Transient Intermediates in the Folding of Dihydrofolate Reductase As Detected by Far-Ultraviolet Circular Dichroism Spectroscopy†

Kunihiro Kuwajima,‡ Edward P. Garvey,§,∥ Bryan E. Finn,§,⊥ C. Robert Matthews,*,§ and Shintaro Sugai‡,#

Department of Polymer Science, Faculty of Science, Hokkaido University, Kita-ku, Hokkaido 060, Japan, and Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

Received January 29, 1991; Revised Manuscript Received April 16, 1991

ABSTRACT: The kinetics of the reversible folding and unfolding of *Escherichia coli* dihydrofolate reductase have been studied by stopped-flow circular dichroism in the peptide region at pH 7.8 and 15 °C. The reactions were induced by concentration jumps of a denaturant, urea. The method can detect various intermediates transiently populated in the reactions although the equilibrium unfolding of the protein is apparently approximated by a two-state reaction. The results can be summarized as follows. (1) From transient circular dichroism spectra measured as soon as the refolding is started, a substantial amount of secondary structure is formed in the burst phase, i.e., within the dead time of stopped-flow mixing (18 ms). (2) The kinetics from this burst-phase intermediate to the native state are multiphasic, consisting of five phases designated as τ_1, τ_2, τ_3, τ_4, and τ_5 in increasing order of the reaction rate. Measurements of the kinetics at various wavelengths have provided kinetic difference circular dichroism spectra for the individual phases. (3) The τ_5 phase shows a kinetic difference spectrum consistent with an exciton contribution of two aromatic residues in the peptide CD region. The absence of the τ_5 phase in a mutant protein, in which Trp 74 is replaced by leucine, suggests that Trp 74 is involved in the exciton pair and that the τ_5 phase reflects the formation of a hydrophobic cluster around Trp 74. From the similarity of the kinetic difference spectrum to the difference between the native spectra of the mutant and wild-type proteins, it appears that Trp 47 is the partner in the exciton pair and that the structure formed in the τ_5 phase persists during the later stages of folding. (4) The later stages of folding show kinetic difference spectra that can be interpreted by rearrangement of secondary structure, particularly the central β sheet of the protein. The pairwise similarities in the spectrum between the τ_3 and τ_4 phases, and between the τ_1 and τ_2 phases, also suggest the presence of two parallel folding channels for refolding. (5) The unfolding kinetics show three to four phases and are interpreted in terms of the presence of multiple native species. The total ellipticity change in kinetic unfolding reaction, however, agrees with the ellipticity difference between the native and unfolding states, indicating the absence of the burst phase in unfolding. These results on the formation of secondary structure are compared with those on the tertiary structure obtained in previous studies by absorption and fluorescence spectroscopies and provide a more detailed picture of the folding of dihydrofolate reductase.

Kinetic studies of protein folding reactions provide access to transient intermediates which play key roles in folding but which are not highly populated at equilibrium. Until recently, most of the information available on events in the millisecond time range involved the development of tertiary structure. The sensitivity of aromatic chromophores, especially tyrosine and tryptophan, to their environments and the availability of commercial stopped-flow absorbance and fluorescence instruments were largely responsible for this focus. While data on the formation of tertiary structure is certainly critical, a complete understanding of how the amino acid sequence of a protein determines its three-dimensional structure also requires information on the formation of the secondary structure. This requirement is emphasized by the framework model of folding, which proposes that elements of secondary structure form early and direct subsequent events in folding (Kim & Baldwin, 1982; 1990).

The development of both stopped-flow circular dichroism (CD) spectroscopy (Bayley, 1981; Pflumm et al., 1986; Labhardt, 1986; Kuwajima et al., 1987) and hydrogen-exchange pulse-labeling NMR[1] spectroscopy (Udgaonkar & Baldwin, 1988; Roder et al., 1988) now permits access to information on the formation of secondary structure in the millisecond time range. The NMR experiment has the advantage of being able to monitor the formation of stable hydrogen bonds at individual, assignable amide protons. However, it suffers from the disadvantage that only amide hydrogens which are slowly exchanging in the native conformation, e.g., those that are involved in stable hydrogen bonds, can be examined. Nonnative secondary structure which might appear during folding would be very difficult to detect. The CD technique is capable of detecting nonnative secondary structure; however, the decomposition of the spectrum into individual contributions from

† This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (61420050 and 01580258) to K.K. and S.S., a grant from the Kurata Foundation to K.K., a grant from the National Science Foundation (DMB-9004707) to C.R.M., and a National Research Service Award from the National Institutes of Health (GM11936) to E.P.G.

* Author to whom correspondence should be adressed.

‡ Hokkaido University.

§ The Pennsylvania State University.

∥ Present address: Division of Experimental Therapy, Burroughs-Wellcome, Research Triangle Park, NC 27709.

⊥ Present address: Department of Physical Chemistry II, Chemical Center, University of Lund, S-221 00 Lund, Sweden.

Present address: Institute of Life Science, Soka University, Tangicho, Hachoji, Tokyo 192, Japan.

[1] Abbreviations: CD, circular dichroism; DHFR, dihydrofolate reductase; $NaDodSO_4$, sodium dodecyl sulfate; NMR, nuclear magnetic resonance; RMS, root mean square error; UV, ultraviolet; V75Y, Val 75 → Tyr; W74L, Trp 74 → Leu.

0006-2960/91/0430-7693$02.50/0

helix, strand, turn, and coil is not without ambiguity, and the assignment to specific residues is impossible. Thus, the two methods provide complementary information on the formation of secondary structure during folding.

Dihydrofolate reductase (DHFR) is an excellent system to which these techniques can be applied. DHFR from *Escherichia coli* is a monomeric protein of 159 amino acids arranged in an α/β motif. The X-ray structure shows that eight β strands form a hydrophobic structural core upon which four α helices, two 3_{10} helices, and eight hairpin turns dock (Bolin et al., 1982). Because over 85% of the sequence is involved in these helices, strands, and turns, DHFR has the potential to provide a detailed view of the formation of secondary structure during folding.

Previous studies on the equilibrium and kinetic properties of the urea-induced unfolding and refolding reactions have employed a combination of absorbance and fluorescence spectroscopy to monitor the behavior of the five tryptophan residues (Touchette et al., 1986; Perry et al., 1989). The first detectable event in refolding occurs in the 200–500-ms time range and corresponds to the burial of a single tryptophan, Trp 74, in a hydrophobic pocket (Garvey et al., 1989). Subsequent slow reactions lead to nativelike intermediates or native forms which are capable of binding methotrexate, a competitive inhibitor (Touchette et al., 1986), or which are enzymatically active (Frieden, 1990).

To address the issue of the formation of secondary structure during folding, we now report the results of a stopped-flow CD study on DHFR. A folding intermediate with significant secondary structure appears within the dead time of mixing (18 ms). Subsequent slower reactions involve the rearrangement of the polypeptide backbone as well as the development of tertiary structure previously detected by tryptophan fluorescence measurements (Touchette et al., 1986).

Materials and Methods

Chemicals. Urea and 2-mercaptoethanol were of specially prepared reagent grade from Nacalai Tesque, Inc. (Kyoto). The stock solution of urea was deionized on a mixed-bed column of Amberlite IR-120B and IRA-402 resin and used within a day. The concentration of urea was determined from the refractive index at 589 nm (Pace, 1986). All other chemicals were of guaranteed reagent grade.

Protein Source and Purification. (A) Construction of the W74L Mutant DHFR. The plasmid containing the W74L mutation in DHFR, pBF2/W74L, was constructed by using site-directed mutagenesis on plasmid pBF2. pBF2 is a derivative of plasmid pTZ19R (Mead et al., 1986) in which the wild-type *E. coli* DHFR gene has been inserted at the *Hin*dIII site behind the T7 promoter. The uracil-phage method of Kunkel (Kunkel et al., 1987) was used for mutagenesis. An appropriate oligonucleotide coding for the desired TGG to TTG codon change was employed. Mutants were selected by direct sequencing of plasmid purified from individual colonies by using the Sequenase (US Biochem) DNA sequencing kit. The entire gene of the mutant DHFR was sequenced to confirm that no other changes occurred.

(B) Purification of Wild-Type and Mutant DHFR. Wild-type DHFR was isolated from *E. coli* strain MV 1190 containing either plasmid pBF2 or pTP64-1. pTP64-1 is a plasmid containing the gene for wild-type DHFR under control of a consensus promoter and an engineered Shine–Dalgarno sequence (M. Iwakura, personal communication). The W74L mutant protein was purified from strain MV 1190 containing plasmid pBF2/W74L (see above). Wild-type and W74L DHFRs were purified according to the method of Baccanari (Baccanari et al., 1981) with the following modifications: (1) The elution of the protein from the methotrexate–agarose resin with 0.2 M potassium borate and 1 M KCl, pH 9.0, was omitted. (2) Chromatography on the LKB AcA 54 column was omitted. (3) Folate was removed by chromatography on a DEAE-Sephacel column (Pharmacia/LKB) equilibrated in 10 mM Tris-HCl, pH 7.2, by using a logarithmic, 0–0.5 M KCl gradient.

