# Folding of Dihydrofolate Reductase from *Escherichia coli*†

Nancy A. Touchette,‡ Kathy M. Perry,§ and C. Robert Matthews*

*Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802*

*Received June 21, 1985; Revised Manuscript Received May 15, 1986*

ABSTRACT: The urea-induced equilibrium unfolding transition of dihydrofolate reductase from *Escherichia coli* was monitored by UV difference, circular dichroism (CD), and fluorescence spectroscopy. Each of these data sets were well described by a two-state unfolding model involving only native and unfolded forms. The free energy of folding in the absence of urea at pH 7.8, 15 °C is 6.13 ± 0.36 kcal mol$^{-1}$ by difference UV, 5.32 ± 0.67 kcal mol$^{-1}$ by CD, and 5.42 ± 1.04 kcal mol$^{-1}$ by fluorescence spectroscopy. The midpoints for the difference UV, CD, and fluorescence transitions are 3.12, 3.08, and 3.18 M urea, respectively. The near-coincidence of the unfolding transitions monitored by these three techniques also supports the assignment of a two-state model for the equilibrium results. Kinetic studies of the unfolding and refolding reactions show that the process is complex and therefore that additional species must be present. Unfolding jumps in the absence of potassium chloride revealed two slow phases which account for all of the amplitude predicted by equilibrium experiments. Unfolding in the presence of 400 mM KCl results in the selective loss of the slower phase, implying that there are two native forms present in equilibrium prior to unfolding. Five reactions were observed in refolding: two slow phases designated $\tau_1$ and $\tau_2$ that correspond to the slow phases in unfolding and three faster reactions designated $\tau_3$, $\tau_4$, and $\tau_5$ that were followed by stopped-flow techniques. The kinetics of the recovery of the native form was monitored by following the binding of methotrexate, a tight-binding inhibitor of dihydrofolate reductase, at 380 nm. For refolding in the presence of saturating methotrexate, the four slower reactions, $\tau_1$–$\tau_4$, lead to the binding of the inhibitor; the $\tau_5$ phase does not. The increase in fluorescence intensity observed for the $\tau_5$ phase is opposite to the changes observed for the $\tau_1$–$\tau_4$ phases and is contrary to what would be expected for a progressive folding reaction. This result suggests that the early stages of folding may involve a hydrophobic cluster that protects tryptophan residues from solvent-associated quenching mechanisms. The urea dependence of the $\tau_1$–$\tau_5$ reactions is characteristic of folding, and the activation energies range from 13 to 37 kcal mol$^{-1}$. A folding model which incorporates all of these results is proposed that involves a multiplicity of native, intermediate, and unfolded forms.

Although it has long been known that the amino acid sequence of a protein determines its three-dimensional structure (Anfinsen, 1973), the mechanism by which this complex conformational change takes place is not known in detail for any protein. The high cooperativity and short time frame ($10^{-3}$–$10^{1}$ s) for this process have made it extremely difficult to characterize with current spectroscopic techniques. A new approach toward solving this problem, made possible by the spectacular advances in recombinant DNA technology, is to investigate the effects of single amino acid substitutions on protein folding and stability. The replacement of amino acids that play key roles should have observable effects on the kinetic and equilibrium properties of the folding transition. Analysis of these results in terms of the X-ray structure may provide a better understanding of the mechanisms by which proteins fold.

Dihydrofolate reductase (DHFR, EC 1.5.1.3)[1] from *Escherichia coli*, a relatively small globular protein ($M_r$ 17 680), appears to be an ideal candidate for the mutagenic approach toward solving the folding problem. This enzyme whose key role in metabolism has led to extensive studies of its activity and inhibition by folate analogues (Cayley et al., 1981; Baccanari et al., 1982; Blakley, 1984; Stone & Morrison, 1984) has several important advantages:

(1) DHFR is monomeric and has no prosthetic groups or disulfide bonds. Thus, refolding kinetics will be unimolecular, and there will be no constraints placed on the conformations available to the unfolded form by intramolecular cross-links.

(2) The gene for *Escherichia coli* DHFR has been cloned into a high copy number plasmid, the DNA sequence determined, and efficient expression achieved (Smith & Calvo, 1980). Site-directed mutagenesis (Craik, 1985; Kunkel, 1985) provides a means for synthesizing point mutations in cloned genes.

(3) The 1.7-Å, refined X-ray structure of *Escherichia coli* DHFR (Bolin et al., 1982; Filman et al., 1982) provides a basis for the quantitative analysis of the observed effects.

Recently, the synthesis of point mutations at three different sites in DHFR has been reported (Howell et al., 1986; Villafranca et al., 1983). Replacements at Asp-27 and Gly-95 implicate these residues in important roles in the structure and function of this enzyme. The replacement of Pro-39 by Cys-39 was intended to permit the introduction of a disulfide bond between this residue and Cys-85. Preliminary evidence suggests that the disulfide bond can form. Chen et al. (1985) have

† This work was supported by National Science Foundation Grant PCM-81-04495. C.R.M. is the recipient of Research Career Development Award K04 AG00153 from the U.S. Public Health Service.

‡ Present address: Department of Biochemistry, University of California, Berkeley, Berkeley, CA 94720.

§ Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143.

[1] Abbreviations: CD, circular dichroism; KCl, potassium chloride; DHFR, dihydrofolate reductase; $NaDodSO_4$, sodium dodecyl sulfate; $K_2EDTA$, dipotassium ethylenediaminetetraacetic acid; MTX, methotrexate; NMR, nuclear magnetic resonance; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; NADPH, nicotinamide adenine dinucleotide phosphate, reduced form.

0006-2960/86/0425-5445$01.50/0 

studied mutations of DHFR at residues 45 (His → Gln) and 113 (Thr → Val) to assess interactions of the wild-type enzyme with the pyrophosphate backbone of NADPH and with dihydrofolate, respectively. Analysis of the kinetic behavior of these mutant proteins concludes that product release is unlikely to be the rate-determining step in the wild-type mechanism. These results support the proposition that mutagenesis will prove to be an important tool for understanding the structure–function relationship in proteins. We believe that mutagenesis will also be a very useful tool in unraveling the mechanism of folding.

In this paper, we report the results of a study of the equilibrium and kinetic properties of the urea-induced reversible unfolding transition in wild-type DHFR. From these data, we have developed a model for the folding reaction which can be used to investigate the effects of amino acid replacements. In a subsequent paper, the effect of replacing Leu-28 with Arg, Asp-27 with Asn, Glu-139 with Lys, and Thr-113 with Val on the stability and folding of DHFR will be reported (K. Perry, J. Onuffer, M. Gittelman, N. Touchette, and C. R. Matthews, unpublished results).

