Kinetic Analysis of R67 Dihydrofolate Reductase Folding: From the Unfolded Monomer to the Native Tetramer†

Christophe Bodenreider,‡ Nicolas Kellershohn,§ Michel E. Goldberg,‡ and Annick Méjean*,‡

Unité de Repliement et Modélisation des Protéines, CNRS URA 2185, Institut Pasteur, 28 rue du Dr. Roux 75724 Paris Cédex 15, France, and Imagerie et Dynamique Cellulaires, Université de Paris Sud, 91405 Orsay Cédex, France

Received July 3, 2002; Revised Manuscript Received October 2, 2002

ABSTRACT: R67 dihydrofolate reductase (DHFR) is a homotetrameric enzyme. Its subunit has a core structure consisting of five antiparallel β-strands that form a compact β-barrel. Our interest was to describe the molecular mechanism of the complete folding pathway of this β-sheet protein, focusing on how the oligomerization steps are coordinated with the formation of secondary and tertiary structures all along the folding process. The folding kinetics of R67 dihydrofolate reductase into dimers at pH 5.0 were first examined by intrinsic tryptophan fluorescence, fluorescence energy transfer, and circular dichroism spectroscopy. The process was shown to consist of at least four steps, including a burst, a rapid, a medium, and a slow phase. Measurements of the ellipticity at 222 nm indicated that about 50% of the total change associated with refolding occurred during the 4 ms dead time of the stopped-flow instrument, indicating a substantial burst of secondary structure. The bimolecular association step was detected using fluorescence energy transfer and corresponded to the rapid phase. The slow phase was attributed to a rate-limiting isomerization of peptidyl-prolyl bonds involving 15% of the unfolded population. A complete folding pathway from the unfolded monomer to the native tetramer was proposed and an original model based upon the existence of early partially folded monomeric intermediates, rapidly stabilized in a dimeric form able to self-associate into the native homotetramer was formulated. The rate constants of these various steps were determined by fitting the kinetic traces to this model and supported our mechanistic assumptions.

Efforts to examine the mechanisms of protein folding have largely increased in the past few years, and the folding of a wide variety of protein structures has been studied (*1*−*4*). However, multimeric proteins have remained particularly under-represented in folding studies despite the fact that such proteins are relatively common in biological systems and their oligomeric nature plays an essential role in determining their biological function in vivo. Moreover, oligomeric proteins often do not obey all the folding principles established for smaller proteins (*3*, *5*, *6*), and the role of intramolecular forces in folding and stabilizing secondary and tertiary structure of simple monomeric proteins provides little information on how intermolecular associations stabilize oligomeric systems.

The folding pathways of a few oligomeric proteins have been reported and revealed a variety of mechanisms (*5*, *7*−*11*). Some proteins display folding monomeric or dimeric intermediates as reported for the *Escherichia coli* Trp repressor (*12*), the ATPase SecA (*13*), or the human glutathione transferase A1−1 (*14*), whereas others fold in two-state reactions in which folding and dimerization occur simultaneously, i.e., stable folding intermediates are not detected, as observed for the dimeric P22 Arc repressor (*15*, *16*). However, the proteins for which transient intermediates were detected span a range of complexities in their folding behavior from the simple association-competent monomeric intermediate (*14*) to multiple transient intermediates such as those detected in the folding of the multidomain β2 Trp synthase dimer (*17*−*19*). Obviously, such folding processes can be very complex, and elementary steps have been characterized in only few cases [review by Jaenicke (*5*)]. Thus, the way in which the amino acid sequence of a polypeptide controls both the folding of individual subunits and their assembly to higher-order quaternary structure is far from being well understood.

Another weakness in our understanding of protein folding consists of the relatively few detailed investigations of predominately β-sheet proteins. Such proteins are not only important as models for understanding the formation of β-sheets, but also their folding properties may differ significantly from those of helical proteins since the interactions that stabilize their secondary structure are predominantly nonlocal in nature. The disorders in the formation of β-sheet structures observed in Alzheimer's, Huntington's, and prion diseases (*20*) strengthen the need for a better understanding of the folding mechanism of β-sheet proteins.

In a previous report, we initiated kinetic studies on the model protein R67 DHFR[1] with the aim of providing at the molecular level a dynamic vision of the mechanism of protein−protein interactions (*21*). This oligomeric protein encoded by the plasmid R67 of *E. coli* (*22*, *23*) is interesting

† This work was supported by funds from the Institut Pasteur and the Centre National de la Recherche Scientifique (URA 2185). C.B. was supported by a fellowship from the Ministère de l'Education Nationale de la Recherche et de la Technologie.

* Corresponding author. Tel: 33 1 40 61 32 81. Fax 33 1 40 61 30 43. E-mail: amejean@pasteur.fr.

‡ Institut Pasteur.

§ Université de Paris Sud

10.1021/bi020453b CCC: $22.00
Published on Web 11/16/2002

for several reasons. The R67 DHFR monomer is one of the smallest polypeptide chains (78 residues) (*24*) known to assemble into an enzymatically active quaternary structure. Though it catalyzes the same reaction as chromosomal DHFRs, it has, interestingly, no sequence or structural homology with the chromosomal enzyme (*22*). Its activity is responsible for the resistance to the antibiotic trimethoprim (*22*, *24*, *25*). The R67 DHFR is active only in the tetrameric state (*26*) and undergoes a pH-dependent equilibrium, with a stable dimeric form below pH about 5.0. The crystal structures of the dimer and tetramer have been solved (*27*, *28*). The homotetramer catalyzes the NADPH-dependent reduction of 7,8-dihydrofolate (DHF) into 5,6,7,8-tetrahydofolate (THF) in a single active-site pore, passing through the middle of the doughnut shaped macromolecule (*29*, *30*). The crystal structure shows that each dimer−dimer interface is stabilized by a hydrogen bond formed between the imidazole ring of the histidine 62 from one protomer and the hydroxyl group of the nearby serine 59 from the symmetry related protomer (*28*). Our previous studies (*21*) demonstrated that protonation of this histidine 62 within the tetramer triggers the dissociation into dimers and that the driving force of dimer−dimer association resides in the establishment of four hydrogen bonds per tetramer.

R67 DHFR is quite interesting not only because of its pH-dependent oligomerization equilibrium but also because it is a pure β-sheet protein. Each subunit has a core structure consisting of five antiparallel β-strands (*28*). Two monomers associate to form a dimer with three strands from each monomer meshing to form a compact six-stranded β-barrel at the subunit interface. R67 DHFR thus constitutes an ideal model to address the relative contributions of long range and local interactions in the initiation of β-sheet protein folding and to understand how the coordination of the various levels of structural organization along the folding process of an oligomeric protein proceeds.

Previously reported folding studies (*31*) led to the conclusion that R67 DHFR undergoes at pH 5.0, a simple two state unfolding equilibrium described by the following reaction scheme:

$$D \rightleftharpoons 2U$$

where U and D represent the unfolded monomer and the native dimer, respectively. No evidence was found for folded monomers as intermediate species in the unfolding of dimeric R67 DHFR since the equilibrium unfolding transition curves monitored by fluorescence and CD were superimposable and monophasic. These results suggested that the folded monomer is not stable as an isolated species and that folding and dimerization are tightly coupled events. Moreover, with increasing pH, dimeric R67 DHFR associates to form a tetramer which constitutes the true native state, and equilibrium folding studies on tetrameric R67 DHFR at pH 8.0 (*32*) supported the simplest folding mechanism with only one dimeric intermediate according to the following scheme:

$$4U \rightleftharpoons 2D \rightleftharpoons T$$

where T represents the native tetramer. The free energy value of the dimer−dimer association (2D → T) was deduced from the global analysis of the equilibrium unfolding transitions at pH 8.0 according to the scheme described above and was consistent with the previous $\Delta G°$ values determined from the R67 DHFR dimer−tetramer equilibrium experiments.

