuHCl with and without Ligand Present and from (b) pH 2.25[a]

| | fast phase | | slow phase | |
|---|---|---|---|---|
| | rate ($s^{-1}$) | amplitude (%) | rate ($s^{-1}$) | amplitude (%) |
| | | (a) 6 M GuHCl | | |
| SF fluorescence | 46.6 ± 2.2 | 90.2 ± 2.8 | 8.5 ± 2.8 | 9.8 ± 2.8 |
| SF near-UV CD | 30.8 ± 4.4 | 88.1 ± 6.6 | 4.9 ± 2.4 | 11.9 ± 6.2 |
| SF far-UV CD | 41.5 ± 9.4 | 90.5 ± 11.4 | 2.4 ± 2.4 | 9.5 ± 4.1 |
| ESI-MS[b] | 33.8 ± 4.6 | 100 | | |
| SF fluorescence$_{ligand}$ | 53.5 ± 3.4 | 87.7 ± 2.8 | 10.9 ± 2.4 | 12.3 ± 3.2 |
| | | (b) pH 2.25 | | |
| SF fluorescence | 39.5 ± 1.4 | 92.6 ± 1.6 | 5.8 ± 2.0 | 7.4 ± 1.4 |
| SF near-UV CD | 36.5 ± 18.6 | 89.3 ± 24.4 | 2.5 ± 5.0 | 10.7 ± 11.0 |
| SF far-UV CD[b] | 38.8 ± 6.8 | 100 | | |

[a] The uncertainties have been calculated by the fitting procedure under the assumption that the errors of all data points have the same normal distribution, and represent 95% confidence intervals. SF: stopped-flow. [b] These data were fitted to a single exponential as no slow phase could be revealed at the present sensitivity of the experiment.

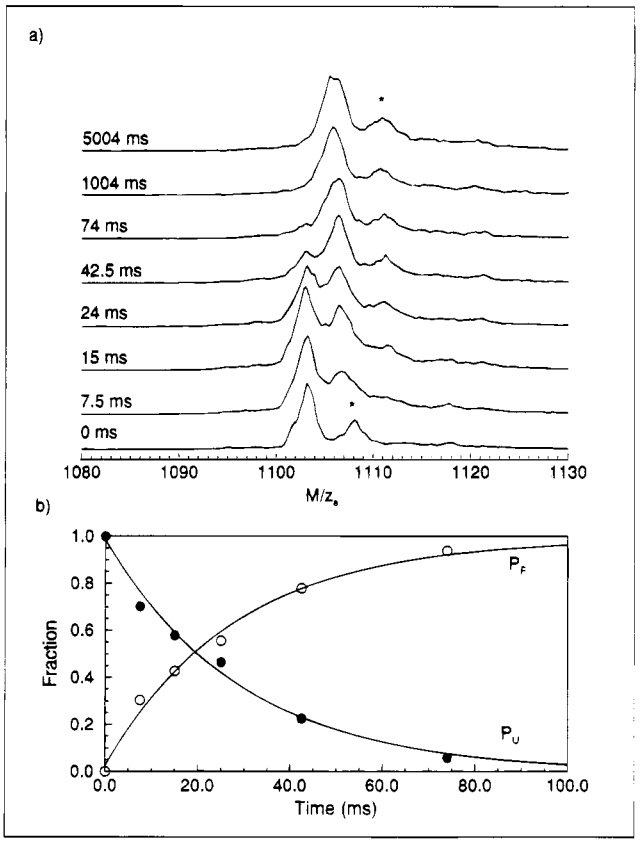


FIGURE 4: Time evolution of (a) the electrospray ionization mass spectra of recombinant bovine ACBP from hydrogen exchange pulse-labeling samples at different times during refolding and (b) populations of the unlabeled (representing the unfolded state) (●) and partially deuterated (representing the fully folded state) (○) ACBP species of each mass spectrum fitted to a single exponential. The spectrum at $t = 0$ represents that of the unlabeled ACBP species (mass = 9920 Da), and the spectrum at $t = 5004$ ms represents that of the partially deuterated ACBP species (mass = 9950 Da). The increased mass of the unlabeled ACBP species (8 Da) is due to the isotopic composition of the refolding buffer. The side peaks (*) are recombinant molecules which are *N*-terminally acetylated; line broadening due to exchange is observed for long refolding times ($t > 1$s).