Purity was determined by the observation of a single band on $NaDodSO_4$ and native polyacrylamide gels. The concentration of wild-type DHFR was determined by using a molar extinction coefficient at 280 nm (ϵ_{280}) of 3.11×10^4 M^{-1} cm^{-1} (D. Baccanari, personal communication). The ϵ_{280} for the W74L mutant was determined by titration of the protein with methotrexate of a known concentration and was found to be 2.85×10^4 M^{-1} cm^{-1} (ϵ_{280} for wild-type DHFR was 3.15×10^4 M^{-1} cm^{-1} according to this method). The specific activities of the proteins were determined by using the method of Hillcoat et al. (1967). The activities of the wild-type and W74L mutant DHFRs were 90 and 54 units/mg, respectively.

Equilibrium and Kinetic Measurements of CD Spectra. All measurements were carried out at pH 7.8 and 15 °C in the presence of 10 mM potassium phosphate and 1 mM 2-mercaptoethanol. The protein solutions also contained the indicated amounts of urea as a denaturant. All the CD data were reported as mean residue ellipticity by taking 111 g/mol as the mean residue molecular weight.

Equilibrium CD spectra were taken on a Jasco J-500A spectropolarimeter as described previously (Kuwajima et al., 1985). The path length of the optical cuvette was 5.0, 1.0, or 0.1 mm, depending on the protein concentration. The peptide spectra below 250 nm were investigated.

Kinetic refolding and unfolding reactions were induced by concentration jumps of urea using a stopped-flow apparatus attached to the Jasco J-500A spectropolarimeter (Kuwajima et al., 1987, 1988). For the 1:10 mixing ratio of the protein and the diluent solutions used throughout this study, the dead time of mixing was 18 ms. The reaction kinetics were followed by the ellipticity in the peptide region. The path length of the optical cuvette was 3.7 mm, and the protein concentration in the final solution was 0.12 mg/mL in all the kinetic experiments.

Computer Fitting of Kinetic Data. The kinetic data were fit by the method of nonlinear least-squares to the equation:

$$\theta(t) = \theta(\infty) - \sum \Delta\theta_i \exp(-t/\tau_i) \qquad (1)$$

where $\theta(t)$ and $\theta(\infty)$ are the observed ellipticities at time t and infinite time, respectively; $\Delta\theta_i$ is the amplitude in the ith phase; and τ_i is the associated apparent relaxation time. Iterative computer calculations using the nonlinear least-squares fitting procedure were carried out until $\theta(\infty)$, each $\Delta\theta_i$, and the τ_i values all converged. The nonlinear least-squares fitting program was written in Microsoft FORTRAN 77 (by K.K.), using an algorithm based on the Gauss–Newton procedure, and was run on an NEC 9801 personal computer.

For the refolding kinetic traces of wild-type DHFR, the computer calculations were made in most cases by fixing the relaxation times, τ_i, to those values previously determined by absorption and fluorescence studies of the tryptophan residues (Touchette et al., 1986). The computer fits thus provided estimates of the associated amplitudes, $\Delta\theta_i$. In these cases, eq 1 is reduced to the linear form and the calculation converged after a single step.

Computer Fitting of Peptide CD Spectra. Curve fitting of the peptide CD data was carried out for the equilibrium states and also for a burst-phase transient intermediate in refolding

detected by stopped-flow CD measurements (see below). The observed ellipticity θ_{obs} at any fixed wavelength can be expressed by a linear combination of the ellipticity values of pure structural components. For the three-component analysis, in which the peptide backbone structure is assumed to be classified into the three components, α helix, β structure, and irregular structure, θ_{obs} is represented by the weighted values, f_i, of the appropriate reference spectra as

$$\theta_{obs} = f_\alpha\theta_\alpha + f_\beta\theta_\beta + f_c\theta_c \quad (2)$$

For the four-component analysis, in which β turn is also included as a pure structural component in the calculation, θ_{obs} is given by

$$\theta_{obs} = f_\alpha\theta_\alpha + f_\beta\theta_\beta + f_t\theta_t + f_c\theta_c \quad (3)$$

where f_t and θ_t are the fraction and the reference ellipticity of β turn. Linear least-squares calculations were made in an NEC 9801 personal computer with constraints of $f_\alpha + f_\beta + f_c = 1$ for the three-component analysis and $f_\alpha + f_\beta + f_c + f_t = 1$ for the four-component analysis to obtain the best-fit values for the fractions of secondary structure. The reference CD curves used were the data of Greenfield and Fasman (1969) and Chen et al. (1974) for the three-component analysis and Chang et al. (1978) for the four-component analysis. The average helical length used in the reference spectra of α helix was 11 residues for the latter two methods.

Computer Fitting of Kinetic Difference CD Spectra. The kinetic difference CD spectrum for each phase in refolding is given by the wavelength dependence of $\Delta\theta_i$ in eq 1. The fitting of the difference CD spectrum was also carried out for each kinetic phase, to obtain an estimate of the fractional change of each secondary structure component which occurs in that phase. The observed ellipticity change $\Delta\theta_{obs}$ at any fixed wavelength is given by equations analogous to eqs 2 and 3. For the three-component analysis

$$\Delta\theta = \Delta f_\alpha\theta_\alpha + \Delta f_\beta\theta_\beta + \Delta f_c\theta_c \quad (4)$$

and for the four-component analysis

$$\Delta\theta = \Delta f_\alpha\theta_\alpha + \Delta f_\beta\theta_\beta + \Delta f_t\theta_t + \Delta f_c\theta_c \quad (5)$$

where Δf_α, Δf_β, Δf_t, and Δf_c are the fractional changes of the secondary structure. The constraints in the linear least-squares calculations for the difference spectra were $\Delta f_\alpha + \Delta f_\beta + \Delta f_c = 0$ and $\Delta f_\alpha + \Delta f_\beta + \Delta f_c + \Delta f_t = 0$ for the three- and four-component analyses, respectively. The same reference CD spectra were used as in the fitting of the peptide CD spectra, described above.

Results

Equilibrium CD Spectra. The peptide CD spectra of DHFR in the native state and in the unfolded state in 7 M urea are shown in Figure 1. As the refolding kinetics of the W74L mutant will be compared with those of the wild-type protein (see below), the spectra of the mutant protein are also presented. The spectrum of the native wild-type protein has a broad negative band at 220 nm, a crossover to positive ellipticity at 203 nm, and a positive peak at 195 nm. Although the mutant also shows intense negative ellipticity between 211 and 235 nm, the detailed spectral features in the native state are different from those of wild-type DHFR. The spectrum is less intense between 211 and 227 nm and more intense above 227 nm for the mutant. If both the wild-type and mutant proteins have the same backbone structures in the native state, such differences in the peptide spectra must be due, at least in part, to an aromatic contribution of Trp 74 that is absent when replaced by leucine in the mutant. This tryptophan residue may be in an asymmetric environment in the native

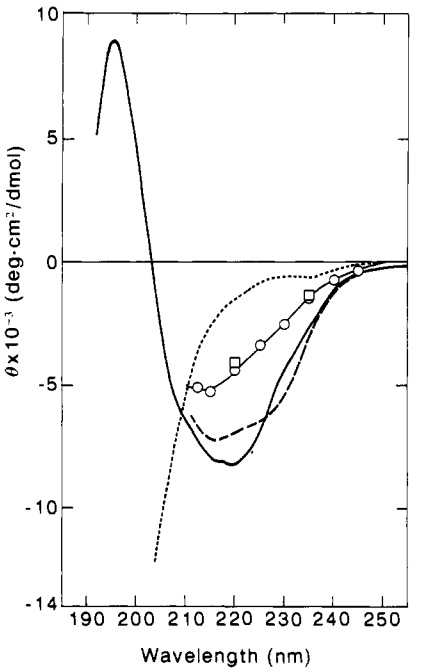


FIGURE 1: Peptide CD spectra of DHFR (pH 7.8 and 15 °C). (—) Native state of the wild-type protein; (– –) native state of the W74L mutant; and (- - -) unfolded state in 7 M urea for both the wild-type and the mutant proteins. Circles and squares are $\theta(0)$ values observed in refolding at 0.4 M urea for the wild-type and the mutant proteins, respectively.

structure of the wild-type protein brought about by the acquisition of the tertiary structure. In fact, the spectra of the two proteins are essentially identical with each other in 7 M urea (Figure 1), indicating that the differences in ellipticity between the proteins disappear after unfolding.