The aim of the present study was to gain a better insight into the molecular steps involved in R67 DHFR refolding and to search for intermediates that could provide direct information on the folding mechanisms. For that purpose, fluorescence and CD stopped-flow methods were applied to monitor the refolding kinetics. To simplify the kinetic analysis, we chose to investigate first the refolding into dimers at pH 5.0, where tetramerization could not occur. The refolding kinetics will be reported and analyzed in terms of a four steps model which takes into account the heterogeneity of the unfolded state, the dimerization step, and the first very rapid phase. The rate constants obtained by fitting the experimental data to this model will be reported. The existence of partially folded monomeric intermediates will be discussed. The complete refolding of R67 DHFR into tetramers at pH 8.0 was then studied and will be described in terms of a general refolding pathway leading from the unfolded monomer to the native tetramer.

MATERIALS AND METHODS

Materials and Protein Purification. Ultrapure guanidine hydrochloride (GuHCl) was purchased from ICN Biomedicals Inc. and used without further purification. Fluorescent labeling reagents were purchased from Molecular Probes (Leiden, The Netherlands). The cyclophilin (eCyp) was kindly provided by Dr. J. M. Betton, Institut Pasteur, Paris, France. Wild-type R67 DHFR was overexpressed in BLi5 bacteria (corresponding to BL21 lacIQpACYC Cm^R derived from Novagen) transformed with the *wt* R67 *dhfr* gene subcloned in the expression vector pET19b (Novagen) and purified using the same protocol as previously described (*21*). Protein concentrations were routinely measured spectrophotometrically using an extinction coefficient of 15 400 M^{-1} cm^{-1} at 280 nm (*31*) and were expressed in monomer molarity. Dihydrofolate reductase activity was assayed as previously described (*33*).

Stopped-Flow Fluorescence Measurements. All the kinetic experiments were carried out with an SFM3 mixing device from Bio-Logic (Pont de Claix, France) equipped with two large (18 mL) syringes injecting through the first mixer and a small syringe (5 mL) injecting into the second mixer. The mixing device, equipped with an F15 (1.5 mm × 1.5 mm cross section) fluorescence cell, was combined with the optical bench and detection module of Bio-Logic. Two photomultipliers were used, one for the detection of the emitted light through a 350 nm high pass filter for intrinsic fluorescence measurements and the second for the detection of the excitation light (295 nm). The ratio of the two signals was electronically determined in the Bio-Logic dual amplifier and recorded by means of the Bio-Kine software package of Bio-Logic. Kinetic measurements of R67 DHFR refolding

[1] Abbreviations: DHFR, dihydrofolate reductase; MTA buffer, 50 mM MES, 100 mM Tris, 50 mM acetic acid polybuffer; GuHCL, guanidinium hydrochloride; IANBD, *N*,*N*′-dimethyl-*N*-(iodoacetyl)-*N*′-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)ethylenediamine; DCIA, 7-diethylamino-3-((4′-(iodoacetyl) amino) phenyl)-4-methylcoumarin; IANBD−DHFR, R67 DHFR monomer covalently linked to IANBD; DCIA−DHFR, R67 DHFR covalently linked to DCIA; FRET, fluorescence resonance energy transfer; eCyp, *Escherichia coli* cyclophilin; CD, circular dichroism.

were obtained by 40-fold dilution of 20 μL (from the small syringe) of denatured protein in MTA buffer, pH 5.0, 4 M GuHCl with 780 μL (390 μL from each large syringe) MTA buffer, pH 5.0, containing different amounts of GuHCl for achieving final concentrations ranging between 0.1 and 0.8 M GuHCl. The final concentration of DHFR R67 in the observation cell ranged from 0.5 to 30 μM. The duration of the injection was 75 ms resulting in a 5 ms dead time. Recording of the fluorescence intensity was triggered at the end of the injection. All the solutions were filtered and extensively degassed immediately before they were used. The stopped-flow temperature was controlled by means of an external thermostated water bath and a high flux pump to circulate the water between the bath and the stopped-flow apparatus. Standard temperature was 25 °C.

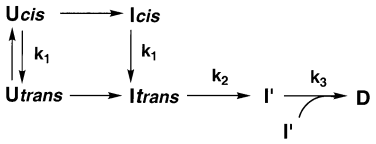
Fluorescence resonance energy transfer (FRET) experiments were conducted on proteins labeled with DCIA as the donor or IANBD as the acceptor. A premix of equimolar quantities of the two denatured labeled proteins, DCIA−DHFR and IANBD−DHFR, was introduced into the small syringe and the refolding kinetics were initiated in the stopped-flow as described above for the intrinsic fluorescence. The quenching of fluorescence of the donor was monitored by exciting at 384 nm and measuring the light emitted above 450 nm using a high pass filter. The sampling interval and filtering time constants will be specified in the legend of figures.

Time-Resolved CD Experiments. CD experiments were performed on a CD6-circular dichroism spectropolarimeter (Jobin-Yvon ISA, Longjumeau, France) equipped with a SFM-3 stopped-flow module (Bio/Logic, Claix, France). The stopped-flow unit and the observation cell (path length of 1.5 mm) were thermostated at 25 °C. Kinetic measurements of R67 DHFR refolding were obtained by rapid mixing (45 ms injection time) of 15 μL of denatured protein in MTA buffer, pH 5.0, 4 M GuHCl with 585 μL of MTA buffer, pH 5.0. Two final protein concentrations were used: 6 and 30 μM. The dead time of the instrument was 3 ms. The ellipticity at 222 nm was recorded, using a filtering time constant equal to the sampling interval. For each sample, several traces were averaged (from 20 to 400, depending on the signal). Controls with refolding buffer and GuHCl solutions alone were recorded for correcting the kinetic traces of any instrument offsets.

Double-Jump Experiments. Classical double jump experiments (native−unfolded−native) were performed by fluorescence stopped-flow measurements. In the first mixing step, denaturation was initiated by the dilution of native dimeric R67 DHFR in MTA buffer pH 5.0 to final denaturing conditions of 10 mM acetate, pH 1.5, 4 M GuHCl, 5 mM DTT. The second mixing step was applied after varying denaturation times (1 s to 45 min) and transferred the protein into renaturing conditions (MTA buffer, pH 5.0) to monitor refolding. The final concentration of protein was 0.6 μM. The excitation wavelength was 295 nm, and the emission was measured using a high-pass filter with a 350 nm cutoff.

Fluorescence Labeling and Fluorescence Spectra. The R67 DHFR was reduced for 1 h (50 mM sodium phosphate buffer, pH 7.0, 5 mM DTT) and denatured by dialysis against 10 mM acetate, pH 5.0, 4 M GuHCl. The single reactive SH residue was labeled by incubating reduced and denatured R67 DHFR (0.1 mM) with 1 mM IANBD or DCIA at pH 6.5 for 5 h at 20 °C. Excess reagent was removed by filtration on a PD10 column (Pharmacia) equilibrated with 50 mM sodium phosphate buffer, pH 6.5, 4M GuHCl. The amount of covalently bound probe was found to be respectively 1 mole of IANBD group per mole of DHFR R67 monomer and 0.75 mol of DCIA per mol of DHFR R67 monomer. The molar extinction coefficients used to determine the labeling yield were 25 000 M^{-1} cm^{-1} at 478 nm for IANBD and 31 000 M^{-1} cm^{-1} at 384 nm for DCIA. The fluorescence spectra of the labeled proteins were monitored at 20 °C in an LS5B (Perkin-Elmer) spectrofluorometer in 50 mM MES, 100 mM Tris, 50 mM acetic acid buffer (MTA) at pH 5.0. The protein concentration was 2 μg/mL. The excitation and emission bandwiths were 5 nm, and the scan speed was 30 nm/min.