similar (Figure 3 and Table 1). There is no evidence for any significant events occurring in the dead time of the measurements, typically less than 4 ms. The refolding of ACBP is fast and has a profile of two distinct exponential kinetic phases: a first and fastest phase of refolding has a time constant of around 25 ms and accounts for approximately 90% of the kinetic amplitude in each experiment, and a second and slower phase of refolding, having a time constant of around 200 ms, accounts for the remaining 10% of the amplitude (Table 1).

The similarity of the kinetic profiles recorded with the different techniques shows that the secondary and tertiary structural features of ACBP develop effectively simultaneously during folding in an apparently two-state process. Importantly, no differences in the rates or in the amplitudes of folding from the two differently obtained denatured states were observed, indicating that the refolding process of ACBP is essentially independent of the nature of the unfolded state and the unfolding agent. This situation resembles that found recently for lysozyme (Kotik *et al.*, 1995) using similar techniques and is consistent with conclusions from earlier kinetic studies (Kato *et al.*, 1981), but differs from results observed for cytochrome *c* (Sosnick *et al.*, 1994). In the latter case, however, the acid-denatured state adopts a highly structured compact conformation, a situation shown in this work not to be the case for ACBP. For cytochrome *c*, this state can convert rapidly into the native state (Sosnick *et al.*, 1994). Furthermore, the results for ACBP reported here demonstrate that the measured rates are not limited by diffusion of denaturants from the unfolded polypeptide chain.

Further evidence concerning the refolding of ACBP was obtained in three additional sets of experiments. In the first of these, ESI-MS was combined with hydrogen exchange pulse-labeling techniques in order to monitor the development of persistent structure during folding, and to search for indications of intermediates or lack of cooperativity (Miranker *et al.*, 1993). The group of amide probes that remain protected after formation of the native state to detection in the mass spectrometer, which includes the time frame for desalting of the sample, is restricted to those amides that exchange slowly in the native state of ACBP. Such exchange in ACBP has previously been thoroughly analyzed, and the helical backbone amides located in all four helices have been shown to have relatively large protection factors in excess of $10^4-10^7$ at pH 6.65, 25 °C (B. B. Kragelund, J. Knudsen, and F. M. Poulsen, unpublished experiments). During folding, only two mass-specific populations are present (Figure 4a); the unlabeled and the partly deuterated species corresponding to populations of the unfolded and the fully folded states, respectively, present at the time of the labeling pulse. A plot of these shows that cooperative protection of the approximately 30 amides protected in the

native structure at pH 5.0 after 20 h at 5 °C occurs with a time constant very close to that detected by the optical techniques (Table 1 and Figure 4b) and that no subsequent steps generate additional detectable protection. No very slow phase corresponding to those seen in the optical experiments could, however, be revealed by the present resolution of the experiment. The presence of the small fraction of N-terminal acetylated molecules does not affect the function of ACBP (Rosendal *et al.*, 1993) and structurally, only a minor stabilization of the N-terminal of helix A1 is observed (Andersen *et al.*, 1991). From data obtained in this study, no apparent differences in the protection of amides in these molecules compared to the unacetylated molecules were found.

In the second additional set of experiments, the generation of functionality of ACBP was probed. ACBP binds long-chain acyl-CoA esters with high affinity in an intricate one-to-one binding fashion (Rosendal *et al.*, 1993; Rasmussen *et al.*, 1994) that involves residues from all four helices (Kragelund *et al.*, 1993). When palmitoyl-CoA is bound in the center of ACBP, its adenine ring stacks with Tyr31 of ACBP, resulting in a quenching of the native fluorescence (Mikkelsen & Knudsen, 1987; Kragelund *et al.*, 1993). Again, the formation of the hydrophobic core of ACBP was monitored by native tryptophan and tyrosine fluorescence, but this time the pH-stable ligand *S*-hexadecyl-CoA was included in the refolding buffer. The rate of development of the ACBP core was not significantly altered by the presence of the ligand (Table 1), and the expected quenching of the native signal by the bound ligand occurred concurrently with formation of the core and not in an isolated step. Correspondingly, ACBP was unfolded in the presence of the ligand, and refolded and analyzed by ESI-MS. Analysis of the mass spectra showed the presence of two sets of peaks, one arising from the complex and one from the protein; only two mass-specific populations were again found to be present during folding. Furthermore, the overall rate of refolding was the same within experimental error as in the previous experiments. These results provide good evidence that the binding site in ACBP must be generated effectively simultaneously with the formation of the helices, and does not, therefore, represent a step subsequent to the overall folding of ACBP.