## Materials and Methods

*Protein Source and Purification.* Dihydrofolate reductase (DHFR, EC 1.5.1.3) was isolated from two sources: (1) a trimethoprim-resistant strain of *Escherichia coli* B (RT-500) and (2) the *Escherichia coli* strain, MC294, which contains the plasmid pTY1 (D. Tu and T. Yaegashi, unpublished results). This vector was constructed by excising the DHFR gene from pCV29 (Smith & Calvo, 1980) and inserting it into a derivative of pBR322 which lacks the *Eco*RI site (a gift from D. Shortle). The only difference between the amino acid sequences of DHFR from these two sources involves the assignment of the residue at position 118. The DNA sequence of the plasmid-derived DHFR indicates Glu (Smith & Calvo, 1980) while amino acid sequencing of the RT-500-derived DHFR indicates Gln (Stone et al., 1977). We believe that both sequences actually have Glu at this position for the following reasons: (1) reexamination of the evidence for the assignment of Gln from amino acid sequencing shows that there is no conclusive evidence for Gln [Tables VIIIS and XS in Baccanari et al. (1981)], and (2) native polyacrylamide gel electrophoresis of purified samples of both proteins shows that they migrate at identical rates (M. Matthews, unpublished results). A Glu/Gln difference at this solvent-exposed position would be expected to affect migration on native gels. This apparent discrepancy, if real, has no observed functional effects on folding. The equilibrium and kinetic parameters for folding are identical for protein isolated from these two sources.

Protein purification was performed according to the method of Baccanari et al. (1975, 1977). Purity was demonstrated by the observation of single bands on both $NaDodSO_4$ and native polyacrylamide gels. Protein concentration was measured by the absorbance at 280 nm, using a molar extinction coefficient of $3.11 \times 10^4$ $M^{-1}$ $cm^{-1}$ (D. Baccanari, personal communication). Enzymatic activity was monitored as described previously (Hillcoat et al., 1967). The specific activity of various samples used in these experiments ranged from 73 to 95 units $mg^{-1}$ in 0.1 M imidazole chloride (pH 7.0), 30 °C. The reported activity under these conditions is 85 units $mg^{-1}$ (Baccanari et al., 1977).

*Spectroscopic Methods.* Equilibrium unfolding as a function of urea was monitored by three spectroscopic techniques: (1) UV difference spectroscopy, where the change in tryptophan absorbance was measured at 293 nm on a Cary 118 spectrophotometer; (2) circular dichroism (CD) by following the loss in secondary structure at 220 nm on a Jasco J-20 spectropolarimeter; and (3) fluorescence spectroscopy by exciting the tryptophans at 290 nm and monitoring the emission at 378 nm on a Perkin-Elmer MPF-66 fluorometer. All samples were allowed to fully equilibrate at the appropriate final urea concentration at 15 °C before spectra were taken.

Kinetics of folding were also followed by UV difference spectroscopy using manual mixing methods which have a dead time of 10–15 s and by fluorescence spectroscopy using a Durrum 110 stopped-flow spectrophotometer which has a dead time of 10 ms. Fluorescence measurements on the Durrum instrument were made by exciting at 290 nm with a slit width of 5 mm. The emission intensity was monitored at wavelengths above 340 nm by using a Corning C.S. 0-52 ground-glass filter. To facilitate kinetic measurements, both the Cary 118 and Durrum 110 spectrophotometers were interfaced to a PDP 11/23 computer. Equilibrium and kinetic experiments in the absence of KCl were performed in 10 mM potassium phosphate, 0.2 mM $K_2EDTA$, and 1 mM 2-mercaptoethanol at pH 7.8. The buffer used in CD experiments was identical; however, the $K_2EDTA$ was deleted. Final protein concentrations ranged from 0.15 to 0.60 mg $mL^{-1}$.

*Inhibitor Binding and Salt Studies.* Methotrexate (MTX) binding was monitored at 380 nm where an extremum in the difference spectrum between bound and unbound MTX exists. MTX concentration was determined by using a molar extinction coefficient of $2.21 \times 10^4$ $M^{-1}$ $cm^{-1}$ at 302 nm in 0.1 M NaOH (Dawson et al., 1969). Unfolding jumps to a final concentration of 50 mM KCl were performed with protein dialyzed against 15 mM Tris-HCl, 0.2 mM $K_2EDTA$, and 1 mM 2-mercaptoethanol at pH 7.8 and either 0 m or 400 mM KCl initially.

*Computer Fitting of Equilibrium and Kinetic Data.* Nonlinear least-squares fits to kinetic data were obtained by using the Statistical Analysis Systems program NLIN (SAS Institute Inc., Cary, NC) and the equation:

$$A(t) = \sum_i A_i \exp(-t/\tau_i) + A_\infty$$

where $A(t)$ is the total amplitude at time $t$, $A_\infty$ is the amplitude at infinite time, $A_i$ is the amplitude corresponding to the individual phase, $i$, at zero time, and $\tau_i$ is the associated relaxation time. The data for the apparent fraction unfolded, $F_{app}$, were fit to the equation:

$$F_{app} = \exp(-\Delta G_{app}/RT)/[1 + \exp(-\Delta G_{app}/RT)]$$

where $\Delta G_{app}$ is the apparent free energy difference between the native and unfolded forms in a two-state model, $R$ is the gas constant, and $T$ is the temperature. The apparent free energy difference was assumed to have a linear dependence on the urea concentration (Schellman, 1978).

*Reagents.* Ultrapure urea was obtained from Schwarz/Mann and used without further purification. MTX was a gift from R. Bales of Microbiological Associates, Inc., or purchased from Sigma. MTX affinity resin was purchased from Pierce. All other chemicals were obtained from commercial sources and were reagent grade.

## Results

*Equilibrium Studies.* The urea-induced unfolding transition of DHFR was followed by difference UV, CD, and fluorescence spectroscopy. Comparison of the transition curves obtained by these three techniques provides a method of determining the proper equilibrium unfolding model. For a simple two-state model involving only a single native and single unfolded form, the transition curves obtained from different

spectroscopic methods are coincident; noncoincidence implies that additional stable forms are present (Wong & Tanford, 1973; Waheed et al., 1977).

The change in the extinction coefficient at 293 nm caused by increasing concentrations of urea at pH 7.8, 15 °C is shown in Figure 1A. The linear increase in extinction coefficient up to 2 M urea reflects the change in solvent composition (Yanari & Bovey, 1960) and indicates that the native form is stable up to 2 M urea. The sigmoidal decrease between 2 and 5 M urea can be attributed to a cooperative unfolding transition. The linear increase in extinction coefficient above 5 M urea is again due to changes in solvent composition and indicates that the unfolded conformation is populated in this region. From the magnitude of the observed change in extinction coefficient obtained by extrapolation to 0 M urea, $-4300\ M^{-1}\ cm^{-1}$, one can estimate that a net of 2.7 Trp residues become exposed to solvent upon unfolding (Donovan, 1973). The transition is fully reversible as judged both by optical measurements (Figure 1A) and by the recovery of over 95% of the original activity after exposure to 6 M urea for 1 h at 25 °C. Panels B and C of Figure 1 show the change in the molar ellipticity at 220 nm and fluorescence intensity at 378 nm at increasing urea concentrations, respectively. Similar to the difference UV spectroscopy results, a cooperative unfolding transition occurs between 2 and 5 M urea.