Scheme 1



Data Analysis. For each set of experimental conditions, 10−30 kinetic traces were obtained, and the accumulated kinetic data were averaged and converted into an ASCII file. In the first approach, the resulting file was analyzed with the Figure P software (version 2.7 for Windows, Biosoft, Cambridge, U.K.) on a model consisting of a sum of exponentials.

Several refolding kinetic schemes were then tested for compatibility with the experimental data mechanism. Various programs were routinely used: the fitting routine PowerBASIC v. 2.1, an in-house software program using nonlinear least-squares method and the Simplex and Newton-Gauss minimization algorithms, the program Scientist (Micromath Inc., USA), and the program DynaFit by Petr Kuzmic (*34*). The final kinetic model was refined until no significant improvement of the fit was observed as judged from the residual errors. All parameters (rate constants k_1, k_2, k_3, species concentrations) were then determined according to the reaction scheme proposed in this paper (Scheme 1).

In the case of FRET experiments, fitting of DCIA−DHFR fluorescence extinction data was performed with the PowerBASIC v. 2.1 program assuming various dimerization models.

The model of an irreversible association from two distinct molecular species X_1 and X_2 stemming from an isomerization, with the mixed formation of the heterodimer X_1X_2 and the two homodimers X_1X_1 and X_2X_2, was selected. The reaction scheme associated to this system is represented as follows:

$$\begin{aligned} X_1 + X_2 &\xrightarrow{k_a} D \\ X_1 + X_2 &\xrightarrow{k_a} X_1X_1 \\ X_2 + X_2 &\xrightarrow{k_a} X_2X_2 \end{aligned} \qquad \text{with} \quad \begin{cases} X_1^* \xrightarrow{k_i} X_1 \\ X_2^* \xrightarrow{k_i} X_2 \end{cases}$$

where X_1 and X_2 are respectively the trans prolyl isomers of the DCIA and IANBD labeled monomers, D is the heterodimer, X_1X_2, X_1X_1, and X_2X_2 are the homodimers, and k_a is the association rate constant. X_1* and X_2*, the cis prolyl

isomers of X_1 and X_2, are isomerized into the trans conformer with a rate constant k_i which will be fit and the cis/trans proportion is allowed to vary between 10% and 15% (see Results).

If $[X_{1T}]$ and $[X_{2T}]$ represented, respectively, the total molarity of the DCIA−DHFR monomer and the total molarity of the IANBD−DHFR monomer, $[X_{1D}]$ the concentration of X_1 in the homodimer X_1X_1, and $[X_{2D}]$ the concentration of X_2 in the homodimer X_2X_2, the differential equations system was

$$\frac{d[D]}{dt} = k_a[X_1][X_2] = k_a([X_{1T}] - (2[X_{1D}] + [D])) \times ([X_{2T}] - (2[X_{2D}] + [D])) \quad (1)$$

$$\frac{d[X_{1D}]}{dt} = k_a[X_1]^2 = k_a([X_{1T}] - (2[X_{1D}] + [D]))^2 \quad (2)$$

$$\frac{d[X_{2D}]}{dt} = k_a[X_1]^2 = k_a([X_{2T}] - (2[X_{2D}] + [D]))^2 \quad (3)$$

where $[X_{1T}] = [X_1] + 2[X_{D1}] + [D]$ and $[X_{2T}] = [X_2] + 2[X_{D2}] + [D]$.

The time course of fluorescence change was described by the equation

$$F(t) = \epsilon_1[X_1] + \epsilon_2[X_2] + \epsilon_D[D] + \epsilon_{1D}[X_{1D}] + \epsilon_{2D}[X_{2D}] \quad (4)$$

where $F(t)$ is the total amplitude at time t; ϵ_1,ϵ_2 , ϵ_{1D}, ϵ_{2D}, and ϵ_D are, respectively, the specific fluorescences of X_1, X_2, X_{1D}, X_{2D}, and D, with $\epsilon_1 \neq \epsilon_{1D} > 0$ and $\epsilon_2 = \epsilon_{2D} = 0$.

RESULTS

Equilibrium Unfolding Studies. Equilibrium unfolding/folding of R67 DHFR was investigated at pH 5.0 by monitoring the tryptophan fluorescence intensity at 340 nm as function of guanidine hydrochloride (data not shown). As previously described (*32*), the transition of R67 DHFR from native to denatured conditions resulted in a red shift of the emission spectrum with an overall decrease in fluorescence intensity. The denaturation and renaturation reactions were shown to be reversible at 3.5 μM protein concentration, 25 °C, and pH 5.0, and transition curves were consistent with a two-state equilibrium between native dimers and unfolded monomers without any detectable thermodynamically stable intermediates. The free energy of unfolding in the absence of GuHCl were calculated to be 12.3 ± 0.4 kcal/mol in agreement with previous $\Delta G°$ values reported in the literature ($\Delta G°$ = 12.3 kcal/mol) (*31*, *32*).

Refolding Kinetics Monitored by Tryptophan Fluorescence and Circular Dichroism. Refolding kinetics of R67 DHFR were examined in stopped-flow experiments by analyzing tryptophan fluorescence and ellipticity changes during the refolding process. Unfolded proteins at 4 M GuHCl and pH 5.0 were diluted 40-fold to induce refolding. Figure 1 shows typical kinetic traces monitored by intrinsic fluorescence (λ_{exc} = 295 nm; emission above 350 nm) (Figure 1A) and ellipticity at 222 nm (Figure 1B). The kinetics in fluorescence were complex. The fluorescence first very rapidly decreased (see inset, Figure 1A) and then increased to finally reach the fluorescence of native dimers. To detect a potential burst phase during the dead time of the stopped-flow instrument, the initial fluorescence intensity of the unfolded protein was determined at 0.1 M GuHCl by extrapolation from the