The fractions of secondary structure in the different conformational states of DHFR were estimated from the spectra in Figure 1 by the methods of Greenfield and Fasman (1969), Chen et al. (1974), and Chang et al. (1978) [see also Yang et al. (1986)]. The latter two methods give results that most closely agree with the secondary structure observed for the native conformation of the wild-type protein. The fits for the native conformation of the wild-type protein were made to data collected between 192 and 250 nm. For Chen's method, $f_\alpha = 22 \pm 2\%$, $f_\beta = 43 \pm 5\%$, and $f_c = 35 \pm 3\%$; for Chang's method, $f_\alpha = 26 \pm 2\%$, $f_\beta = 31 \pm 6\%$, $f_t = 20 \pm 3\%$, and $f_c = 23 \pm 1\%$; and from the X-ray structure (Bolin et al., 1982), $f_\alpha = 31\%$ including two 3_{10} helices, $f_\beta = 35\%$, $f_t = 20\%$, and $f_c = 14\%$. Note that the differences between the predicted and observed secondary structure must reflect in part the omission of the Trp side-chain contribution described above.

The estimation of the fractions of secondary structure in the unfolded forms and in transient folding intermediates is made difficult by the presence of urea and by the necessity of having to use a low protein concentration, 0.12 mg/mL, to avoid aggregation during refolding. Because the fractions depend not only on the spectral shape but also significantly on the wavelength region chosen [see Kuwajima et al. (1985) and Johnson (1990)], the calculations for these species were restricted to data in the range from 212.5 to 250 nm. The results obtained by the three methods are summarized in Table I. These structural fractions are only considered to be qualitative in discussions of the structure in a single conformation. However, the structural fractions for the native conformation of the wild-type protein do not differ appreciably from the fractions calculated from data to 192 nm by using Chang's method (see above).

Table I: Secondary Structure Fractions of DHFR Calculated from the Peptide CD Spectra by Three Sets of Reference CD Data (pH 7.8, 15 °C)[a]

	Greenfield and Fasman (1969)				Chen et al. (1974)				Chang et al. (1978)				
	f_α (%)	f_β (%)	f_c (%)	RMS[b]	f_α (%)	f_β (%)	f_c (%)	RMS	f_α (%)	f_β (%)	f_t (%)	f_c (%)	RMS
wild-type DHFR													
native state	8 ± 1	39 ± 2	53 ± 1	341	26 ± 1	16 ± 2	58 ± 2	601	21 ± 2	24 ± 3	18 ± 2	37 ± 2	445
unfolded state in 7 M urea	−7 ± 1	39 ± 3	68 ± 1	380	6 ± 1	18 ± 1	76 ± 1	344	5 ± 2	23 ± 3	32 ± 2	41 ± 2	340
burst-phase intermediate	−1 ± 2	41 ± 5	60 ± 3	291	15 ± 1	16 ± 4	70 ± 3	515	13 ± 3	20 ± 6	28 ± 4	40 ± 3	340
W74L mutant													
native state	17 ± 1	17 ± 3	66 ± 2	424	26 ± 1	5 ± 2	69 ± 2	586	26 ± 1	−3 ± 2	35 ± 2	42 ± 2	291

[a] CD data from 250 to 212.5 nm were used. [b] RMS values in deg·cm²/dmol.

Comparison of the fractions of β structure in the native conformations of the wild-type and W74L mutant proteins suggests that the β-sheet core is substantially disrupted in the mutant (Table I). Position 74 is the second residue in a short, three-residue strand at the edge of the sheet. Given that the mutant retains 60% of the activity of the wild-type protein, two more probable explanations for the difference are the aromatic contribution of Trp 74 discussed above and the error introduced by the omission of the spectral data below 212.5 nm.

The data in Table I also suggest that DHFR in 7 M urea may not be in a fully "random coil" form. Approximately 20% of the protein is predicted to be in the β form under conditions in which no further changes in absorbance, fluorescence, or circular dichroism are observed (Touchette et al., 1986). This high content of predicted β structure reflects the negative ellipticity between 210 and 220 nm and does not change appreciably when the spectrum measured to 203 nm is included in the calculation (data not shown). The reference spectra for the pure "random" form for all of the methods of analysis have more positive ellipticity in this region than observed for unfolded DHFR. In particular, the reference spectra in the three-component analyses such as Chen's method exhibit a positive CD band around 220 nm (Yang et al., 1986). The perspective adopted in this paper is that the β structure predicted in 7 M urea may be an artifact which arises from the choice of the reference spectrum for the random coil (see Discussion).

Folding and Unfolding of DHFR Studied by the Peptide CD Spectra. To determine suitable conditions for kinetic CD studies of the urea-induced unfolding and refolding of DHFR, the equilibrium unfolding transition of DHFR was first investigated by monitoring the ellipticity at 220 nm (Figure 2). As reported previously (Touchette et al., 1986; Perry et al., 1989), the unfolding begins at 2 M urea and is essentially complete at 4 M urea. The data are well described by a two-state model, N ↔ U, in which the free energy difference between the native, N, and unfolded, U, conformations depends linearly on the denaturant concentration:

$$\Delta G = \Delta G_{H_2O} - m[\text{urea}] \quad (6)$$

where ΔG is the apparent free energy difference at any urea concentration, ΔG_{H_2O} is the free energy difference in the absence of denaturant, and m is a parameter describing the cooperativity of the transition. At pH 7.8 and 15 °C, the midpoint of the transition is found to occur at 3.07 M urea, ΔG_{H_2O} = 7.31 ± 0.78 kcal/mol, and m = 2.38 ± 0.25 kcal/(mol·M). These results are essentially identical with those reported previously (Touchette et al., 1986), given the magnitudes of the errors, and are considered to be more accurate because of improvements in instrumentation.

(*a*) *Refolding Kinetics.* Figure 3 shows kinetic progress curves of refolding produced by a concentration jump of urea from 4.5 to 0.4 M and measured by θ_{220} and θ_{235} at pH 7.8

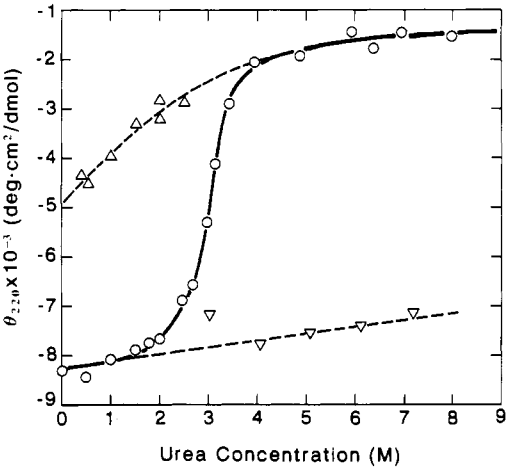


FIGURE 2: Unfolding transition of DHFR measured by the ellipticity at 220 nm as a function of urea concentration (pH 7.8 and 15 °C). Broken lines indicate the $\theta(0)$ values observed in kinetic refolding (△) and unfolding (▽) studies.

and 15 °C. Comparison of the results at these two wavelengths may highlight backbone and side-chain contributions to the signal and provides an internal check on the validity of the computer fits of the data. At both 220 and 235 nm, the kinetics are represented by multiphasic processes. At 220 nm the ellipticity first decreases sharply (within 1 s) and then more slowly as the reaction proceeds. At 235 nm a lag phase lasts until ~0.4 s, after which successive decreases in ellipticity are observed.

The kinetic curves were first fit to two- and three-exponential functions of eq 1 by using a nonlinear least-squares calculation. Both curves are reasonably well fit with a three-term equation (Table II). The apparent relaxation time of each phase in the three-exponential fit is, however, markedly different between the two curves; the relaxation times observed at 220 nm are 2- to 5-fold larger than the corresponding relaxation times at 235 nm. This inconsistency indicates that there must be more than three phases in the refolding process, even though the individual curves can be fit to three exponentials.

In a previous study of the refolding kinetics by a combination of absorbance and fluorescence spectroscopies (Touchette et al., 1986), it was found that at least five exponentials are required to obtain an adequate fit of the data. Given this precedent, the two refolding curves of Figure 3 were then fit to four and five exponentials. The calculations for the kinetic curve obtained at 235 nm converged with both four and five exponentials while the curve obtained at 220 nm did not show convergence in the nonlinear least-squares fitting procedure. The kinetic parameters for the 235-nm data are presented in Table II. The theoretical curves with four and five exponentials reproduce the observed curve at 235 nm better than that with three exponentials with respect to both RMS deviations (Table II) and residual plots (Figure 3). The apparent

Table II: Kinetic Parameters of Refolding of DHFR at 0.4 M Urea, pH 7.8, and 15 °C

	RMS (deg·cm²/dmol)	τ_1 (s)	τ_2 (s)	τ_3 (s)	τ_4 (s)	τ_5 (s)	$\Delta\theta_1$ (deg·cm²/dmol)	$\Delta\theta_2$ (deg·cm²/dmol)	$\Delta\theta_3$ (deg·cm²/dmol)	$\Delta\theta_4$ (deg·cm²/dmol)	$\Delta\theta_5$ (deg·cm²/dmol)	$\theta(0)$ (deg·cm²/dmol)	$\theta(\infty)$ (deg·cm²/dmol)
wild-type DHFR													
CD at 220 nm													
$n = 3$[a]	113	50.4	7.16	0.418			1004	484	1745			−4596	−7829
$n = 5$	115	(164	37	5.7	0.88	0.20)[b]	180	1005	258	837	1220	−4364	−7864
CD at 235 nm													
$n = 3$	37	23.1	1.40	0.143			389	1000	−257			−1451	−2584
$n = 4$	34	57.1	5.21	0.843	0.279		231	418	1041	−508		−1421	−2602
$n = 5$	34	165	29	4.3	0.78	0.30	112	191	414	1075	−583	−1419	−2628
abs and fluorescence[c]													
$n = 5$		164	37	5.7	0.88	0.20							
W74L mutant													
CD at 220 nm													
$n = 2$	191	58.0	7.20				1033	726				−4087	−5846
CD at 235 nm													
$n = 2$	51	60.9	7.25				877	496				−1366	−2739

[a] The number of phases denoted by n. [b] The relaxation times are fixed to the values obtained by the absorption and fluorescence kinetics. [c] The values of relaxation time are from Touchette et al. (1986).