In the third and final additional set of experiments, the rates of both refolding and unfolding were measured as a function of the final concentration of GuHCl in the refolding/unfolding buffer. The natural logarithm of the observed rate constants shows a V-shaped dependence on residual denaturant (Figure 5) with a minimum (2.32 M) corresponding closely to the midpoint of the unfolding transition measured at equilibrium (2.36 M). On the basis of a two-state reaction, with only the unfolded and the fully folded form of the protein significantly populated, the kinetics of folding and unfolding of ACBP can be fitted to the equation (Jackson *et al.*, 1993):

$$\ln k_{obs} = \ln[k_f^{H_2O} \exp(-m_{k_f}[\mathrm{GuHCl}]) + k_u^{H_2O} \exp(m_{k_u}[\mathrm{GuHCl}])] \quad (1)$$

From this fit, values for $k_f^{H_2O}$, $k_u^{H_2O}$, $m_{k_f}$ and $m_{k_u}$ were determined as 279.7 ± 29.4 $s^{-1}$, 1.0 × $10^{-4}$ ± 2 × $10^{-5}$ $s^{-1}$, 3.9 ± 0.2 $M^{-1}$, and 2.5 ± 0.2 $M^{-1}$, respectively.

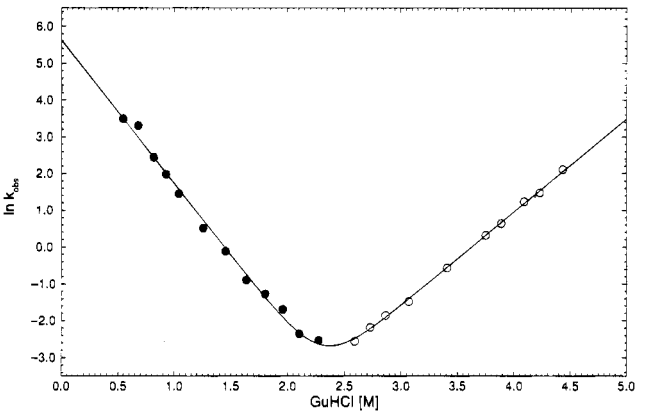


FIGURE 5: Denaturant dependence of the natural logarithm of the observed rate constants for folding (●) and unfolding (○) of ACBP at 5 °C, pH 5.3, measured by stopped-flow fluorescence (excitation at 280 nm, emission above 320 nm). The solid line represent the best fit to the two-state model (eq 1). The fast refolding of ACBP hampered the recording of data below 0.56 M GuHCl.

Equilibrium values for $\Delta G_{U-F}^{H_2O}$ and $m$ can be calculated from the apparent equilibrium constant for unfolding, $K = k_f^{H_2O}/k_u^{H_2O}$, and from the slopes of the V-shaped denaturant dependence, $m = m_{k_f} + m_{k_u}$, taking into account the 10% slow-folding population molecules. Values of $\Delta G^{H_2O}$ and of $m$ were found to be 8.1 ± 1.8 kcal $mol^{-1}$ and 3.5 ± 0.1 kcal $mol^{-1}$ $M^{-1}$, respectively. The values obtained in these kinetic experiments are within experimental error of those obtained from the equilibrium unfolding studies of ACBP described above. As found previously for CI2 (Jackson & Fersht, 1991b), this high level of agreement between kinetic and equilibrium data provides further evidence that ACBP behaves kinetically as a two-state system of folding; at least 90% of the protein molecules acquire all characteristics of the fully native protein efficiently and concurrently.