These data were converted to a form suitable for comparison by the equation $F_{app} = (Y_{obsd} - Y_N)/(Y_U - Y_N)$. The term $F_{app}$ refers to the fractional conversion of the native form, N, to the unfolded form, U. $Y_{obsd}$ refers to the observed extinction coefficient, molar ellipticity, or fluorescence intensity, and $Y_N$ and $Y_U$ refer to the respective spectroscopic properties of the native and unfolded forms at the particular urea concentration. Because all three techniques show small linear dependences on urea concentration in the appropriate base-line regions (Figure 1A–C), the values for $Y_N$ and $Y_U$ in the unfolding transition region were estimated by linear extrapolation.

The normalized versions of these three transition curves are shown in Figure 1D. The difference UV, CD, and fluorescence data sets can be independently fit to a two-state model to determine the stability of DHFR. The free energy of unfolding in the absence of urea calculated from the difference UV data is 6.13 ± 0.36 kcal $mol^{-1}$ and is 5.32 ± 0.67 kcal $mol^{-1}$ from CD and 5.42 ± 1.04 kcal $mol^{-1}$ from fluorescence measurements. The midpoint in the unfolding transition obained from difference UV, CD, and fluorescence spectroscopy occurs at 3.12 ± 0.02, 3.08 ± 0.04, and 3.18 ± 0.08 M urea, respectively, at pH 7.8, 15 °C. Considering the intrinsic errors involved, including the necessity of extrapolating the spectroscopic properties of the native and unfolded forms into the transition region in order to calculate $F_{app}$, these differences are not significant. If folding intermediates are present, it is likely that they comprise a small fraction of the total population under these conditions.

*Kinetic Studies.* Although the results of the equilibrium studies suggest that the folding follows a two-state model, kinetic studies are required to test this hypothesis. A combination of absorbance and fluorescence data was collected to determine the transient response of DHFR in unfolding and refolding reactions. A representative kinetic trace of $\Delta A_{293}$ for an unfolding jump at pH 7.8, 15 °C is shown in Figure 2A. For unfolding jumps ending at 5.4 M urea, two kinetic phases were observed that accounted for all the absorbance changes expected from the equilibrium data. At 5.4 M urea, the slower phase which will be designated as the $\tau_1$ phase has a relaxation time of 414 s, while the faster phase, designated

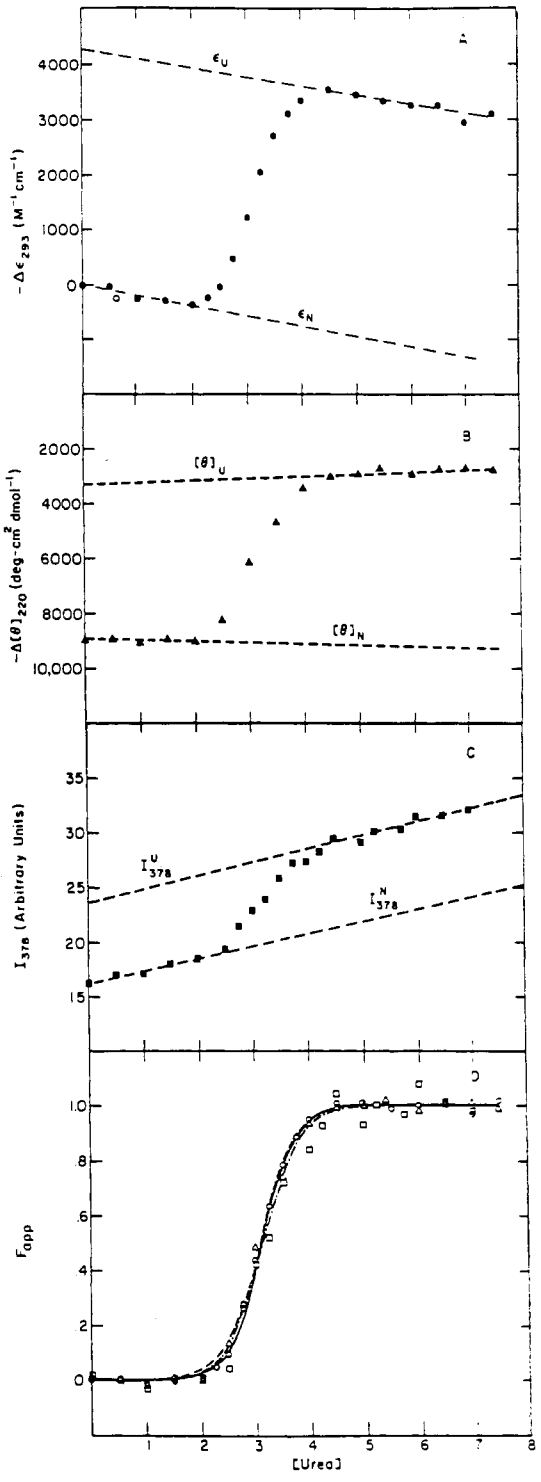



FIGURE 1: Equilibrium unfolding of DHFR monitoring urea-induced unfolding by (A) tryptophan absorbance at 293 nm using UV difference spectroscopy (●), (B) molar ellipticity at 220 nm (▲), and (C) tryptophan fluorescence at 378 nm, excited at 290 nm (■). The solvent for all studies contained 10 mM potassium phosphate and 1 mM 2-mercaptoethanol at pH 7.8, 15 °C; the solvent for the UV and fluorescence studies also contained 0.2 mM $K_2EDTA$. The dashed lines represent the assumed dependence of the optical properties of the native and unfolded forms on the urea concentration. The reversibility of the unfolding transition was demonstrated by incubating a sample of the protein at 5.4 M urea for 1 h to ensure complete unfolding and then diluted to 0.54 M urea; (○) represents the difference UV spectrum at 293 nm. (D) Urea concentration dependence of the fractional change, $F_{app}$, calculated from the difference UV (○, —), CD (△, – –), and fluorescence (□, –·–) data in Figure 1A–C. The lines indicate the fits to a two-state model, assuming a linear dependence of the free energy of folding on the urea concentration (Schellman, 1978).