Table III: Fractional Changes of Secondary Structure Calculated from the Kinetic Difference CD Spectra of the Five Phases in Refolding of DHFR by Three Sets of Reference CD Data (pH 7.8, 15 °C)[a]

	Greenfield and Fasman (1969)				Chen et al. (1974)				Chang et al. (1978)				
	Δf_α (%)	Δf_β (%)	Δf_c (%)	RMS[b]	Δf_α (%)	Δf_β (%)	Δf_c (%)	RMS	Δf_α (%)	Δf_β (%)	Δf_i (%)	Δf_c (%)	RMS
τ_1	−0.1 ± 0.7	1.1 ± 1.6	−1.0 ± 0.9	98	0.6 ± 0.2	−0.2 ± 0.7	−0.4 ± 0.6	95	0.1 ± 1.0	0.3 ± 2.0	−0.8 ± 1.4	0.4 ± 1.1	105
τ_2	−1.2 ± 1.0	6.4 ± 2.3	−5.3 ± 1.3	140	2.0 ± 0.3	2.9 ± 1.0	−4.9 ± 0.9	142	0.7 ± 1.5	7.3 ± 3.0	−5.3 ± 2.1	−2.7 ± 1.6	157
τ_3	4.8 ± 1.3	−8.5 ± 2.8	3.7 ± 1.6	172	2.0 ± 0.1	−4.2 ± 0.4	2.2 ± 0.4	60	2.4 ± 0.6	−8.6 ± 1.1	3.3 ± 0.8	2.8 ± 0.6	60
τ_4	9.2 ± 4.5	−15.7 ± 9.9	6.5 ± 5.5	605	4.5 ± 0.7	−9.5 ± 2.7	5.0 ± 2.4	379	6.9 ± 2.8	−23.4 ± 5.4	10.3 ± 3.7	6.2 ± 2.9	285
τ_5	−5.3 ± 4.2	15.0 ± 9.1	−9.7 ± 5.1	558	0.9 ± 1.0	7.7 ± 3.6	−8.6 ± 3.2	504	−6.6 ± 3.2	26.2 ± 6.2	−17.1 ± 4.3	−2.5 ± 3.3	329

[a] Kinetic CD data from 250 to 212.5 nm were used. [b] RMS values in deg·cm²/dmol.

relaxation times in the five-exponential fit are in good agreement with those obtained in the absorption and fluorescence kinetics (Table II).

Presuming that the failure of the five-exponential fit for the 220-nm data may be due to small amplitudes of some of the phases, the refolding curve at 220 nm was fit to a five-exponential function in which the relaxation times were fixed to the values obtained in the absorbance and fluorescence kinetic studies. For this five-exponential fit, the observed curve is reproduced as well as for the three-exponential fit using the criteria of RMS deviation and residual plots. This constrained approach also reproduced the kinetic curves at a variety of wavelengths (see below), suggesting that the refolding kinetics of DHFR measured by the peptide CD spectra are fully consistent with the kinetics measured by absorption and fluorescence spectra. The $\theta(\infty)$ values in the five-exponential fit are essentially identical with the ellipticity values in the native state (Figure 1), demonstrating the reversibility of the unfolding transition.

(*b*) *The Burst-Phase Intermediate in Refolding.* It can be seen in Table II that the $\theta(0)$ values in the five-exponential fit are remarkably smaller than the values for the unfolded protein, θ_U, in 4.5 M urea; $\theta(0) = -4364$ and $\theta_U = -2000$ deg·cm²/dmol at 220 nm and $\theta(0) = -1419$ and $\theta_U = -700$ deg·cm²/dmol at 235 nm. This difference is a clear indication that there is a burst phase in refolding occurring within the dead time of the stopped-flow measurement (18 ms). The peptide CD spectrum of the burst-phase species was investigated by measuring the refolding kinetics at various wavelengths from 212.5 to 245 nm. All of the kinetic curves at these different wavelengths were fit well with the five-exponential function in which the relaxation times were fixed to the values previously obtained in the absorption and fluorescence kinetics. The $\theta(0)$ values thus obtained represent the CD spectrum of the burst-phase intermediate and are plotted as a function of wavelength in Figure 1. Although the number of data points is limited, the comparison of this spectrum with that for the native conformation shows that the spectrum of the intermediate is not simply that of the native which has been reduced in intensity. It appears that this burst-phase species has a different mix of secondary structure and/or side-chain interactions than native. According to all three predictive schemes, this early intermediate has a lower α helix content and either a higher coil (Greenfield & Fasman, 1969; Chen, 1974) or turn (Chang, 1978) content (Table I).

(*c*) *Kinetic Difference CD Spectra.* The five-exponential fit of the refolding curves measured at various wavelengths provides the kinetic difference CD spectra of individual kinetic phases. The difference spectrum for the ith phase is given by the wavelength dependence of the amplitude, $\Delta\theta_i$ (eq 1). The difference spectra are related to the fractional changes of secondary and/or tertiary structure components, which occur during the respective kinetic phases. The difference spectra are shown in Figure 4, and the fractional changes of the secondary structure components for the five phases obtained by eqs 4 and 5 are summarized in Table III.

As shown by solid and broken lines in Figure 4, the theoretical curves drawn by the use of these parameters represent well the experimentally observed difference CD spectra except for the τ_5 phase. The amplitude of the difference spectra for the τ_1 phase is so small that it is difficult to assess the secondary structure changes occurring in this phase; the fractional changes presented in Table III are all within the standard errors. The theoretical fits of the difference spectrum are somewhat worse for the τ_5 phase, implying that a spectral contribution other than those from changes in the secondary structure may be significant in this phase. Among the other phases, both the τ_3 and τ_4 phases show difference spectra with

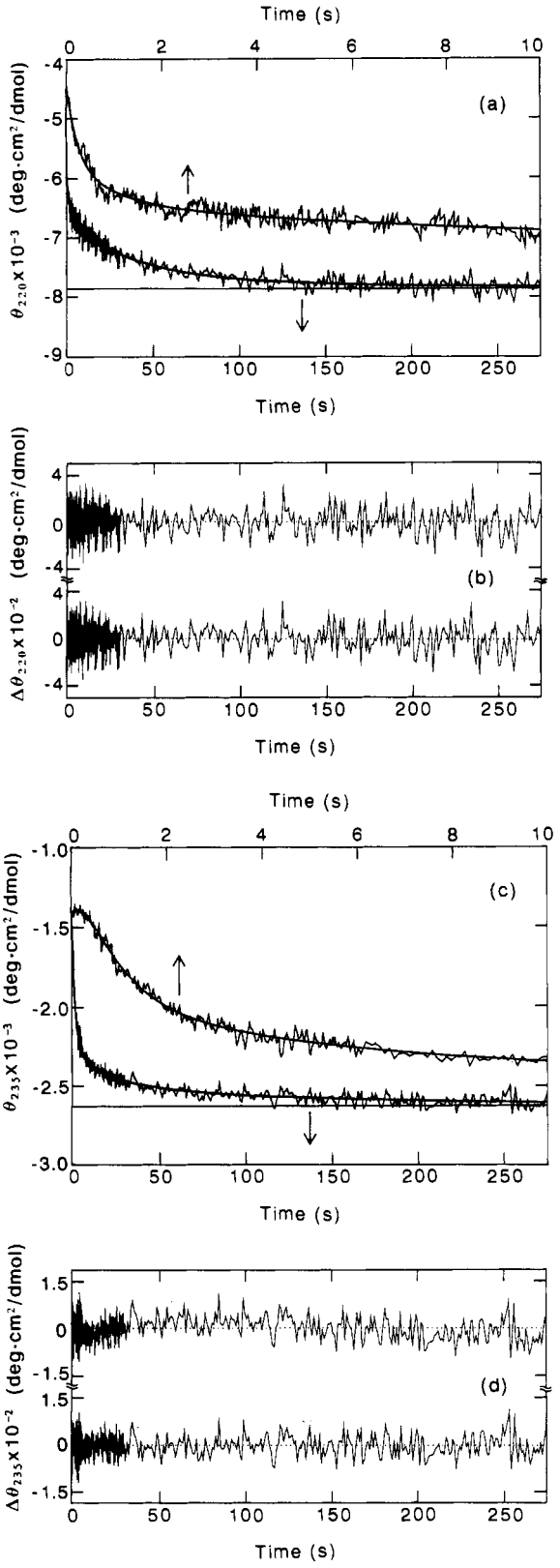


FIGURE 3: Kinetic refolding curves of DHFR measured by the ellipticity at 220 nm (a) and 235 nm (c), and their residual plots at 220 nm (b) and 235 nm (d) (pH 7.8 and 15 °C). The refolding was induced by a concentration jump of urea from 4.5 to 0.4 M. (a and c) The kinetics are shown in different time scales, and thick solid lines represent the theoretical curves by five-exponential fits with the parameter values shown in Table II. (b and d) Residual plots for fitting with three exponenials (upper) and five exponentials (lower).