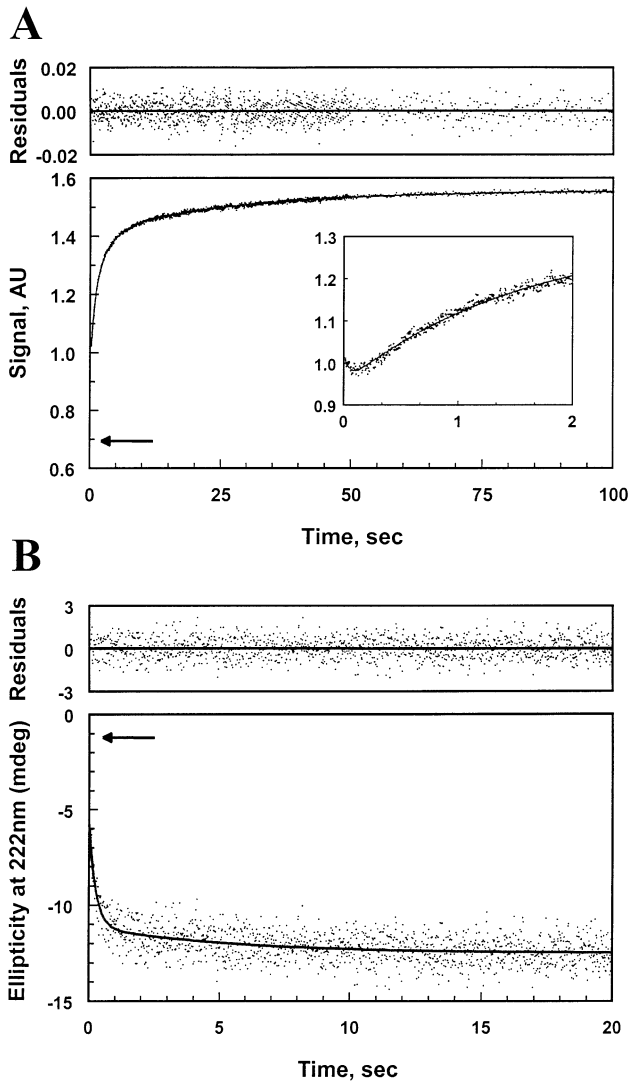


FIGURE 1: Refolding kinetics of R67 DHFR at pH 5.0 and 25 °C monitored by fluorescence (panel A) and circular dichroism (panel B). The kinetic traces shown result from the average of 10 recordings in fluorescence and 200 in circular dichroism. The solid lines represent the best fits of bi- or triphasic sequential kinetics to the experimental points, and the residuals of the fits are shown above panel A and B, respectively. Panel A shows a typical trace observed by stopped-flow fluorescence and, in the inset, an enlarged view of the firsts two seconds. An excitation wavelength of 295 nm was used, and fluorescence was collected above 350 nm using a high pass cutoff filter. The concentration of R67 DHFR in the observation cell was 0.6 μM. The horizontal arrow indicates the fluorescence intensity value for the denaturated protein at 0.1 M GuHCl, pH 5.0, extrapolated from the variation with the final GuHCl concentration, of the fluorescence intensity measured at the end of the dead time of the stopped-flow apparatus. Recording of the fluorescence was triggered at the end of the injection for a duration of 100 s with a sampling interval of 100 ms on one channel and for 10 s with a sampling interval of 10 ms on a second channel (see inset). For both recordings, the filtering constant was 10 ms. The fluorescence intensity is in arbitrary units. Panel B shows a kinetic trace of ellipticity changes at 222 nm. The concentration of R67 DHFR in the observation cell was 30 μM. The horizontal arrow indicates the value of the ellipticity of unfolded R67 DHFR at pH 5.0 obtained by repeated injections of R67 DHFR in 4.0 M GuHCl, pH 5.0. The sampling interval and filtering constant were 5 ms.

variation of the first measurable fluorescence intensity of R67 DHFR at GuHCl concentrations ranging between 2 and 4 M. The extrapolation yielded a fluorescence intensity value at time zero corresponding to 35% of the total fluorescence of native R67 DHFR, indicating that the rapid phase of fluorescence decrease was in fact preceded by an even faster one that occurred during the 4 ms dead time of the machine and resulted in a significant fluorescence variation. The low amplitude of the first observable phase and the very limited number of experimental points available prevented a quantitative analysis of this phase.

Unlike the rapid fluorescence decrease, the fluorescence increase could be satisfactorily fitted to a sum of two exponential functions with a random distribution of the residuals (see Figure 1 above the kinetic trace). The rate constants of the two corresponding phases were found to be about 1.7 s^{-1} for the medium phase and 0.04 s^{-1} for the slow phase at 6 μM protein and 25 °C. The relative amplitudes of these medium and slow phases accounted for about, respectively, 65% and 35% of the visible part of the kinetics. Thus, an initial burst phase followed by three kinetic phases was identified that gave rise to fluorescence variations of different signs and amplitudes.

The kinetics of R67 DHFR refolding were also analyzed by monitoring the ellipticity at 222 nm (Figure 1B). As expected from the spectra of the unfolded monomer and native dimer, the refolding was reflected by a gain of ellipticity amplitude. 50% of the total ellipticity amplitude was reached at the end of the dead time, indicating that a very rapid phase was completed within the first 4 ms. Then the signal decreased more slowly to reach the ellipticity of the native dimeric protein. As observed for the fluorescence increase, the ellipticity change could be fitted to a sum of two exponential functions. The procedure provided good fits of the kinetic traces (low X^2 and random residuals) and allowed to estimate the rate constants to be about 2.6 s^{-1} and 0.05 s^{-1}, close to those determined for the slow and medium phases observed in fluorescence at the same protein concentration. Thus, the circular dichroism signal obeyed a complex kinetic model that could be decomposed in at least three distinct phases: a burst, a medium and a slow phase.

Concentration Dependence Experiments. Since native R67 DHFR at pH 5.0 is a dimer, its folding must include a bimolecular reaction and hence a step with a rate constant dependent on the protein concentration. To find out which of the observed phases might correspond to the dimerization step, refolding experiments were carried out at various protein concentrations. In fluorescence as in circular dichroism, only the medium phase was dependent on the protein concentration whereas the rate constant of the slow phase appeared to be concentration-independent. The relative amplitude of the burst phase observed in fluorescence and CD, and that of the rapid phase detected in fluorescence were unchanged with the protein concentration. Figure 2 shows the plot of the medium rate constant as a function of protein concentration. Fitting the portion of the plot below 4 μM of monomer to a straight line provided a slope of 0.56×10^6 M^{-1} s^{-1}, which can be considered as a rough estimate of the second-order association rate constant at 25 °C and pH 5.0. Therefore, though fitting the fluorescence kinetics to the sum of exponential terms is obviously inconsistent with the presence of a bimolecular reaction, this simple fitting procedure was used for a qualitative characterization of the different kinetic phases.

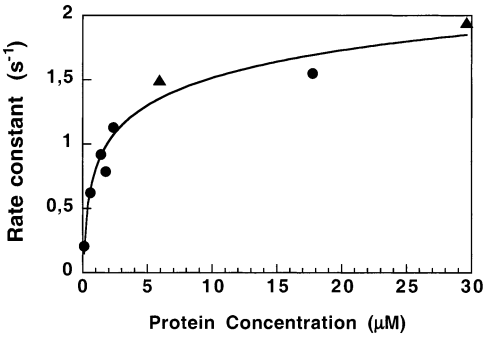


FIGURE 2: Dependence of the rate constant of the medium refolding phase on the protein concentration. Experiments similar to those depicted in Figure 1 monitored in fluorescence (●) or in circular dichroism (▲) were repeated at various protein concentrations. The kinetics were analyzed as indicated in Figure 1, and the resulting rate constants obtained for the medium phase are plotted as a function of protein concentration. An approximate second-order rate constant is determined in the low protein concentration range (below 4 μM) from the slope of the plot and yields a rate constant value of about 0.56×10^6 M^{-1} s^{-1}.

The Dimerization Step. To investigate specifically the dimerization step on the pathway of R67 DHFR folding, the physical association of the monomers had to be monitored directly. Energy transfer between fluorescent probes bound to the isolated monomers seemed a suitable signal for monitoring the association (*35*, *36*). To introduce a covalent fluorescent label on each monomer, we took advantage of the presence of a unique cysteinyl residue (C47) per monomer exposed to the solvent. The distance of 19 Å between the α-carbons of cystein 47 and cystein 147 (or between C247 and C347) in the R67 DHFR tetramer structure (*28*) is compatible with a potential energy transfer. Moreover, the C47 residue is the first amino acid residue belonging to the β-strand involved in the intermonomer β barrel and thus should be a direct reporter group for the dimerization.