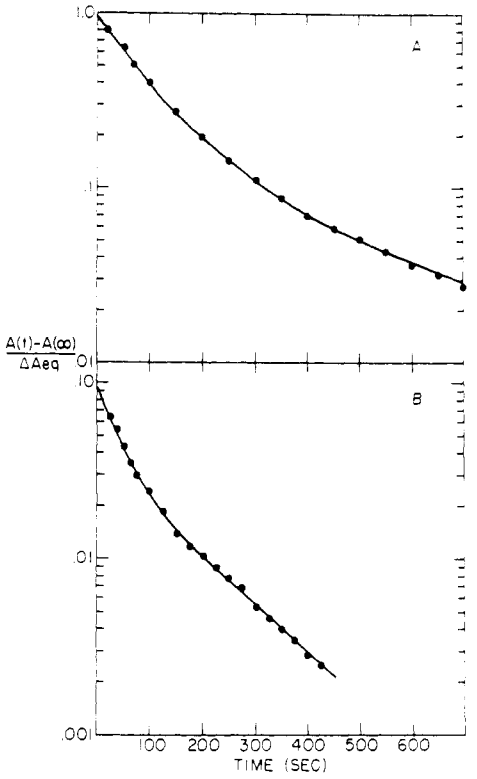



FIGURE 2: Semilog plot of the change in absorbance at 293 nm normalized to the value expected from equilibrium measurements as a function of time for (A) the unfolding reaction starting at 0 M urea and ending at 5.4 M urea and (B) the refolding reaction starting at 5.4 M urea and ending at 0.54 M urea. The solvent contained 10 mM potassium phosphate, 0.2 mM $K_2EDTA$, and 1 mM 2-mercaptoethanol at pH 7.8; the temperature was 15 °C. The solid lines represent theoretical fits of the data, in each case, to two exponentials.

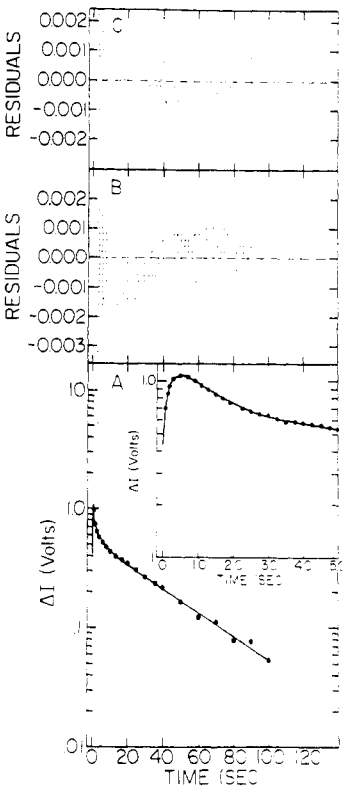



FIGURE 3: (A) Log of the fluorescence change vs. time for a refolding jump from 5.4 M to 0.54 mM urea measured by stopped-flow. The tryptophans were excited at 290 nm, and the emitted light was monitored at wavelengths above 340 nm. The inset shows the initial stages of the folding reaction. The solid line represents the calculated fit to the sum of four exponentials. (B) and (C) are plots of the residuals corresponding to fits to three and four exponentials, respectively. Solvent conditions are identical with those described in Figure 2. The ordinate indicates photomultiplier tube voltage.

as the $\tau_2$ phase, has a relaxation time of 125 s.

The refolding of DHFR to urea concentrations less then 1.5 M displayed a complex pattern that required a minimum of five exponentials to obtain an adequate fit. The slower reactions, designated $\tau_1$ and $\tau_2$, were monitored by difference UV spectroscopy at 293 nm and are shown in Figure 2B. The reactions whose relaxation times, in decreasing order, are designated by $\tau_2$, $\tau_3$, $\tau_4$, and $\tau_5$ were monitored by fluorescence spectroscopy on a stopped-flow instrument. The rapid refolding reactions could not be monitored by difference UV spectroscopy at 293 nm on the stopped-flow instrument because the protein transiently aggregated at the concentrations required to obtain a detectable signal. This aggregation effect disappeared after a few seconds, permitting the measurement of slower ($\tau > 10$ s) folding reactions by absorption spectroscopy on the Cary 118 spectrophotometer. A typical trace of a refolding jump from 5.4 to 0.54 M urea obtained from stopped-flow studies is shown in Figure 3A. Following refolding, the fluorescence intensity first *increases* rapidly with a relaxation time, $\tau_5$, of 233 ms at 0.54 M urea. This burst in fluorescence is then followed by successive decreases in intensity which are well described by three relaxation times; $\tau_4$, $\tau_3$, and $\tau_2$. At 15 °C and 0.54 M urea, $\tau_4 = 1.20$ s, $\tau_3 = 6.80$ s, and $\tau_2 = 44$ s. Difference UV measurements using manual mixing techniques gave values of 37 amd 165 s for $\tau_2$ and $\tau_1$, respectively, under these conditions. The validity of fitting the fluorescence data in the time range from 10 ms to 200 s to four exponentials, $\tau_2$–$\tau_5$, is supported by residual plots shown in Figure 3B,C. The validity of the fit is further supported by the good agreement for the values for the $\tau_2$ phase determined by difference UV and fluorescence spectroscopies at a series of final urea concentrations (Figure 4). The relaxation time for the $\tau_1$ phase was too long to be accurately measured by stopped-flow fluorescence spectroscopy.

The urea dependence of the relaxation times and amplitudes of these phases was determined from a series of unfolding and refolding jumps to various final conditions. A semilogarithmic plot of the relaxation times as a function of urea concentration is shown in Figure 4. The $\tau_1$ and $\tau_2$ phases in unfolding connect smoothly with the $\tau_1$ and $\tau_2$ phases in refolding. This result permits the correct correlation of the unfolding and refolding reactions and demonstrates the reversibility of this process. The linear increase in ln $\tau$ for both $\tau_1$ and $\tau_2$ phases up to a maximum near 3.5 M urea and the subsequent decrease at higher urea concentrations are typical of the behavior observed for folding reactions in other proteins (Tanford, 1973; Crisanti & Matthews, 1981; Kelley et al., 1986). The relaxation times for the $\tau_3$ and $\tau_4$ phases also decrease as the final urea concentration is decreased. The $\tau_4$ phase could only be resolved below 1.5 M urea (Figure 4).