a minimum around 230 nm. These are interpreted in terms of a decrease in f_β and concomitant increases in the other structural components, notably in α helices (Table III). For

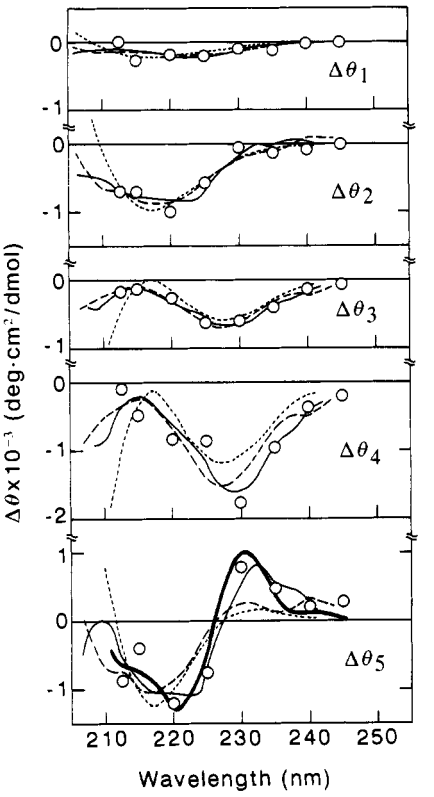


FIGURE 4: Kinetic difference CD spectra for the five phases in refolding of DHFR at 0.4 M urea, pH 7.8, and 15 °C. Observed difference spectra are shown by (O), and theoretical best-fit curves drawn by the parameter values of Table III with three sets of reference data are as follows: (---) Greenfield and Fasman (1969), (--) Chen et al. (1974), and (—) Chang et al. (1978). A thick solid line in the panel for $\Delta\theta_5$ shows the equilibrium difference between the native spectra of the wild-type and the W74L mutant proteins. The equilibrium difference spectrum coincides with the kinetic difference spectrum (see text).

the τ_2 phase, the difference spectrum has minimum around 220 nm and is interpreted in terms of an increase in f_β and concomitant decreases in f_t and f_c. These results thus demonstrate that there must be significant rearrangement between different structural components, particularly concerning the increase or decrease in f_β, during the late stages of refolding.

(d) Unfolding Transition of Burst-Phase Intermediate. Because the burst-phase reaction is at least an order of magnitude faster than subsequent refolding phases, the corresponding intermediate becomes fully populated before being depleted. As a result, the value of $\theta(0)$ is proportional to the concentration of the burst-phase species. The stability of this intermediate can be examined by monitoring the dependence of $\theta(0)$, i.e., the concentration, on the final urea concentration in refolding. Similar procedures have been used to study the stability of burst-phase species in α-lactalbumin, lysozyme, parvalbumin, and staphylococcal nuclease (Ikeguchi et al., 1986; Kuwajima et al., 1988; Sugawara et al., 1991). The results of such a study on DHFR at urea concentrations from 0.4 to 2.5 M are shown in Figure 2. The refolding transition curve for the burst-phase species has a much weaker dependence on the urea concentration than the equilibrium transition curve linking the native and unfolded conformations. Although the saturation curve shape precludes a quantitative analysis of the free energy difference between the intermediate and the unfolded form, the burst-phase species must have a lower stability than the native conformation.

(e) Unfolding Kinetics. Figure 5 shows a kinetic progress curve for the unfolding reaction produced by jumping the urea

Table IV: Kinetic Parameters of Unfolding of Wild-Type DHFR at 6.1 M Urea, pH 7.8, and 15 °C

	RMS (deg·cm²/dmol)	τ_1 (s)	τ_2 (s)	τ_3 (s)	$\Delta\theta_1$ (deg·cm²/dmol)	$\Delta\theta_2$ (deg·cm²/dmol)	$\Delta\theta_3$ (deg·cm²/dmol)	$\theta(0)$ (deg·cm²/dmol)	$\theta(\infty)$ (deg·cm²/dmol)
CD at 220 nm									
$n = 3$[a]	156	313	81.9	2.73	−552	−4247	−848	−7424	−1774
absorption at 293 nm[b]									
$n = 2$		310	85						

[a] The number of phases. [b] Touchette et al. (1986).

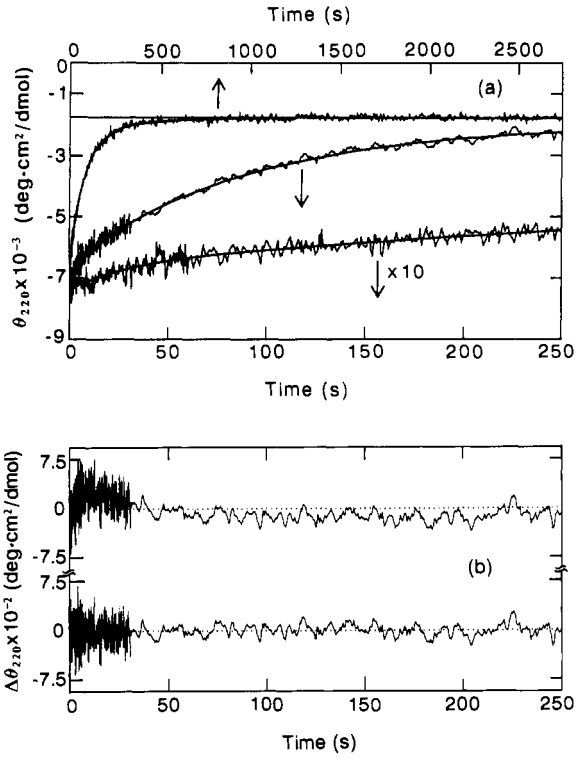


FIGURE 5: Kinetic unfolding curve of DHFR measured by the ellipticity at 220 nm (a) and its residual plots (b) (pH 7.8 and 15 °C). The unfolding was induced by a concentration jump of urea from 0 to 6.1 M. (a) The kinetics are shown in the three different time scales, 2750 s (upper), 250 s (middle), and 25 s (lower). Thick solid lines represent the theoretical curve with the parameter values of Table IV. (b) The residual plots for fitting with two exponentials (upper) and three exponentials (lower).

concentration from 0 to 6.1 M at pH 7.8 and 15 °C. The nonlinear least-squares analysis with eq 1 has shown that at least three exponentials are required for an adequate fit to the observed curve; the kinetic parameters obtained are summarized in Table IV. The values of τ_1 and τ_2 are in good agreement with those reported in the previous study of the kinetics measured by tryptophan absorption (Touchette et al., 1986). The observation of the τ_3 phase in unfolding is, however, different from the previous observation in which the entire change in absorption was accounted for by only the τ_1 and τ_2 phases. Although the amplitude of the τ_3 phase is only 15% of the total ellipticity change in unfolding, its neglect clearly makes the theoretical fit unsatisfactory (see the residual plots in Figure 5).

A recent reexamination of the unfolding of wild-type DHFR with stopped-flow fluorescence spectroscopy (Finn and Matthews, unpublished results) confirms the presence of this τ_3 phase and suggests that a fourth, minor phase also exists. This latter phase comprises only 5% of the total signal and may be difficult to detect in the CD experiment (Jennings, Finn, and Matthews, manuscript in preparation). The τ_3 phase in unfolding has been previously detected by fluorescence spectroscopy for the DHFR mutant in which valine at position 75 is replaced by tyrosine (Garvey & Matthews, 1989). The relaxation time for this phase is in good agreement with the value for the wild-type protein obtained in the CD study.