Analogues of the coumarin and benzoxadiazol fluorescent probes were chosen as fluorescent donor and acceptor, respectively. The reduced R67 DHFR was specifically labeled with DCIA (iodoacetyl derivative of coumarin) or IANBD (iodoacetyl derivative of benzoxadiazol) in the unfolded state and renatured either separately or as a stoichiometric mixture. Figure 3 shows the fluorescence emission spectra (λ_{exc} = 384 nm, corresponding to the maximum excitation wavelength of the donor group) of the two individual labeled proteins and their mixture. It can be seen that DCIA−DHFR alone exhibited a strong emission band with a maximum at 470 nm whereas a weak fluorescence emission of IANBD−DHFR was centered around 535 nm and undetectable at 470 nm. When the two labeled proteins were mixed and renatured together, the fluorescence spectrum showed a clear decrease of the donor (DCIA−DHFR) emission, consistent with an energy transfer occurring between the donor and the acceptor.

The refolding kinetics of IANBD−DHFR alone were investigated by monitoring the intrinsic tryptophan fluorescence above 350 nm (λ_{exc} = 295 nm), and the rate constants of the two major refolding phases were determined. The rate constants were unchanged as compared to those obtained

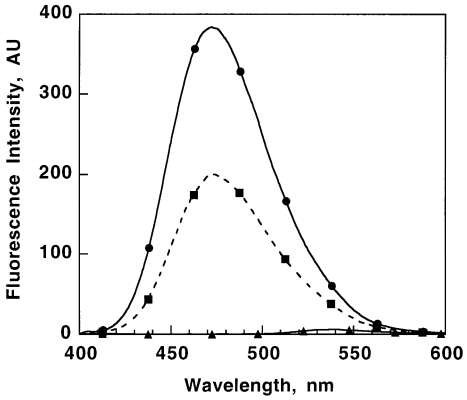


FIGURE 3: Comparison of the fluorescence emission spectra (excitation at 384 nm) of the labeled proteins DCIA−DHFR and IANBD−DHFR, either isolated or renatured as a mixture. All labeled proteins were in MTA buffer pH 5.0 at the protein concentration of 2 μg/mL. The DCIA−DHFR (●) and IANBD−DHFR (▲) spectra are described by solid line and the mix (1:1) (■) of DCIA−DHFR and IANBD−DHFR by dashed lines.

with the unlabeled protein, confirming that alkylation of the C47 residues did not affect the R67 DHFR refolding kinetics.

Thus, the extinction of the emission fluorescence of the donor (DCIA−DHFR) above 450 nm (λ exc = 384 nm) could be used to monitor the kinetics of the monomer−monomer association. Figure 4A shows a typical kinetic trace from a FRET experiment at 0.1 M GuHCl and 25 °C. The DCIA emission fluorescence first rapidly decreased and then slowly reached a stable value. Various attempts to fit the FRET kinetic data led us to retain an irreversible dimerization model from two molecular species X_1 and X_2, arising from an isomerization reaction, with the mixed formation of heterodimer (X_1X_2) and homodimer (X_1X_1) and (X_2X_2). In this model (described in Materials and Methods), we assumed that the first order isomerization rate constant was the same for both species, the reactions of heterodimerization and homodimerization occurred concurrently with the same association rate constant (k_a), and the fluorescence decrease was only due to the heterodimer appearance (X_1X_2). Using the non linear least squares curve fitter PowerBASIC v. 2.1, FRET experimental data (Figure 4A) were satisfactory fitted according to this reaction scheme as judged by the residual errors (see above Figure 4A). They provided a first-order isomerization rate constant $k_i = 0.037\ s^{-1}$ and a bimolecular rate constant of $k_a = 2.8 \times 10^6\ M^{-1}\ s^{-1}$, in reasonable agreement with the approximate value estimated above from the concentration dependence of the apparent first-order rate constant of the medium phase (Figure 2). The proportion of cis species was found to be 9%.

Denaturant Dependence of the Refolding Kinetics of R67 DHFR. Since the R67 DHFR monomer contains five proline residues in trans conformation in the native dimer, proline isomerization reactions could be expected to occur during the refolding process. To test the existence of proline isomerization, the kinetic behavior of the refolding phases was analyzed at different denaturant concentrations. The apparent amplitude of the rapid phase increased while increasing the final GuHCl concentration, whereas the amplitude of the two other phases seemed unaffected (data not shown). This made it possible to perform a satisfactory fit of the complete kinetic traces obtained in the presence of

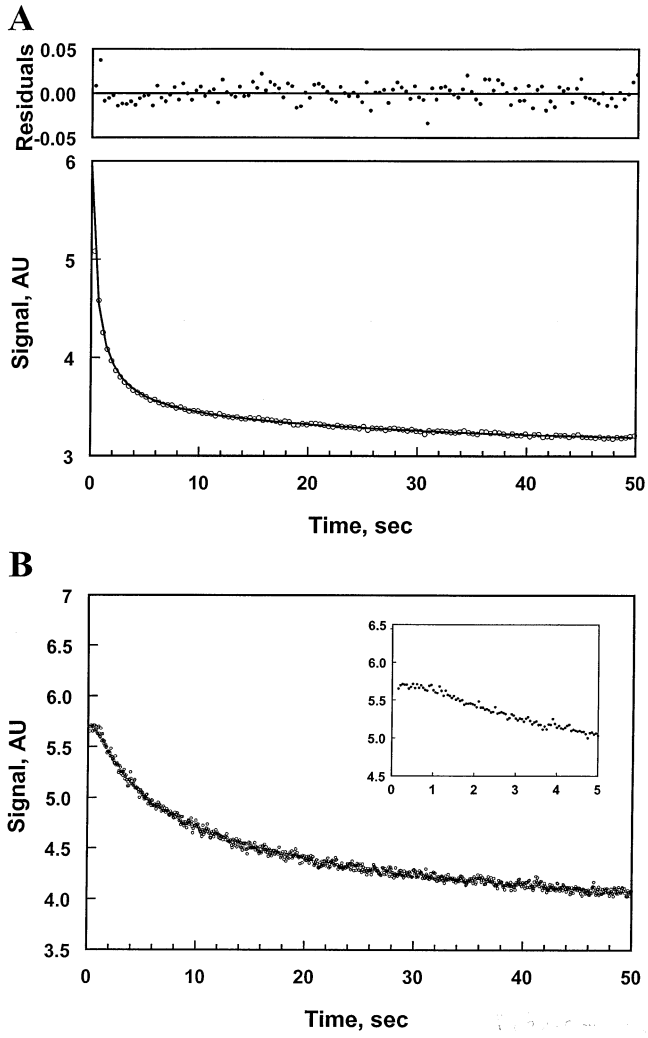


FIGURE 4: Kinetics of quenching of fluorescence during refolding upon association of DCIA−DHFR and IANBD−DHFR monomers in standard conditions (panel A) and at 12 °C and 0.7 M GuHCl final (panel B). An excitation wavelength of 384 nm was used, and the fluorescence was collected above 450 nm. Each trace results from the accumulation and averaging of 10 kinetic runs. Under standard conditions (panel A) a premix of the two denatured labeled DCIA−DHFR (20 μM) and IANBD−DHFR (27 μM) in 4M GuHCl, pH 5.0 (from the small syringe), was diluted at 25 °C 40-fold in MTA buffer pH 5.0 (from the large syringe). The fluorescence was recorded for 50 s with a sampling interval and a filtering constant of 50 ms. However, because of limitations of the fitting routine PowerBASIC used for FRET data analysis, the 1000 experimental data points were averaged two by two in 512 points represented by open circles. The solid line represents the best fit to the experimental points according to an irreversible dimerization from two species X_1 and X_2, stemming from an isomerization reaction, with the mixed formation of homodimers and heterodimers (see Material and Methods). The residuals of this fit are shown above the kinetic trace. In panel B, the premix of the two denatured labeled DHFR species was diluted at 12 °C 40-fold in MTA buffer 0.6 M GuHCl, pH 5.0, to reach a GuHCl final concentration of 0.7 M. The fluorescence was recorded for a duration of 50 s on one channel and for 5 s (see inset) on a second channel. For both recordings, the sampling interval and the filtering constant were 50 ms.