The $\tau_5$ relaxation process must be quite different from its slower counterparts since it reflects an *increase* in fluorescence intensity; the $\tau_2$, $\tau_3$, and $\tau_4$ phases involve decreases in intensity. Comparison of the fluorescence emission spectra of the native and unfolded protein shows that the total intensity of fluorescence above 340 nm is greater for unfolded DHFR (data not shown). Therefore, a progressive change from the unfolded to the native form during refolding would be expected to result in a decrease in fluorescence intensity. The burst phase shows that folding does not follow a simple, monotonic course. The urea dependence of the $\tau_5$ relaxation time is shown

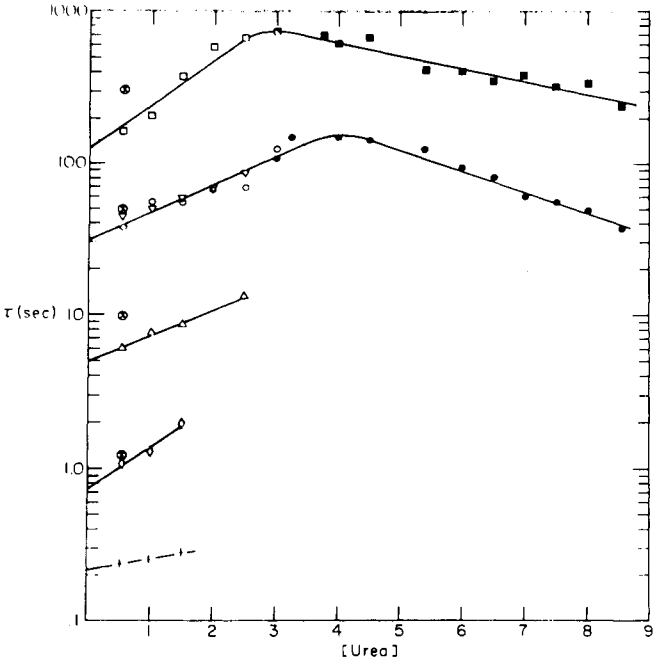



FIGURE 4: Dependence of the observed relaxation times on the final urea concentration. The $\tau_1$ (□, ■) and $\tau_2$ (○, ●) phases were measured on a Cary 118 spectrophotometer. The open symbols correspond to refolding and closed symbols to unfolding; (◪) represents unfolding and refolding jumps to the same final urea concentration. The $\tau_2$ (▽), $\tau_3$ (△), $\tau_4$ (◇), and $\tau_5$ (+) refolding phases were observed by fluorescence changes on a Durrum 110 stopped-flow spectrophotometer. Solid lines are drawn to guide the eye and are not fits to the data. Measurements of $\tau_1$, $\tau_2$, $\tau_3$, and $\tau_4$ in refolding by MTX binding on the stopped-flow instrument at 380 nm are indicated by (⊗). These symbols indicate the average relaxation times for refolding jumps from 5.4 to 0.54 M urea over a range of MTX to DHFR ratios (see Table I). The conditions are identical with those described in Figure 2.

in Figure 4. Above 1.5 M urea, this reaction is not well described by a single exponential, indicating that it is complex. The results for the single-exponential fits are shown in Figure 4.

That these changes in fluorescence intensity are due to folding reactions and not artifacts such as light scattering was demonstrated by changing the excitation wavelength. No fluorescence emission was observed when the sample was excited at 320 nm, where tryptophan absorption is negligible.

The dependences of the amplitudes of the $\tau_1$ and $\tau_2$ phases in unfolding and refolding on urea concentration are shown in Figure 5. Around 6 M urea where the protein is fully unfolded (Figure 1A), the $\tau_1$ phase comprises 15% and the $\tau_2$ phase 85% of the total expected change in extinction coefficient at 293 nm. Above 6 M urea, the amplitudes of both phases decrease slightly, suggesting either that a fast phase is also occurring that accounts for the missing amplitude or that the linear extrapolation of the urea concentration dependence of the extinction coefficient for the native conformation to high urea concentrations is not correct (Figure 1A). Stopped-flow fluorescence spectroscopy was used to determine the correct explanation. No fast phases were detected, demonstrating that additional phases in unfolding, if present, are complete in less than 10 ms. Therefore, it appears that the unfolding of DHFR is governed by two rather slow reactions.

Below 1.5 M urea, the protein is fully folded, and the $\tau_1$ relaxation time accounts for 9% and the $\tau_2$ phase for 12% of the expected change in absorbance (Figure 5). The dependences of the amplitudes of the $\tau_1$ and $\tau_2$ phases in refolding from absorbance measurements are complex in the transition region.

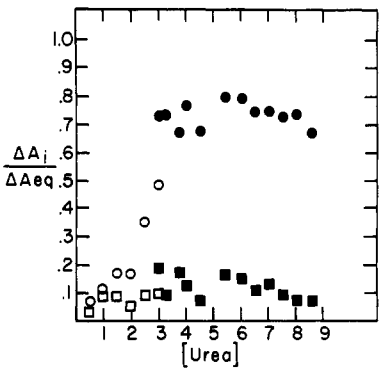



FIGURE 5: Amplitudes for the $\tau_1$ (□, ■) and $\tau_2$ (○, ●) phases (see Figures 2 and 4) normalized to the total change predicted by the UV difference equilibrium curve (Figure 1A); closed symbols correspond to unfolding and open symbols to refolding reactions.

For fluorescence refolding experiments ending in the native base-line region, the amplitudes of the $\tau_2$, $\tau_3$, $\tau_4$, and $\tau_5$ phases are independent of the final urea concentration (data not shown). The $\tau_2$, $\tau_3$, and $\tau_4$ phase amplitudes are 0.53, 0.14, and 0.89 V, respectively. The amplitude of the $\tau_5$ phase has an opposite sign and is −1.13 V.

The amplitudes and relaxation times for both unfolding and refolding were found to be independent of the protein concentration over the range from 0.3 to 1.5 mg $mL^{-1}$ (N. Touchette and K. Perry, unpublished results). Therefore, these reactions are unimolecular and reflect actual protein folding and not dimerization or other higher order aggregation.

The $\tau_1$ and $\tau_2$ phases in unfolding and the $\tau_1$–$\tau_5$ phases in refolding were further characterized by the temperature dependence of the various relaxation times. The activation energies of the $\tau_1$ and $\tau_2$ phases for unfolding jumps ending at 5.4 M urea were 18 and 26 kcal $mol^{-1}$, respectively. For refolding jumps ending at 0.54 M urea, the activation energies for the $\tau_1$ and $\tau_2$ phases detected by absorption measurements were 29 and 18 kcal $mol^{-1}$, respectively, and those for the $\tau_2$, $\tau_3$, $\tau_4$, and $\tau_5$ phases detected by fluorescence measurements were 18, 37, 13, and 26 kcal $mol^{-1}$, respectively.

The complexities in the kinetics for both the unfolding and refolding reactions show that more than the two species deduced from the equilibrium studies must be present. To determine the appropriate kinetic model, two types of experiments were performed. One explanation for the multiple phases in both unfolding and refolding is that more than a single species is present at equilibrium in either the native base-line region or the unfolded base-line region or both. Since the amplitudes of individual phases in folding depend upon the relative populations of species present initially (Hagerman & Baldwin, 1976), a change in the solution conditions prior to the jump could shift the putative preexisting equilibrium and result in a change in the relative amplitudes of the kinetic phases.