Using a three-phase fitting procedure, the unfolding kinetics of DHFR were investigated at various final urea concentrations from 3.0 to 7.2 M. The apparent relaxation times were found to decrease with increasing denaturant concentration in the same manner as reported previously for the wild-type and V75Y mutant proteins (data not shown) (Touchette et al., 1986; Garvey & Matthews, 1989). When the value of the ellipticity at zero time, $\theta(0)$, is plotted as a function of the urea concentration (Figure 2), this parameter is seen to vary linearly with urea concentration and be coincident with the linear extrapolation of the native base-line value. Therefore, in contrast with the refolding reaction, there is no burst phase occurring in the dead time of the measurement for the unfolding reaction to any urea concentration. The omission of a fourth exponential in the fitting procedure does not affect this conclusion because its amplitude is very small.

Refolding Kinetics of the W74L Mutant DHFR. The τ_5 phase in refolding detected by fluorescence has been assigned to the burial of a specific tryptophan residue, Trp 74, in a hydrophobic pocket (Garvey et al., 1989). The replacement of this residue with phenylalanine was shown to eliminate the τ_5 phase in fluorescence-detected refolding. Because the kinetic difference spectrum of the τ_5 phase detected by CD spectroscopy resembles that expected for aromatic side chains (Adler et al., 1973; Brahms & Brahms, 1980), it was of interest to determine whether the replacement of the Trp would also eliminate the τ_5 phase in CD. The introduction of leucine at this position removes any ambiguities arising from the benzyl side chain of the original phenylalanine mutant. Phenylalanine also has a substantial CD signal in this region.

The refolding of the W74L mutant was initiated by jumping the urea concentration from 4.5 to 0.4 M, conditions sufficient to bring the protein from the unfolded to the native conformation (Finn and Matthews, unpublished results). The reaction curves obtained by CD at 220 and 235 nm are shown in Figure 6. Compared with the refolding curves of the wild-type protein (Figure 3), the lag phase observed at 235 nm and the sharp decrease in ellipticity observed at 220 nm both disappear in the mutant (Figure 6). The refolding curves at both wavelengths can be well fit to two exponentials; the kinetic parameters are listed in Table II. In contrast to the previous analysis for the wild-type protein where either two or three exponentials gave different relaxation times at different wavelengths, the relaxation times of the mutant at 220 and 235 nm are the same within experimental error. Therefore, only two phases are sufficient to interpret the CD kinetics of the W74L mutant in the time range studied, i.e., up to 275 s. The slower relaxation time is between that of the τ_1 and τ_2 phases for the wild-type protein while that for the faster phase is close to the wild-type τ_3 phase (Touchette et al., 1986).

Parallel studies on this mutant using fluorescence spectroscopy show three phase in refolding. The slowest, τ_1, phase

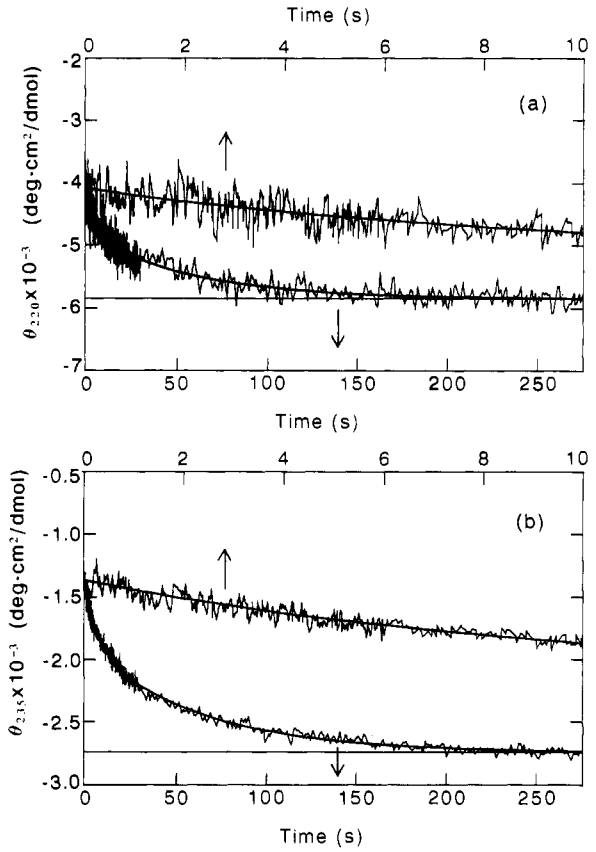


FIGURE 6: Kinetic refolding curves of the W74L mutant measured by the ellipticity at 220 nm (a) and 235 nm (b) (pH 7.8 and 15 °C). The refolding was induced by a concentration jump of urea from 4.5 to 0.4 M. The kinetics are shown in two different time scales; the thick solid lines represent the theoretical curves with parameter values shown in Table II.

occurs over a longer time range than was sampled by CD. The τ_2 and τ_3 phases have the same relaxation times as the two phases detected by CD spectroscopy (Finn and Matthews, unpublished results). The existence of a slower phase in CD is supported by the observation that the $\theta(\infty)$ calculated from the kinetics up to 275 s is significantly larger than the ellipticity of the native protein (−6900 deg·cm²/dmol) at 220 nm. The τ_5 phase seen in the wild-type protein, which is unique in that its fluorescence intensity increases in refolding, must be absent in the mutant. The one-to-one matching of the phases detected by CD with corresponding phases in fluorescence and the absence of the τ_5 phase in fluorescence are most simply interpreted in terms of the loss of the τ_5 phase in CD. Either the τ_3 or τ_4 phase is also lost in the W74L mutant protein.

The coincidence of the $\theta(0)$ values at 220 and 235 nm in refolding for the W74L mutant with the values for the wild-type (Figure 1) shows that both proteins form the same burst-phase species. Comparison of the kinetic difference CD spectrum of the τ_5 phase with the difference spectrum generated by subtracting the native spectrum of the mutant DHFR from that for the wild-type DHFR shows that the two difference spectra are essentially identical (Figure 4). The equivalence implies that the asymmetry in the environment which gives rise to the CD effect for Trp 74 appears in the τ_5 phase and persists unchanged into the native conformation.

Discussion

The stopped-flow CD spectroscopy in the peptide region detected various intermediates which are transiently populated during refolding of DHFR. The method has provided new information about changes in secondary structure occurring at the individual steps of kinetic refolding. The purpose of this study is to obtain a more detailed picture of the folding of DHFR by comparison of the results with those obtained from absorption and fluorescence spectroscopies which monitor principally changes in tertiary structure (Touchette et al., 1986; Perry et al., 1987; Garvey & Matthews, 1989; Garvey et al., 1989).

CD Spectral Analysis. Before discussing each step in refolding, it is important to consider the accuracy of secondary structure fractions predicted from the experimental CD spectra. This is especially important for the spectrum of the unfolded state. As documented by many researchers (Sears & Beychok, 1973; Woody, 1977; Cantor & Timasheff, 1982), there is no reason to believe that there is a unique CD spectrum which is characteristic of the unordered polypeptide. The spectrum in concentrated urea need not be identical with the "random" spectrum used in the CD analyses. The reference spectra for the pure "random" form have positive ellipticity between 210 and 220 nm. It has frequently been noted, however, that extensively unfolded proteins like DHFR in 7 M urea do not, generally, show a positive band in this region (Sears & Beychok, 1973) although a few exceptions have been reported (Perlmann & Grizzuti, 1971; Kuwajima et al., 1988). Two possible sources for the negative ellipticity in the unfolded state are (1) the presence of residual structure in 7 M urea and (2) the direct interaction of urea with peptide groups which may affect the optical activity. Independent evidence for the existence of some residual structure in unfolded DHFR is required before the result can be interpreted.

Some modern methods of the CD analysis require the CD data extending to the vacuum UV region (Brahms & Brahms, 1980; Hennessey & Johnson, 1981; Yang et al., 1986; Johnson, 1988). It has been suggested that accurate estimates for at most five different classes of secondary structure can be obtained only if data extend to 184 nm (Johnson, 1990). In this study, however, the presence of urea (≥0.4 M in refolding conditions) and the necessity of using a low protein concentration precluded CD measurement below 212.5 nm. Development of sufficiently negative ellipticity in the region between 212.5 and 245 nm and the consistency of the results among different analysis methods are thus used as indications for formation of ordered secondary structure. The limitation caused by truncation of the CD data at 212.5 nm should be kept in mind, however, when quantitative accuracy for estimates of structural changes is concerned.

The Burst-Phase Intermediate, a Common State at the First Stage in Protein Folding. A folding intermediate with significant secondary structure appears within the dead time of mixing of stopped-flow CD spectroscopy, i.e., within 18 ms. The far-UV spectrum of this burst phase species suggests that the intermediate has both α helix and β structure (Table I). In contrast, in the same time range there is no evidence for the formation of tertiary structure (Garvey & Matthews, 1989); no corresponding burst phase is detected by fluorescence spectroscopy.