## DISCUSSION

For a number of proteins whose folding has been studied in detail, a variety of steps are apparent in the folding process (Baldwin, 1993; Matthews, 1993; Evans & Radford, 1994). Evidence has been obtained for the rapid formation of extensive secondary structure within or close to the dead-times of the experimental techniques, and prior to the formation of persistent structure detected by amide hydrogen exchange pulse labeling experiments, and such structure is associated with, or preceded by, the formation of a compact state having a hydrophobic core in several cases shown to be capable of binding fluorescent probes such as ANS (Kuwajima *et al.*, 1987; Bycroft *et al.*, 1990; Ptitsyn *et al.*, 1990; Elöve *et al.*, 1992; Radford *et al.*, 1992; Varley *et al.*, 1993; Jennings & Wright, 1993; Itzhaki *et al.*, 1994). The formation of such a state, frequently called a molten globule (Bychkova & Ptitsyn, 1993), through a process of collapse that brings together the hydrophobic residues to give a rudimentary native-like structure but without persistent tertiary interactions, offers a model for the generation and stabilization of secondary structure and for limiting the extent of conformational space that needs to be explored in order to achieve the characteristics of the fully native structure (Chan & Dill, 1990; Bychovka & Ptitsyn, 1993; Dobson, 1994; Peng & Kim, 1994). Support for such a model comes from theoretical as well as experimental studies, including simulations involving simplified lattice models of the folding

process (Sali *et al.*, 1994; Baldwin, 1994; Fersht *et al.*, 1994).

The use of the combination of independent time-resolved techniques applied in the present study has revealed important features of the folding of ACBP, in particular the high cooperativity of the folding process which is reflected in the closely similar kinetics measured by the different techniques. This similarity suggests that the secondary structure, the hydrophobic core, and at least the large majority of stable tertiary interactions are formed effectively concomitantly for at least 90% of the ACBP molecules. The results provide no evidence for significant structure formation in any burst phase prior to kinetically detectable events, nor for any significantly populated states intermediate between the fully unfolded and fully folded ones. Furthermore, the finding that the capability for binding a specific ligand by ACBP develops on a similar time scale to that observed by means of the physical probes of structure suggests strongly that no subsequent slow steps are involved in the folding process. The high cooperativity observed in the folding of ACBP is not, however, unique and has been reported for other small proteins, notably the 64 residue truncated form of CI2 and the 76 residue protein ubiquitin (Jackson & Fersht, 1991a,b; Briggs & Roder, 1992). Interestingly, the kinetics of folding of these two proteins are closely similar to those measured here for ACBP. In contrast to ACBP, which is all α-helical, both CI2 and ubiquitin have extensive β-structure and only a single α-helix. This implies that efficient and cooperative folding is not simply attributable to the type of secondary structure within a protein.

During folding of ACBP, some 10% of the molecules were detected to fold much more slowly than the majority of molecules although still without any evidence for well-detected intermediates. Similar kinetic diversity accounting also for approximately 10−20% of the total amplitude was observed in the study of the refolding of both CI2 and ubiquitin. The origin of the very slow folding behavior of this fraction of molecules has not yet been determined for ACBP. The slower phase of CI2 folding has been shown to be catalyzed by peptidylprolyl isomerase (Jackson & Fersht, 1991b), and for ubiquitin hydrogen exchange results suggest that it is a result of structural heterogeneity in the unfolded state around the three prolines (Briggs & Roder, 1992). ACBP contains two prolines: Pro19 which is in the loop between helices A1 and A2 and Pro44 which is in the overhand loop between helices A2 and A3. Although the kinetics of even the slow component of ACBP folding are relatively fast compared to those of proline isomerism in simple peptides (Nall *et al.*, 1978), it is possible that *cis/trans* isomerization of these two proline peptide bonds is responsible for the existence of a slow folding population of ACBP.