It was found that if 400 mM KCl was present before the unfolding reaction was initiated, only a single kinetic phase was observed. The relaxation time of this phase was 286 s when the final conditions were 5.4 M urea and 50 mM KCl. This value is midway between those for the $\tau_1$ and $\tau_2$ phases in the absence of KCl, suggesting that the KCl also alters the unfolding rates. This was shown to be the case by initiating unfolding jumps in the absence of salt and ending in 50 mM KCl at 5.4 M urea. The observed relaxation times were 1100 and 250 s for the $\tau_1$ and $\tau_2$ phases, respectively. Thus, if 400 mM KCl is present initially, the $\tau_1$ phase is eliminated. This result implies a preexisting equilibrium between two native forms.

Table I: Amplitudes and Relaxation Times for Refolding of DHFR Monitored by MTX Binding at 380 nm[a]

| MTX:DHFR ratio | $\tau_1$ (s) | $\tau_2$ (s) | $\tau_3$ (s) | $\tau_4$ (s) | $\alpha_1$[b] | $\alpha_2$[b] | $\alpha_3$[b] | $\alpha_4$[b] |
|---|---|---|---|---|---|---|---|---|
| 1:1 | 220 | 55 | 11 | 1.3 | 0.07 | 0.09 | 0.25 | 0.59 |
| 2:1 | 280 | 40 | 8.3 | 1.3 | 0.05 | 0.17 | 0.17 | 0.61 |
| 3:1 | 390 | 41 | 9.4 | 1.3 | 0.06 | 0.15 | 0.22 | 0.58 |
| 4:1 | 320 | 60 | 11 | 1.3 | 0.05 | 0.12 | 0.27 | 0.55 |
| av[c] | 300 ± 70 | 49 ± 10 | 9.9 ± 1.3 | 1.3 ± 0.0 | 0.06 ± 0.01 | 0.13 ± 0.03 | 0.23 ± 0.04 | 0.58 ± 0.02 |

[a] Refolding jumps from 5.4 to 0.54 M urea at 15 °C in the standard 10 mM potassium phosphate buffer. The final protein concentration was 15 μM. [b] $\alpha_i = \Delta A_i^{380}/\Delta A_{total}^{380}$. [c] The errors are the standard deviation.

The initial conditions for unfolding were also varied by adding urea up to 1.5 M, where the protein is still in the native conformation (Figure 1A). This change in initial conditions has no effect on the amplitudes of the $\tau_1$ and $\tau_2$ phases in unfolding (data not shown). For refolding, the relative amplitudes of the $\tau_1$ and $\tau_2$ phases were found to be independent of the initial urea concentrations in the range from 3 to 6 M urea (data not shown).

The second type of experiment used to determine the folding model of DHFR monitored the binding of an inhibitor during folding. MTX is a potent inhibitor of DHFR which appears to bind to the same site as the substrate, dihydrofolate; however, the pteridine ring system is inverted (Kimber et al., 1977; Bolin et al., 1982). Upon binding to DHFR, an absorbance maximum in MTX at 373 nm shifts to shorter wavelengths. A difference spectrum of bound vs. free MTX reveals an extremum at 380 nm. Because there are no contributions from the protein difference spectrum in this region, changes in absorbance at 380 nm can be assigned to MTX binding. The binding isotherm at 15 °C, pH 7.8 is that expected for a one to one, tight-binding inhibitor; $\Delta\epsilon_{380} = -4942\ M^{-1}\ cm^{-1}$ (data not shown).

A kinetic trace of $\Delta A_{380}$ accompanying the binding of MTX to DHFR during a refolding jump from 5.4 to 0.54 M urea is shown in Figure 6A. A total of four exponentials are required to adequately fit the data. Figure 6B,C demonstrates the validity of the four-phase fit over that obtained from the sum of three exponentials. The amplitudes and relaxation times of these phases for several ratios of MTX to protein are shown in Table I.

The conclusions that can be drawn from these data are: (1) The relaxation times and amplitudes in the MTX binding studies do not depend on the MTX to DHFR ratio. (2) Inhibitor binding kinetics parallel those for the protein folding reactions for the $\tau_1$ through the $\tau_4$ phases. The average relaxation times from MTX binding agree well with those from protein folding by fluorescence and UV difference spectroscopy considering the complexity of the analysis. The average values from the MTX binding are plotted in Figure 4 for comparison.

It should be noted that the absorbance changes at 380 nm that accompany the binding of MTX to fully folded DHFR were complete in the dead time of mixing (<10 ms) at the concentrations of enzyme and inhibitor employed. Therefore, the above results for methotrexate binding reflect conformational changes that occur in the protein during folding.

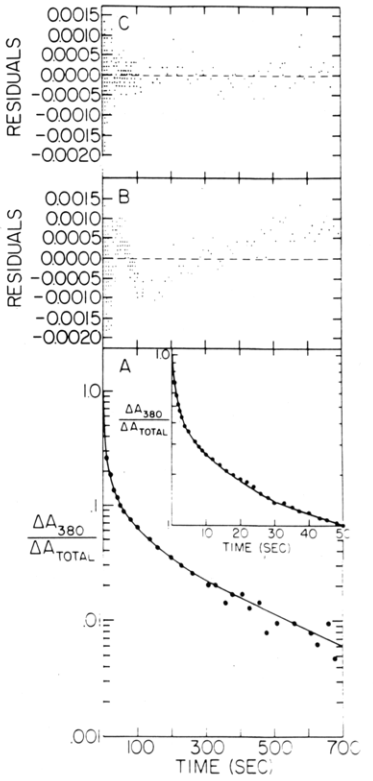
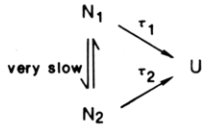



FIGURE 6: Normalized change in the absorbance at 380 nm for refolding from 5.4 to 0.54 M urea in the presence of a MTX:protein ratio of 4:1 by stopped-flow. The solid line represents a fit to the data to four exponentials; the relaxation times at a variety of MTX:protein ratios are shown in Table I. The inset emphasizes the early binding phases. The protein concentration was 15 μM, and solvent conditions are identical with those described in Figure 2. (B) and (C) are plots of the residuals corresponding to fits of three and four exponentials, respectively.

## Discussion

*Folding Models. (1) Unfolding Reaction.* The observation of two slow kinetic phases in unfolding which account for all of the expected absorbance change and which have amplitudes which depend upon the presence of KCl is best explained by the model shown in Scheme I.