residual GuHCl using a three steps model and the sum of three exponential functions

Figure 5 shows a linear denaturant dependence at 25 °C of the logarithm of the refolding rate constants of the rapid, medium, and slow folding phases. The apparent rate constant

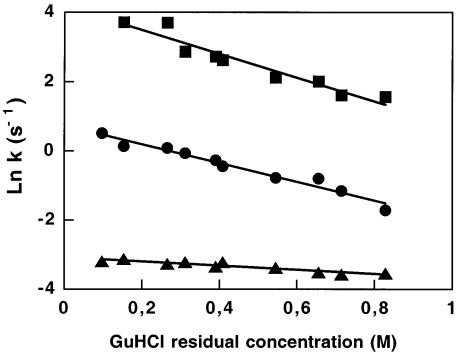


FIGURE 5: Denaturant dependence of rate constants measured by intrinsic fluorescence at pH 5.0 and 25 °C in stopped-flow fluorescence. Experiments similar to those depicted in Figure 1A were repeated in which the buffer contained in the two large syringes was supplemented with GuHCl. The kinetics were analyzed as indicated in Figure 1 and the resulting rate constants obtained at each GuHCl concentration. The logarithm of the rate constant corresponding to the rapid (■), medium (●), and slow (▲) phases are plotted as a function of the GuHCl concentration.

Table 1: Catalytic Effect of eCyp on the Rapid and Slow Refolding Rate Constants[a] during Refolding of DHFR R67 at 25 °C and pH 5.0

	eCyp concentration (nM)					
rate constant (s^{-1})	0	12	30	120	300	600
medium	0.65	0.69	0.93	0.79	0.86	0.69
slow	0.04	0.05	0.09	0.11	0.16	0.15

[a] Rate constants are determined by fitting to a two exponential process the refolding traces obtained in the presence of various concentrations of eCyp. The DHFR R67 concentration was 1 μM, and the final buffer conditions were 0.1 M GuHCl, MTA buffer, pH 5.0.

of the rapid phase was clearly slowed in the presence of residual GuHCl, and its value at 0.1 M residual GuHCl could be reliably determined by extrapolation from the denaturant dependence of the rate constant. A value of 46 s^{-1} was obtained. The apparent first-order rate constant of the medium phase decreased almost 10-fold as the concentration of denaturant increased from 0.1 to 0.8 M GuHCl, whereas the rate constant for the slow phase showed only a weak denaturant dependence. The weak dependence of the refolding rate of the slow phase on denaturant concentration suggested that this kinetic phase might correspond to a peptidyl-prolyl isomerization step (*37*). To test this hypothesis, refolding kinetics were performed in the presence of a peptidyl-prolyl isomerase (*39*). The periplasmic cyclophilin-type cis-trans peptidyl-prolyl isomerase from *E. coli* (eCyp) is an effective peptidyl-prolyl isomerase that catalyses the slow isomerization of peptide−proline bond (*39*). Table 1 shows the effect of increasing concentrations of eCyp on the refolding kinetic rate constants of the medium and slow phases. The relative amplitudes of the two kinetic phases were unaffected by the presence of eCyp, and the rate constants were determined as previously described. While the rate constant of the medium phase was practically unaffected by the presence of eCyp, the rate constant of the slow phase was accelerated approximately 4-fold in the presence of 0.3 μM eCyp suggesting the involvement of a cis−trans isomerization as a rate limiting step. Thus the slow phase clearly arose from a subpopulation of unfolded molecules with non-native cis peptidyl−proline bonds, the rate-limiting step for folding of this species being the isomerization of at least one *cis*-proline into the native trans conformation.

Refolding Model and Data Fitting. As already pointed out, the above analysis according to a sequential pathway, including only monomolecular reactions did not take into account for bimolecular reaction of monomer−monomer association. Nor did it include the heterogeneity of the denatured state. A more complex model therefore had to be imagined to provide a rigorous quantitative description of the folding process. This was done on the basis of the following observations: (i) the refolding kinetics were a multiphasic process with a burst phase followed either by two other phases (medium and slow) in CD or by three phases (rapid, medium and slow) in fluorescence. (ii) The medium phase was the unique phase dependent on the protein concentration which implies that the monomer−monomer association step occurred during the medium phase and not before. Hence monomeric folding intermediates were generated before the association step by processes occurring during the burst and rapid phases. (iii) The 200 ms lag observed in the FRET experiments confirmed the existence of an event occurring before the dimerization step. (iv) A proline isomerization was responsible for the slow refolding phase of R67 DHFR, as shown by the lack of dependence on denaturant concentration of the slow rate constant and by its acceleration in the presence of eCyp.

With these considerations in mind, several sequential and parallel refolding pathways have been simulated and a model describing the R67 DHFR refolding pathway at pH 5.0 could be proposed and is depicted in Scheme 1. The experimental evidences on which this model is based will be discussed later.

In Scheme 1, it is assumed that two subpopulations, named U_{cis} and U_{trans}, of monomeric intermediates are present in the unfolded state and differ from one another by the conformation of their prolyl residues.

Since the reactions $U_{cis} \rightarrow U_{trans}$ and $I_{cis} \rightarrow I_{trans}$ are spectroscopically indistinguishable, the proline isomerization, responsible for the slow phase, might occur either in the unfolded state or in the corresponding I_{cis} and I_{trans} burst intermediates generated extremely rapidly (less than 4 ms). The burst intermediate with all proline residues in trans conformation would be directly converted into a second intermediate I′ able to associate to form the native dimer, while the cis conformers would have to first isomerize to trans conformers before being converted into I′. From Scheme 1, a simplified fitting model 1 was formulated:

$$2S_{cis} \xrightarrow{k_1} 2I_{trans} \xrightarrow{k_2} 2I' \xrightarrow{k_3} D$$

where S_{cis} represents the slow folder species with proline residues in cis conformation (I_{cis} and U_{cis} from Scheme 1) and I_{trans} represents a monomeric intermediate with proline residues in trans conformation, I′ represents a second monomeric intermediate and D the native dimer. k_1 is the slow cis−trans proline isomerization rate constant, k_2 the rapid intermediate rate constant, and k_3 the dimerization rate constant. According to this model, the typical kinetic traces described in Figure 1A,B were processed using the nonlinear least-squares fitting described in Materials and Method with

Table 2: Comparison of the Rate Constant Values of the Three Refolding Kinetic Phases of R67 DHFR at pH 5.0[a]

	fitting method 1[b]	fitting method 2[c]
phase	rate constant	rate constant
rapid (k_2)	46 s^{-1}	23 s^{-1}
medium (k_3)	0.56×10^6 M^{-1} s^{-1} [d]	$0.66 \pm 0.03 \times 10^6$ M^{-1} s^{-1}
slow (k_1)	0.040 s^{-1}	0.036 ± 0.012 s^{-1}

[a] All rate constants are determined from the fitting of the experimental kinetic traces described in Figure 1 at 25°C in the standard refolding buffer conditions (0.1 M GuHCl in MTA buffer, pH 5.0), except the k_2 value extrapolated at 0.1 M residual GuHCl from the denaturant dependence of the rapid phase rate constant. [b] The fitting method 1 consists of the fitting of the experimental data on a sum of two or three exponential functions. [c] The fitting method 2 is based on the model 1 and described in Materials and Methods. [d] The bimolecular rate constant k_3 is calculated from the slope in the linear portion of the protein concentration dependence of the corresponding first-order rate constant (Figure 2).

the relative cis−trans protein population fixed to 15% (see below, Double-Jump Experiments). This procedure provided good fits of the fluorescence and CD traces with low values for the sum of the residuals squared and random distribution of the residuals (data not shown) and allowed to estimate simultaneously the two rate constants k_1 and k_3 at 25 °C and pH 5.0 (Table 2).