The fact that the folding behavior, with the exception of the event associated with proline isomerism of small proteins such as ACBP, CI2, and ubiquitin, appears to differ from that of many other proteins is interesting in light of experimental evidence indicating that slow steps in the folding of other proteins might involve reorganization steps occurring as a result of events taking place during the collapse stage of the folding process (Radford *et al.*, 1992; Sosnick *et al.*, 1994). Such a conclusion has been drawn from the observation that the folding of hen lysozyme involves parallel pathways (Radford *et al.*, 1992; Miranker *et al.*, 1993; Kotik *et al.*, 1995). While some 30% of the molecules appear to be able to form persistent native-like structure in 5 ms or less, the remainder exist in a well-defined collapsed state for 50 ms or longer prior to formation of such structure. This suggests that the fast folding population of lysozyme arises from intermediates whose structure is amenable to further folding, the remaining molecules requiring reorganizational steps prior to further folding (Radford *et al.*, 1992; Miranker *et al.*, 1993; Kotik *et al.*, 1995). Evidence for misfolding events deferring folding has come also from the fact that the folding of cytochrome *c* under some conditions is much slower than under others, an observation that has been attributed to events involving incorrect heme ligation that generates barriers to formation of the native structure (Sosnick *et al.*, 1994). This has been used to propose that misfolding of proteins, rather than intrinsic barriers, may frequently be rate-limiting events in protein folding (Creighton, 1994; Dobson, 1994; Sosnick *et al.*, 1994). Furthermore, the greater complexity involved in forming correct secondary structure in a β-sheet rather than an α-helix has been suggested as the origin of slow folding in the all-β protein interleukin 1β (Varley *et al.*, 1993; Gronenborn & Clore, 1994). Here the need for rearrangements arises as a consequence of the necessity of "unzippering" incorrectly formed sheet structure prior to continued folding (Dill *et al.*, 1993).

The driving force for the collapse of the polypeptide chain during folding is likely to arise at least in part from hydrophobic interactions (Dill, 1985) and hence will normally be greater for larger proteins, which have a higher ratio of internal to external residues (Miller *et al.*, 1987). Evidence for this has been suggested from comparative studies on CI2 and barnase showing a greater potential for the smaller protein to have an unperturbed folding process whereas the larger protein has kinetically detectable intermediates along its folding path (Otzen *et al.*, 1994). It may be that relatively small proteins will only form a stable collapsed state if the latter has a correct overall topology and hence many of the native-like interactions in place. The early stages of folding might then involve the search for appropriate contacts between distinct regions and specific residues of the polypeptide chain, perhaps by a diffusion-collision process (Karplus & Weaver, 1976, 1994) before collapse to a stable compact structure is initiated. Once collapse has occurred, however, rapid conversion to the fully native state is then able to take place (Sosnick *et al.*, 1994). As the possibilities for indiscriminate collapse of at least some regions of the polypeptide chain will increase with size, it seems likely that misfolding will be more probable for larger proteins. Whatever the origin of the slower steps in the folding of other proteins, however, the present results of the four-helix bundle protein, ACBP, reinforce the view that fast, efficient folding may be a characteristic of small proteins with simple folds regardless of the nature of their secondary structure.

## ACKNOWLEDGMENT

We thank Dr. Sheena E. Radford, Dr. Kim Vilbour Andersen, Dr. Robin T. Aplin, Christian Rischel, and Charles Morgan for valuable help and discussions. Erling Knudsen and Pia Skovgaard are thanked for skilled technical assistance.

## REFERENCES


Andersen, K. V., & Poulsen, F. M. (1992) *J. Mol. Biol. 226*, 1131−1141.

Andersen, K. V., & Poulsen, F. M. (1993) *J. Biomol. NMR 3*, 271−284.

Andersen, K. V., Ludvigsen, S., Mandrup, S., Knudsen, J., & Poulsen, F. M (1991) *Biochemistry, 30*, 10654−10663.

Arcus, V. L., Vuilleumier, S., Freund, S. M. W., Bycroft, M., & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9412−9416.

Bai, Y., Milne, J. S., Mayne, L., & Englander, W. S. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.

Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 84−91.

Baldwin, R. L. (1994) *Nature 369*, 183−184.

Briggs, M. S., & Roder, H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2017−2021.

Buck, M., Radford, S. E., & Dobson, C. M. (1994) *J. Mol. Biol. 237*, 247−254.

Bychkova, V. E., & Ptitsyn, O. B. (1993) *Chemtracts 4*, 133−163.

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

Chan, H. S., & Dill, K. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6388−6392.