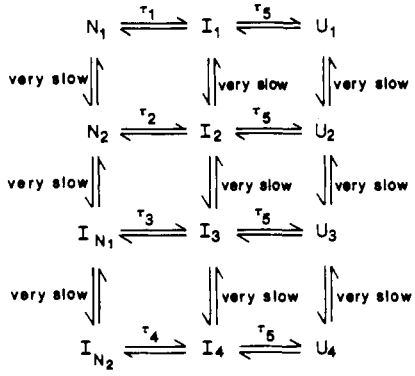
Scheme I




In this model, $N_1$ and $N_2$ are two native forms that are populated prior to the unfolding jump, and U is the unfolded form. The interconversion between the two native species is sufficiently slow that each unfolds independently to the unfolded form. The urea dependence of both the $\tau_1$ and $\tau_2$ relaxation times is that expected for folding reactions, and the lack of dependence of the amplitudes on the initial urea concentration can be accounted for by the inability of urea to perturb the $N_1 \leftrightarrow N_2$ equilibrium. The loss of the $\tau_1$ phase in unfolding in the presence of 400 mM KCl could be explained either by a shift in the $N_1 \leftrightarrow N_2$ equilibrium to $N_2$ or by an increase in the rate of interconversion between $N_1$ and $N_2$ so as to allow unfolding to proceed by the more rapid, $\tau_2$, channel.

Evidence for multiple native forms in DHFR that interconvert by slow processes has been obtained both from NMR

Scheme II




experiments and from studies on the kinetics of binding of substrate and cofactor. The $^{13}C$ NMR spectrum of $^{13}C$-labeled Trp residues in DHFR from *Streptococcus faecium* shows that one of the Trp residues, tentatively Trp-22, can exist in two different conformations that interconvert with a relaxation time much greater than 18 ms (Groff et al., 1981). It was suggested that the motion involved is the cis/trans isomerization of the adjacent Pro residue at position 21. The kinetics of binding of substrate or cofactor to DHFR (Cayley et al., 1981) show a rapid, bimolecular reaction followed by a slow, unimolecular reaction. This result was explained by the slow interconversion of two native forms, only one of which can bind substrate or cofactor. The relaxation time of the interconversion of these two forms at 25 °C, pH 7.0 in 0.5 M KCl can be estimated from the data as 28 s. This value is considerably faster than that required to permit independent unfolding channels. One possible explanation is that the two experiments are detecting different native conformations. Another possibility is that the different conditions, e.g., the presence of cofactor, substrate, or 0.5 M KCl in the binding studies of Cayley et al. (1981), alter the rate of interconversion and/or the equilibrium distribution. Further experiments are required to determine the relationship between the native conformations detected by these two experiments and those detected in folding studies.

(*2*) *Refolding Reaction.* The refolding reaction for DHFR is considerably more complex than unfolding. In addition to the two slow phases observed in unfolding, three faster phases were detected. The assignment of these phases to particular steps in a folding mechanism is assisted by comparison with the MTX binding kinetics. The decrease in absorbance at 380 nm for the $\tau_1$–$\tau_4$ phases in the latter experiment is assumed to reflect the appearance of conformations that are sufficiently nativelike to bind the inhibitor. The burst phase ($\tau_5$) observed in protein folding must reflect a different type of conformational change because the change in fluorescence intensity is opposite to that expected for a simple, progressive folding reaction and because it was not detected in MTX binding experiments.

A folding model which incorporates these results is proposed in Scheme II. In this model, $U_1$, $U_2$, $U_3$, and $U_4$ are a series of unfolded forms that interconvert by reactions that are very slow compared to the folding to a corresponding series of intermediates, $I_1$–$I_4$. These intermediates then fold to the native forms $N_1$ and $N_2$ and two nativelike intermediates $I_{N_1}$ and $I_{N_2}$. The description of these latter two forms as nativelike is based upon their ability to bind MTX. Their classification as intermediates and not actual native forms is based upon the failure to detect their presence in unfolding experiments beginning in the native base-line region. As is the case for the unfolded forms, the intermediates must also interconvert more slowly than any of the folding steps, $\tau_1$–$\tau_4$. In effect, the protein appears to refold by four independent channels.

The $\tau_1$–$\tau_4$ relaxation times correspond to the folding of the intermediates to their respective native or nativelike forms. The $\tau_5$ burst phase is presumed to reflect the folding of the unfolded forms to their respective intermediates. The inverse amplitude of this phase suggests that the unfolded forms may collapse to compact species in which the hydrophobic groups are preferentially found on the interior and the hydrophilic groups are exposed to solvent. The process of burying the Trp residues, which are highly hydrophobic, could transiently increase the fluorescence intensity by removing them from the quenching effect of the solvent. As the folding proceeds, the Trp residues would assume the exposure to solvent normally found in the native conformation, leading to decreases in fluorescence intensity. The observation of only a single burst phase may reflect the near-equality of an appropriate number of phases that link the unfolded and intermediate forms or the lack of dependence of such condensation reactions on the structural features that differentiate the unfolded forms. The inability to fit the burst phase to a single exponential above 1.5 M urea suggests the former explanation is correct. Similar rapid increases in fluorescence intensity upon refolding have also been reported for carbonic anhydrase (Stein & Henkens, 1978) and phosphoglycerate kinase (Betton et al., 1985), suggesting that this phenomenon may have some generality. Dill (1985) has recently advanced a theoretical model for folding that presumes an initial collapse of the unfolded protein to a hydrophobic aggregate. Refolding then proceeds from this condensed state to the native conformation.

The most plausible explanation for the very slow reactions linking the folding channels is cis/trans isomerization at X–Pro peptide bonds (Brandts et al., 1975). According to the Pro isomerization hypothesis as it was first proposed, unfolded species that contain nonnative isomers at X–Pro peptide bonds cannot fold until these peptide bonds isomerize to the isomer found in the native conformation. A later version of this hypothesis, based on folding studies of ribonuclease A, permits folding of species containing nonative Pro isomers to forms which are sufficiently nativelike to bind inhibitor and express enzymatic activity (Cook et al., 1979; Schmid & Blaschek, 1981). The nativelike intermediates, $I_{N_1}$ and $I_{N_2}$, proposed in the above model represent such species.

The inhibitor binding kinetics demonstrate that MTX is binding to native or nativelike forms at the end of the various folding pathways and not to intermediates along the pathways. The lack of dependence of the relaxation times on the MTX/DHFR ratio and the good agreement of these relaxation times with those determined from protein folding experiments support this conclusion.

Scheme II can also account for the unfolding kinetic data. The total amplitude detected in the $\tau_1$ and $\tau_2$ phases in unfolding matches that found in the equilibrium studies. Because the $\tau_1$ and $\tau_2$ reactions are rate limiting, neither the $I_1$ nor the $I_2$ species can accumulate during the course of unfolding.