The appreciable secondary structure, apparent absence of tertiary structure, and little if any cooperativity to urea-induced unfolding suggest that this intermediate may be a molten globule (Ptitsyn, 1987; Kuwajima, 1989). Burst-phase folding intermediates with similar stability properties have been observed in α-lactalbumin and lysozyme (Ikeguchi et al., 1986) and, more recently, in staphylococcal nuclease (Sugawara et al., 1991). For α-lactalbumin and carbonic anhydrase, the species populated at acidic pH or at moderate concentrations of guanidine hydrochloride has been shown to be identical with

the species appearing on the refolding pathway when the protein is diluted from high guanidine hydrochloride (Ikeguchi et al., 1986; Dolgikh et al., 1984). Burst-phase intermediates detected by stopped-flow CD have also reported for β-lactoglobulin, chymotrypsinogen A, ferricytochrome *c*, parvalbumin, and tryptophan synthase β_2 subunit (Kuwajima et al., 1987, 1988; Goldberg et al., 1990; Sugawara et al., 1991).

Formation of a Hydrophobic Core around Trp 74 in the τ_5 Phase of Refolding. The kinetic difference spectrum for the τ_5 folding phase in the wild-type protein does not resemble that expected for any of the standard elements of secondary structure. The poor fits of this spectrum by the secondary structure prediction programs are in line with this observation (Table III). Rather, the approximately equal intensities for the positive and negative bands centered at 228 nm suggest that this is the exciton spectrum for two aromatic residues in DHFR. Given that the splitting for an exciton falls off as R^{-3} (Cantor & Schimmel, 1980), this effect is generally limited to adjacent side chains in the three-dimensional structure. The loss of this kinetic phase in the W74L mutant shows that Trp 74 is one of the partners in the exciton. The identity of the other partner is suggested by the excellent agreement between this kinetic difference spectrum and that for the difference between the native forms of the wild-type and the W74L mutant proteins (Figure 4). The agreement suggests that the asymmetry which gives rise to this effect in the folding intermediate also exists in the native conformation. Tryptophans 47 and 74 are nearest neighbors in the native conformation and are potential candidates for an exciton pair.

The persistence of tertiary structure formed in the τ_5 folding phase is also observed in refolding as monitored by fluorescence spectroscopy (Garvey et al., 1989). In this case, the difference in the fluorescence amplitude (in photomultiplier tube voltage using a Durrum 110 stopped-flow spectrometer) between the native forms of the wild-type and another mutant of Trp 74, W74F (ΔA = 1.5 ± 0.2 V) is equal to the sum of the differences between the two proteins unfolded in 5.4 urea (0.9 ± 0.2 V) plus the amplitude of the τ_5 burst phase in the wild type (0.74 ± 0.05 V). This is consistent with an interpretation which postulates that the structure giving rise to this fluorescence contribution in the τ_5 phase persists in the native form.

Exciton splitting has also been suggested to appear in the spectrum of poly(L-tryptophan) [see Woody (1977)]. A similar difference CD spectrum arising from aromatic residues has been reported for the difference between the chymotrypsinogen and α-chymotrypsin far-UV CD spectra, for which a drastic change in the orientation of Trp 141 on activation of the zymogen is thought to be responsible (Cantor & Timasheff, 1982). Manning and Woody (1989) have invoked electronic coupling between aromatic chromophores in bovine pancreatic trypsin inhibitor to explain the far-UV CD spectrum.

The early association between Trp 47 and Trp 74 may reflect the folding of the entire polypeptide sequence between these two residues. It is striking that this region of DHFR corresponds to three parallel, neighboring β strands (B–C–D) and one intervening α helix (C) (Bolin et al., 1982). The present results suggest that this sequence may act as an autonomous folding unit, whose formation directs the folding of the remainder of the protein (Kim & Baldwin, 1982, 1990). Inspection of the crystal structures of α/β proteins shows that such three-stranded supersecondary structures are rather common (Rao & Rossmann, 1973; Richardson, 1977; Rossmann & Argos, 1981). Rao and Rossmann (1973) were among the first to recognize these units and coined the term "mononucleotide binding fold" to describe them. Thus, these supersecondary structures may play critical roles not only in the final native conformation but also in directing the folding of a whole class of proteins which contain them.

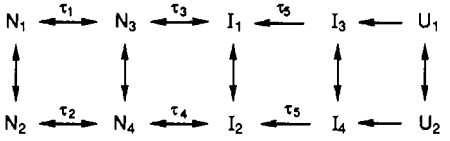
The intermediate which is the product of the τ_5 reaction was first detected by fluorescence spectroscopy and assigned to the burial of Trp 74 in a hydrophobic cluster (Garvey et al., 1989). If the above assignment of the τ_5 phase observed by CD spectroscopy is correct, then this contribution to the far-UV CD spectrum also reflects the development of tertiary, not secondary structure. Taken together, these results show that the formation of secondary structure in the burst phase precedes that for tertiary structure in the early stages of folding of DHFR. These results are consistent with the expectations of the framework model of folding (Kim & Baldwin, 1982, 1990). The observation that some of the interactions present in the native structure also appear in intermediates while others do not is consistent with a model for the later events in folding, the substructure model, proposed by Kuwajima (1989).

τ_1–τ_4 Folding Reactions. These reactions were originally detected by changes in absorbance and fluorescence and by methotrexate binding (Touchette et al., 1986). Thus, the tertiary structure is clearly altered in these rate-limiting steps in folding. The present study shows that the secondary structure is also changing in the τ_1–τ_4 reactions. The kinetic difference spectra indicate that the changes occurring in the τ_1 and τ_2 reactions are fundamentally different from those occurring in the τ_3 and τ_4 reactions (Figure 4). The minimum in the τ_2 phase at 220 nm suggests that the fraction of β structure increases while the fractions of turn and coil decrease. Although the amplitude of the τ_1 phase is too small to assess reliably, the similarity to the kinetic difference spectrum for the τ_2 phase suggests a similar explanation. For the τ_3 and τ_4 phases, the minimum around 230 nm suggests a decrease in the β content and a concomitant increase in other structural components (Table III). These results demonstrate that there must be significant rearrangement between different structural components, particularly the central β sheet, during the late stages of folding.

In support of the above analysis, it is very unlikely that the τ_1–τ_4 difference spectra are caused by side-chain contributions rather than changes in the peptide structure. The theoretical curves generated by the parameters in Table III represent well the observed difference spectra. Although each of the methods is based upon different reference spectra, they give essentially the same results with respect to the increase or decrease in β structure. Also, the aromatic residues in globular proteins are, in most cases, expected to lead to positive contributions to the rotational strength in the 225–250-nm region (Sears & Beychok, 1973; Woody, 1978). In contrast to this expectation, the difference spectra for the τ_1–τ_4 phases all show a negative ellipticity in this region.

Implications for the Mechanism of Folding of DHFR. A kinetic mechanism for the reversible unfolding and refolding reactions of DHFR has been proposed to involve four parallel channels and a series of native, intermediate, and unfolded conformations (Touchette et al., 1986). More recently, Frieden (1990) has proposed that a single channel with a series of intermediates some of which can bind substrate and/or cofactor and are enzymatically active can account for his fluorescence quenching data. Particularly striking about Frieden's simulations is the conclusion that intermediates can bind the substrate without changing the folding relaxation times; only the amplitudes are significantly altered.

With this assumption, the original observation that methotrexate binds in four exponentials whose relaxation times are nearly identical with those of the protein folding reactions could be explained by two folding channels. The pairwise similarities of the kinetic difference spectra for the τ_1 and τ_2 phases, and the τ_3 and τ_4 phases also support a two-channel model:



(7)

U_1 and U_2 are two unfolded conformers which are proposed to interconvert slowly on the folding time scale; proline isomerization may be responsible (Brandts et al., 1975). I_3 and I_4 are intermediates which appear in the dead time of mixing and have the properties of the burst-phase species described above. I_1 and I_2 are intermediates which bury Trp 74 in a hydrophobic cluster. The reactions which lead to these species are sufficiently similar in their relaxation times that they appear as a single reaction. Although the I_1 and I_2 species have tertiary structure, they cannot bind methotrexate and therefore must not have a proper active site pocket. The collection of native conformers, N_1–N_4, bind methotrexate and appear in the final, rate-limiting stages of folding.

All four species, N_1–N_4, are designated as native conformers because they are measurably populated at equilibrium. Preliminary stopped-flow CD (Kuwajima, unpublished results) and fluorescence (Touchette et al., 1986; Finn and Matthews, unpublished results) data on both unfolding and refolding show that all four phases, τ_1–τ_4, are detected in both reactions. The simplest interpretation is that four species are populated at equilibrium as shown above. In order to detect the τ_1 and τ_2 phases in refolding in this model, there must be absorbance, fluorescence, and CD differences between N_1 and N_3 and between N_2 and N_4. Thus, although the N_3 and N_4 species are native-like in their ability to bind methotrexate and in being populated in the absence of denaturant, there must be significant structural differences compared to N_1 and N_2. By any of the optical techniques, N_2 is the predominant species in solution in the absence of denaturant. A more detailed description of these and other data supporting the two channel folding model will be presented elsewhere (Jennings, Finn, and Matthews, manuscript in preparation).