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

Covington, A. K., Robinson, R. A., & Bates, R. G. (1966) *J. Phys. Chem. 70*, 3820−3824.

Creighton, T. E. (1994) *Nat. Struct. Biol. 1*, 135−138.

Dobson, C. M. (1994) *Curr. Biol. 4*, 637−640.

Dill, K. A., (1985) *Biochemistry 24*, 1501−1509.

Dill, K. A, Feibig, K. M., & Chan, H. S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1942−1946.

Elöve, G. A., Chaffotte, A. F., Roder, H., & Goldberg, M. E. (1992) *Biochemistry 31*, 6876−6883.

Evans, P. A., & Radford, S. E. (1994) *Curr. Opin. Struct. Biol. 4*, 100−106.

Fersht, A. R., Itzhaki, L. S., elMasry, N. F., Matthews, J. M., & Otzen, D. E. (1994) *Proc. Natl, Acad. Sci. U.S.A. 91*, 10426−10429.

Gronenborn, A. M., & Clore, G. M. (1994) *Science 263*, 536.

Itzhaki, L. S., Evans, P. A., Dobson, C. M., & Radford, S. E. (1994) *Biochemistry 33*, 5512−5520.

Jackson, S. E., & Fersht, A. R. (1991a) *Biochemistry 30*, 10428−10435.

Jackson, S. E., & Fersht, A. R. (1991b) *Biochemistry 30*, 10436−10443.

Jackson, S. E., elMasry, N., & Fersht, A. R. (1993) *Biochemistry 32*, 11270−11278.

Jennings, P. A., & Wright, P. E. (1993) *Science 262*, 892−895.

Karplus, M., & Weaver, D. L. (1976) *Nature 260*, 404−406.

Karplus, M., & Weaver, D. L. (1994) *Protein Sci. 3*, 650−668.

Kato, S., Okamura, M., Shimamoto, N., & Utiyama, H. (1981) *Biochemistry 20*, 1080−1085.

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

Kotik, M., Radford, S. E., & Dobson, C. M. (1995) *Biochemistry 34*, 1714−1724.

Kragelund, B. B., Andersen, K. V., Madsen, J., Knudsen, J., & Poulsen, F. M. (1993) *J. Mol. Biol. 230*, 1260−1277.

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

Mandrup, S., Højrup, P., Kristiansen, K., & Knudsen, J. (1991) *Biochem. J. 276*, 817−823.

Matthews, B. W. (1993) *Annu. Rev. Biochem. 62*, 139−160.

Mikkelsen, J., & Knudsen, J. (1987) *Biochem. J. 248*, 709−714.

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

Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., & Dobson, C. M. (1993) *Science 262*, 896−899.

Nall, B. T., Garel, J. R., & Baldwin, R. L. (1978) *J. Mol. Biol. 118*, 317−330.

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

Otzen, D. E., Itzhaki, L., ElMasry, N. F., Jackson, S. E., & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10422−10425.

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

Peng, Z., & Kim, P. S. (1994) *Biochemistry 33*, 2136−2141.

Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulayaev, O. I. (1990) *FEBS Lett. 262*, 20−24.

Radford, S. E., Dobson, C. M., & Evans, P. A. (1992) *Nature 358*, 302−307.

Rasmussen, J. T., Færgeman, N. J., Kristiansen, K., & Knudsen, J. (1994) *Biochem. J. 299*, 165−170.

Robinson, C. V., Gross, M., Eyles, S. J., Ewbank, J. J., Mayhew, M., Hartl, F. U., Dobson, C. M., & Radford, S. E. (1994) *Nature 372*, 646−649.

Rosendal, J., Ertbjerg, P., & Knudsen, J. (1993) *Biochem. J. 290*, 321−326.

Sali, A., Shakhnovich, E., & Karplus, M. (1994) *Nature 369*, 248−251.

Sosnick, T. R., Mayne, L., Miller, R., & Englander, S. W. (1994) *Nat. Struct. Biol. 1*, 149−156.

Turner, P. J. (1992) Centre for Coastal and Land-Margin Research, Beaverton, OR.

Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., & Clore, M. (1993) *Science 260*, 1110−1113.


BI950198D