A number of different theoretical models for protein folding assume that local structures appear early in the folding process. These structures are of marginal stability and are rapidly formed and dissipated. Elements of secondary structure such as helices, $\beta$ strands, turns, or hydrophobic groups (Ptitsyn & Finkelstein, 1980) have been suggested as candidates. These local structures then coalesce into larger, more stable species by a growth–merge (Go, 1983), cluster (Kanehisa & Tsong, 1979), or diffusion–collision (Karplus & Weaver, 1976) mechanism. Continuation of this process leads to even larger organized structures and, ultimately, the native conformation.

Rate-limiting steps in such a folding scheme might be expected to occur nearer the native form since the conversion of late intermediates to native could require the disruption of significant secondary and tertiary structure and a correspondingly high activation energy. The kinetic scheme proposed for the folding of DHFR, a fast collapse of unfolded forms to intermediates which then slowly convert to native forms, follows such a model. A similar scheme has been proposed for bovine pancreatic trypsin inhibitor (Creighton, 1977), penicillinase (Robson & Pain, 1976; Carrey & Pain, 1978), the $\alpha$ subunit of tryptophan synthase (Crisanti & Matthews, 1981), the constant-region fragment of an immunoglobin light chain (Goto & Hamaguchi, 1982), and hen egg white lysozyme and $\alpha$-lactalbumin (Kuwajima et al., 1985), indicating that it may have some generality.

## ACKNOWLEDGMENTS

We thank Dr. David Baccanari for his assistance with the purification of DHFR, Dr. J. M. Calvo for the generous gift of plasmid pCV29, Dr. Kenneth A. Johnson and Dr. Neil Tweedy for their help with the stopped-flow studies, and Dr. Robert Baldwin and Dr. Peter Kim for their helpful comments. We also thank Gail Feldman for typing the manuscript.

**Registry No.** DHFR, 9002-03-3.

## REFERENCES

Anfinsen, C. B. (1973) *Science (Washington, D.C.) 181*, 223.
Baccanari, D., Phillips, A., Smith, S., Sinski, D., & Burchall, J. (1975) *Biochemistry 14*, 5267.
Baccanari, D., Averett, D., Briggs, C., & Burchall, J. (1977) *Biochemistry 16*, 3566.
Baccanari, D., Stone, D., & Kuyper, L. (1981) *J. Biol. Chem. 256*, 1738.
Baccanari, D. P., Daluge, S., & King, R. W. (1982) *Biochemistry 21*, 5068.
Betton, J.-M., Dismadril, M., & Yon, J. M. (1985) *Biochemistry 24*, 4570.
Blakley, R. L. (1984) in *Folates and Pteridines* (Blakley, R. L., & Benkovic, S. J., Eds.) Vol. 1, pp 191–253, Wiley, New York.
Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C., & Kraut, J. (1982) *J. Biol. Chem. 257*, 13650.
Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953.
Carrey, E. A., & Pain, R. H. (1978) *Biochim. Biophys. Acta 533*, 12.
Cayley, P. J., Dunn, S. M. J., & King, R. W. (1981) *Biochemistry 20*, 874.
Chen, J.-T., Mayer, R. J., Fierke, C. A., & Benkovic, S. J. (1985) *J. Cell. Biochem. 29*, 73.
Cook, K. M., Schmid, F. X., & Baldwin, R. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6157.
Craik, C. S. (1985) *Biotechniques 3*, 12.
Creighton, T. E. (1977) *J. Mol. Biol. 113*, 295.
Crisanti, M. M., & Matthews, C. R. (1981) *Biochemistry 20*, 2700.
Dawson, R. M. C., Elliot, D. C., Elliot, W. M., & Jones, K. M. (1969) in *Data for Biochemical Research*, p 206, Oxford University Press, Oxford.
Dill, K. A. (1985) *Biochemistry 24*, 1501.
Donovan, J. W. (1973) *Methods Enzymol. 27*, 497.
Filman, D. J., Bolin, J. T., Matthews, D. A., Hamlin, R. C., & Kraut, J. (1982) *J. Biol. Chem. 257*, 13663.
Gō, N. (1983) *Annu. Rev. Biophys. Bioeng. 12*, 183.
Goto, Y., & Hamaguchi, K. (1982) *J. Mol. Biol. 156*, 891.
Groff, J. P., London, R. E., Cocco, L., & Blakley, R. L. (1981) *Biochemistry 20*, 6169.
Hagerman, P. J., & Baldwin, R. L. (1976) *Biochemistry 15*, 1462.
Hillcoat, B. L., Nixon, P. F., & Blakley, R. L. (1967) *Anal. Biochem. 21*, 178.
Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., & Kraut, J. (1986) *Science (Washington, D.C.) 231*, 1123.
Kanehisa, M. I., & Tsong, T. Y. (1979) *Biopolymers 18*, 1375.
Karplus, M., & Weaver, D. L. (1976) *Nature (London) 260*, 404.
Kelley, R. F., Wilson, J., Bryant, C., & Stellwagen, E. (1986) *Biochemistry 25*, 728.
Kimber, B. J., Griffiths, D. V., Birdsall, B., King, R. W., Schudder, J., Feeney, J., Roberts, G. C. K., & Burgen, A. S. V. (1977) *Biochemistry 16*, 3492.
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488.
Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874.
Ptitsyn, O. B., & Finkelstein, A. V. (1980) in *Protein Folding* (Jaenicke, R., Ed.) pp 101–115, Elsevier/North-Holland Biomedical Press, New York.
Robson, B., & Pain, R. J. (1976) *Biochem. J. 155*, 331.
Schellman, J. A. (1978) *Biopolymers 17*, 1305.
Schmid, F. X., & Blaschek, H. (1981) *Eur. J. Biochem. 114*, 111.
Smith, D. R., & Calvo, J. M. (1980) *Nucleic Acids Res. 8*, 2255.
Stein, P. J., & Henkens, R. W. (1978) *J. Biol. Chem. 253*, 8016.
Stone, D., Phillips, A. W., & Burchall, J. J. (1977) *Eur. J. Biochem. 72*, 613.
Stone, S. R., & Morrison, J. F. (1984) *Biochemistry 23*, 2753.
Tanford, C., Aune, K. C., & Ikai, A. (1973) *J. Mol. Biol. 73*, 185.
Villafranca, J. E., Hovell, E. E., Voet, D. H., Strogel, M. S., Ogden, R. C., Abelson, J. N., & Kraut, J. (1983) *Science (Washington, D.C.) 222*, 874.
Waheed, A., Qasim, M. A., & Salahuddin, A. (1977) *Eur. J. Biochem. 76*, 383.
Wong, K. P., & Tanford, C. (1973) *J. Biol. Chem. 248*, 8518.
Yanari, S., & Bovey, F. A. (1960) *J. Biol. Chem. 235*, 2818.