The slow rate constant value was close to that estimated previously according to a monomolecular reaction and was in good agreement with peptidyl proline isomerization rate constants reported in the literature for a variety of proteins (*40*, *41*).

The medium rate constant value was consistent with the bimolecular rate constant value estimated from fitting the FRET experiments (Figure 4A) on an irreversible dimerization and in the range of values reported for monomer−monomer association rate constants of various oligomeric proteins (*5*, *14*, *15*).

The rapid rate constant k_2 could be determined according to model 1 from the kinetic traces obtained in the presence of residual GuHCl. The extrapolation from the denaturant dependence of the rate constant at 0.1 M residual GuHCl yielded a k_2 value of 23 s^{-1} close to that determined from the rate values fitted on an exponential function.

Double-Jump Experiments. To determine the relative amounts of U_{cis} and U_{trans} in the unfolded state, double-jump experiments were performed. The native protein was first rapidly mixed with the denaturing buffer, incubated for various times (ranging between 1 s and 45 min) and diluted into refolding buffer. The refolding kinetics monitored by fluorescence were fit according to model 1 as indicated above except that the fraction of slow folding species (i.e., species in cis conformation) was allowed to vary. After one second in the denaturing buffer, nearly all the proline residues are expected to still be in the trans conformation and the fraction of slow folder species should be negligible. The corresponding refolding kinetic was therefore fit according to model 1 starting from the I_{trans} state. That provided a good fit of the trace, and the resulting rate constants k_2 and k_3, as well as the specific signals of I_{trans}, I′, and D, were used as first estimates for the fitting of the other kinetics (from 2 s to 45 min of denaturation time). The fitting of the complete set of double-jump kinetics allowed the determination of the fraction of slow folding species for each denaturation time.

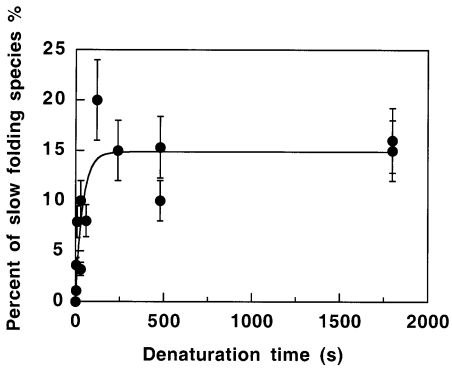


FIGURE 6: Stopped-flow fluorescence double-jump experiments The native protein (5 μM of monomer) was unfolded for varying times as described in Material and Methods and then refolded at 0.1 M residual GuHCl in MTA pH 5.0. The percentage of slow folding species (S_{cis}), determined from the fitting of the refolding kinetics according to the model 1, is plotted as a function of the denaturation time. The refolding kinetics starting after 1 s of denaturation were fit with a fixed value of $S_{cis} = 0$.

Figure 6 shows its variation expressed in percent of the total protein population as a function of the time of denaturation. The fraction of slow folding species reached a plateau corresponding to 15% ± 5% of the total protein population after 250 s of denaturation and did not affect the rate constants k_1, k_2, and k_3. All the fits reported here were therefore started with an initial proportion of 15% of cis species that was allowed to vary between 10% and 20% during the fitting.

From the Unfolded Monomer to the Native Tetramer. To investigate the complete folding of the tetrameric functional R67 DHFR, folding kinetics had to be performed at pH 8.0, where the dimer−tetramer equilibrium is shifted toward the tetramer. It could however be envisaged that the pH might also affect the process of folding into dimer described above. To test this possibility, we used a R67 DHFR variant S59A which folds into a stable dimer but is unable to associate into tetramers whatever the pH (*21*). Refolding kinetics of this variant were investigated both at pH 8.0 and pH 5.0 and showed that the pH affected neither the rate constants nor the amplitudes of the various phases of dimer folding (data not shown). Refolding kinetics of wild-type R67 DHFR at pH 8.0 were therefore investigated in the stopped flow to relate kinetically the dimerization and tetramerization steps. The specific activity of R67 DHFR measured at the outlet of the stopped flow was that of the native enzyme, confirming that the protein obtained was the native tetramer. Monitoring the intrinsic fluorescence provided complex kinetic trace (Figure 7). The fluorescence intensity rapidly increased during the first 2 s, reached a maximum after about an additional 1 s, and finally decreased slowly to reach a plateau after about 100 s. The first fluorescence increase was already observed in Figure 1A and corresponded to the transition of the W45 from a hydrophilic to an hydrophobic environment during monomer refolding and dimer formation. The second part of the kinetics reflected a fluorescence decrease that could be correlated with the tetramer appearance since at 350 nm the tetramer specific fluorescence is about 60% that of the dimer (*21*).

The kinetic traces at pH 8.0 were analyzed according to the following model 2 (see below), corresponding to model

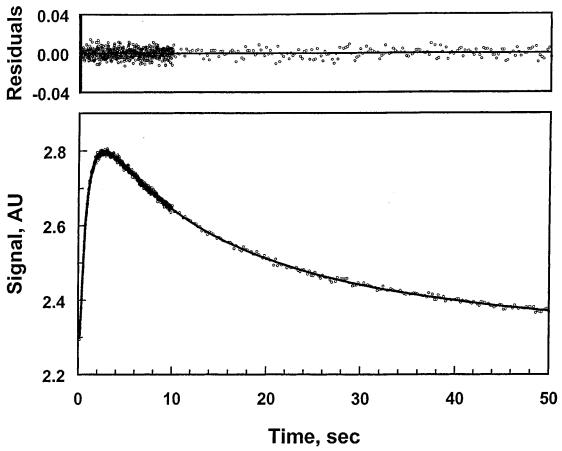


FIGURE 7: Overall refolding kinetic of R67 DHFR at pH 8.0 Experiments similar to those depicted in Figure 1A and monitored in fluorescence (λ_{exc} = 295 nm; emission measured above 350 nm) were performed at pH 8.0 and 25 °C. The trace depicted results from the accumulation and averaging of 10 kinetic runs. The fluorescence was recorded for a duration of 50 s on one channel with a sampling interval of 200 ms and for 10 s with a sampling interval of 10 ms on a second channel. For both recordings, the filtering constant was 10 ms. The solid line represents the best fit to the experimental points according to the model 2 as described in Material and Methods, and the residuals of this fit are shown above the kinetic trace.

1 completed with a final tetramerization step:

$$4S_{cis} \xrightarrow{k_1} 4I_{trans} \xrightarrow{k_2} 4I' \xrightarrow{k_3} 2D \xrightarrow{k_{ass}} T$$

In this model, T represents the native R67 DHFR tetramer and k_{ass} the dimer−dimer association rate constant. Taking as starting estimates the k_1, k_2, and k_3 values obtained from the refolding kinetics at pH 5.0 and the k_{ass} value of 40×10^3 M^{-1} s^{-1} determined in a previous work (*21*), the fitting of the kinetic traces provided good fit and random distribution of the residuals (Figure 7). A reliable estimate of the dimer−dimer association rate constant ($k_{ass} = 58 \times 10^3$ M^{-1} s^{-1}) could thus be obtained. The good agreement between this estimate and the reported value suggested that the tetramer refolding kinetics display a unique visible additional phase compared to the dimer refolding kinetics.