Conclusions. Stopped-flow studies on the folding of DHFR show that secondary structure forms as soon as the denaturant is removed (within milliseconds) and continues to change until the final, rate-limiting steps in folding, orders of magnitude later in time. The initial event appears to involve the formation of helices and probably also β structure while later events involve continued development of helices and the rearrangement of the central β sheet. Tertiary structure arises after the initial event in the secondary structure, consistent with the framework model for folding.

ACKNOWLEDGMENTS

We thank Masahiro Iwakura for helpful discussion and for the gift of plasmid pTP64-1. We thank Gwen Corman for assistance in the preparation of the manuscript.

Registry No. Dihydrofolate reductase, 9002-03-3.

REFERENCES

Adler, A. J., Greenfield, N. J., & Fasman, G. D. (1973) *Methods Enzymol. 27*, 675–735.

Baccanari, D., Stone, D., & Kuyper, L. (1981) *J. Biol. Chem. 254*, 1799–1805.

Bayley, P. M. (1981) *Prog. Biophys. Mol. Biol 37*, 149–180.

Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C., & Kraut, J. (1982) *J. Biol. Chem. 257*, 13650–13662.

Brahms, S., & Brahms, J. (1980) *J. Mol. Biol. 138*, 149–178.

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

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry, Part II*, Freeman, San Francisco.

Cantor, C. R., & Timasheff, S. N. (1982) in *The Proteins* (Neurath, H., & Hill, R. L., Eds.) Vol. V, pp 145–306, Academic Press, New York.

Chang, C. T., Wu, C.-S. C., & Yang, J. T. (1978) *Anal. Biochem. 91*, 13–31.

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

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

Frieden, C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4413–4416.

Garvey, E. P., & Matthews, C. R. (1989) *Biochemistry 28*, 2083–2093.

Garvey, E. P., Swank, J., & Matthews, C. R. (1989) *Proteins 6*, 259–266.

Goldberg, M. E., Semisotnov, G. V., Friguet, B., Kuwajima, K., Ptitsyn, O. B., & Sugai, S. (1990) *FEBS Lett. 263*, 51–56.

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

Hennessey, J. P., Jr., & Johnson, W. C., Jr. (1981) *Biochemistry 20*, 1085–1094.

Hillcoat, B., Nixon, P., & Blakely, R. L. (1967) *Anal. Biochem. 21*, 178–189.

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

Johnson, W. C., Jr. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 145–166.

Johnson, W. C., Jr. (1990) *Proteins 7*, 205–214.

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

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.

Kuwajima, K. (1989) *Proteins 6*, 87–103.

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.

Kuwajima, K., Sakuraoka, A., Fueki, S., Yoneyama, M., & Sugai, S. (1988) *Biochemistry 27*, 7419–7428.

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

Manning, M. C., & Woody, R. W. (1989) *Biochemistry 28*, 8609–8613.

Mead, D. A., Szczesna-Skorupa, E., & Kemper, B. (1986) *Prot. Eng. 1*, 67–74.

Oas, T. G., & Kim, P. S. (1988) *Nature 336*, 42–48.

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

Perlmann, G. E., & Grizzuti, K. (1971) *Biochemistry 10*, 258–264.

Perry, K. M., Onuffer, J. J., Touchette, N. A., Herndon, C. S., Gittelman, M. S., Matthews, C. R., Chen, J.-T., Mayer, R. J., Taira, K., Benkovic, S. J., Howell, E. E., & Kraut, J. (1987) *Biochemistry 26*, 2674–2682.

Perry, K. M., Onuffer, J. J., Gittelman, M. S., Barmat, L., & Matthews, C. R. (1989) *Biochemistry 28*, 7961–7968.

Pflumm, M., Luchins, J., & Beychok, S. (1986) *Methods Enzymol. 130*, 519-534.
Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273-293.
Rao, S. T., & Rossmann, M. G. (1973) *J. Mol. Biol. 76*, 241-256.
Richardson, J. S. (1977) *Nature 268*, 495-500.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700-704.
Rossmann, M. G., & Argos, P. (1981) *Annu. Rev. Biochem. 50*, 497-532.
Sears, D. W., & Beychok, S. (1973) in *Physical Principles and Techniques of Protein Chemistry, Part C* (Leach, S. J., Ed.) pp 445-593, Academic Press, New York.
Sugawara, T., Kuwajima, K., & Sugai, S. (1991) *Biochemistry 30*, 2698-2706.
Touchette, N. A., Perry, K. M., & Matthews, C. R. (1986) *Biochemistry 25*, 5445-5452.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694-699.
Woody, R. W. (1977) *J. Polymer Sci.: Macromol. Rev. 12*, 181-321.
Woody, R. W. (1978) *Biopolymers 17*, 1451-1467.
Yang, J. T., Wu, C.-S., & Martinez, H. M. (1986) *Methods Enzymol. 130*, 208-269.

Self-Association of Human and Porcine Relaxin As Assessed by Analytical Ultracentrifugation and Circular Dichroism

Steven J. Shire,*,‡ Leslie A. Holladay,§ and Ernst Rinderknecht∥

Department of Pharmaceutical Research and Development and Department of Recovery Process Research and Development, Genentech, Inc., South San Francisco, California 94080, and Department of Research, Beckman Instruments Inc., Palo Alto, California 94303-0803

Received February 8, 1991; Revised Manuscript Received May 17, 1991

ABSTRACT: The self-association properties of recombinant DNA derived human relaxin, and porcine relaxin isolated from porcine ovaries, have been studied by sedimentation equilibrium analytical ultracentrifugation and circular dichroism (CD). The human relaxin ultracentrifuge data were adequately defined by a monomer–dimer self-association model with an association constant of $\sim 6 \times 10^5\ M^{-1}$, whereas porcine relaxin was essentially monomeric in solution. An approximate 5-fold increase in weight fraction of human relaxin monomer elicited by dilution of the protein resulted in no change in the far-UV CD spectrum at 220 nm. In contrast, after the same increase in weight fraction of monomer, the near-UV circular dichroism spectra for human relaxin exhibited a significant decrease in the amplitude for the CD bands near 277 and 284 nm. These CD bands, which may be assigned to the lone tyrosine in human relaxin, are superimposed on a broad envelope that is probably due to the three disulfide chromophores. Although both the human and porcine proteins contain two tryptophan residues, the near-UV CD spectra exhibit only a broad shoulder near 295 nm rather than the strong CD bands often found for tryptophan. Moreover, there is little change in this broad band after dilution of human relaxin to concentrations that resulted in a 4-fold increase in monomer weight fraction. These data suggest that dissociation of the human relaxin dimer to monomer is not accompanied by large overall changes in secondary structure or alteration in the average tryptophan environment, whereas there is a significant change in the tyrosine environment. This tyrosine is replaced by an arginine in porcine relaxin, and this substitution may affect the self-association properties of this protein.

Relaxin is a protein hormone that plays a major role in the reproductive biology of various species (Sherwood, 1988; Bryant-Greenwood, 1982). This protein appears to modulate the restructuring of connective tissues in target organs in order to generate the required changes in organ structure during pregnancy and parturition. In particular, relaxin regulates a number of biological responses of reproductive tissues in pregnant animals including softening of cervical tissue and inhibition of uterine myometrial contractions (Steinetz et al., 1959). Some of the important roles for relaxin as a pregnancy hormone include inhibition of premature labor and induction of cervical ripening prior to parturition.

The latter activity has been investigated in a clinical setting with use of relaxin isolated from porcine corpus lutea and ovaries (Evans et al., 1983; MacLennan et al., 1981, 1986). The preparation of porcine relaxin has been well documented (Sherwood & O'Byrne, 1974; Bullesbach & Schwabe, 1985, 1986), and since the sow ovary serves as a relatively plentiful source of relaxin, most studies on this hormone have been conducted with porcine relaxin. Identification of the human gene responsible for expression of human relaxin (Hudson et al., 1984) has resulted in the cloning, expression, and purification of recombinant DNA derived human relaxin (rhRlx) that is active in a cyclic AMP response assay using human endometrial cells (Fei et al., 1990).

Relaxin consists of two polypeptide chains that are linked by inter- and intrachain disulfide bonds. Although the amino acid sequence homology between relaxin and insulin is low (Figure 1), there are several important homologies between these protein hormones. The positions of the disulfide links are similar in both proteins, and there is a conservation of the glycine residues immediately adjacent to the cysteine residues in the B chain (Sherwood, 1988). These residues are crucial to maintaining the proper structural folding for insulin and

‡ Department of Pharmaceutical Research and Development, Genentech, Inc.
§ Department of Recovery Process Research and Development, Genentech, Inc.
∥ Department of Research, Beckman Instruments Inc.

0006-2960/91/0430-7703$02.50/0