Because the cosubstrate NADPH binds to R67 DHFR only when the enzyme is tetrameric, the tetramerization step could be specifically monitored by observing the energy transfer between the tryptophan residues (W38) from the dimer−dimer interface and the NADPH bound inside the tetramer. Refolding kinetics (data not shown) performed at pH 8.0 in the presence of an excess of NADPH and observed at 420 nm (λ_{exc} = 295 nm) showed a 500 ms lag without any fluorescence variation, followed by a fluorescence increase corresponding to the appearance of the tetramer. Analysis of the kinetic trace according to the complete model 2 provided a good fit with a set of rate constants (k_1, k_2, k_3, k_{ass}) close to the previous values. This result brings additional direct evidence indicating that the tetramer results from the association of two dimers such as those obtained at pH 5.0 and that no observable dimer isomerization separates the dimerization and tetramerization steps. Thus, the model 2 we proposed is able to describe the entire refolding process at pH 8.0. This is, however, a minimal model, and it does not exclude that there might exist additional rapid phases that could not be observed in our experimental conditions.

DISCUSSION

As a complement to previously reported thermodynamic investigations performed at equilibrium on an oligomeric protein, the homotetramer of R67 DHFR (*26*, *42*), the experiments reported above deal with kinetic aspects of its refolding. They were aimed at understanding how the development of quaternary interactions is coordinated with the formation of secondary and tertiary structure during the folding reaction and, more specifically, the role of long-range interactions in the folding of an all β-sheet structure.

The R67 DHFR refolding kinetics were performed at pH 5.0 (leading to the stable dimer) and at pH 8.0 (leading to the native tetramer) and were monitored by three independent criteria: far UV CD, which mainly reflects the formation of secondary structure; intrinsic fluorescence, which characterizes the polarity of the environment of the tryptophan residues located at each protein−protein interface; and energy transfer between two labeled monomeric species, which allows direct detection of the monomer−monomer association.

The experiments reported above led to the identification and the characterization of four distinct phases during R67 DHFR dimer folding: a burst, a rapid, a medium, and a slow phase. The burst phase occurs during the 4 ms dead time of our measurements. The rapid phase is a unimolecular process with a rate constant of 23 s^{-1}. The medium phase corresponds to a second-order process with a rate constant of 0.66×10^6 M^{-1} s^{-1} and reflects the dimerization step. The slow phase with a rate constant of 0.036 s^{-1} corresponds to the isomerization of cis prolyl−peptide bonds of either the unfolded proteins, or some very early intermediates into the native trans conformation. The structural changes associated with each phase will now be discussed and integrated into a consistent description of the refolding of R67 DHFR from the unfolded monomer into the native tetramer.

The Burst Intermediate. The first evidence indicating that an early intermediate was formed in a burst phase was provided, in typical refolding kinetics of R67 DHFR, by the strong increase of the negative ellipticity amplitude that occurred in the stopped-flow dead time. Usually, such very fast far UV CD signals are interpreted as an indication that a significant fraction of secondary structures have been building up during the burst phase. However, it was shown that the aromatic side chains of W38 and W45 residues contribute significantly to the far-UV CD signal of R67 DHFR (*43*, *44*). It therefore could also be assumed that the far-UV CD change observed during the burst in R67 DHFR folding might reflect the tight packing of side chains that might occur without formation of secondary structure. This alternative appears unlikely on the basis of the following considerations. For tryptophan side chains to produce significant ellipticity signals, they must be rigidly immobilized in an asymmetrical environment, which usually is strongly hydrophobic. Indeed, during the R67 DHFR folding burst phase, the intrinsic fluorescence increases, indicating that the tryptophan residues enter a hydrophobic environment. However, it is a frequent observation that, when

a far UV−CD burst is observed in the folding of a protein, the near-UV CD signal does not change during the burst phase although tryptophan residues become shielded from the solvent (*40*, *45*). The interpretation is that aromatic residues, though in a hydrophobic environment, are still loosely packed and, hence, cannot give rise to a far UV CD signal at the end of the burst. Thus, the far UV signal observed at the end of the burst phase in R67 DHFR folding most probably arises from secondary structure formation.

Such a very rapid formation of secondary structure in an all-β protein might seem surprising since most β-sheet proteins fold more slowly than all α-proteins or mixed α,β-proteins. Yet some exceptions have been observed, such as the very rapid folding (time constant about 1.5 ms) of the all-β cold shock protein CspB (*46*).

Finally, because the second-order step leading to the dimer is observed later on the pathway (see association step) and because no energy transfer could be detected during the burst in the FRET-monitored kinetics, it can be concluded that the burst intermediate is a monomer. Thus, the intermediate resulting from the burst phase is monomeric and has a significant secondary structure content (native or non-native), and its tryptophan residues shielded from water.

The Rapid Phase. The rapid phase, located in the proposed sequential pathway between the burst phase and the association step, was first detected as a fluorescence decrease of very low amplitude suggesting minor local rearrangements in the environment of tryptophan residues. However, far UV CD failed to unveil it, indicating that either no secondary structure change is involved in this step or the structural change is so small that it is spectroscopically silent. The last possible explanation is that this step involves redistribution, but no change in the total amount of secondary structure.

The conformational change associated to the rapid phase is a prerequisite for the association. This is shown by the refolding kinetics monitored by energy transfer (Figure 4B). Indeed, at 0.7 M GuHCl and 12 °C, where the rapid phase was strongly slowed, FRET-monitored kinetics showed a lag of 1 s that preceded the signal decrease due to association. No association reaction occurred during the first second, while much of the intrinsic fluorescence variation corresponding to the rapid phase occurred within this time. Hence, the formation of the intermediate I′ is a prerequisite for the dimerization step.

It can therefore be concluded that the monomeric intermediate I, formed during the burst phase, though having a significant secondary structure content, is association incompetent, presumably because it has an ill-defined tertiary structure at the monomer−monomer interface. The tertiary rearrangement occurring during the rapid phase would lead to a second monomeric intermediate I′ that is association-competent. One could speculate that this tertiary rearrangement involves the monomer−monomer interface. For association to take place, W45 from one monomer should be exposed to the solvent to be able to interact with W45 from a second monomer. The transient exposure to the solvent of W45 would result in a decrease of its fluorescence above 350 nm that could correspond to the small fluorescence change observed during the rapid phase.

The Association Step. The medium phase was characterized by the largest change in fluorescence and CD signals (65% and 45% of the total observable amplitude change respectively) and was directly visualized by the energy transfer occurring between two labeled monomers. R67 DHFR contains two tryptophans per monomer, one (W45) located at the monomer−monomer interface, the other (W38) at the dimer−dimer interface in the tetrameric protein (*27*, *28*, *31*). Studies on the effect of single-tryptophan mutations on the R67 DHFR fluorescence (*44*) suggested that the fluorescence signal due to W45 should decrease upon refolding and thus could not account for the fluorescence increase observed either during the burst, or during the medium phase. It therefore seems likely that the fluorescence change is dominated by W38. Then, during the association, changes of environment occur not only around W45, but also around W38. This indicates that the association of two partly folded monomers occurs, within the resolution of our stopped-flow experiments, simultaneously with a secondary and tertiary structure rearrangement involving a large part of the polypeptide chain.

Concentration dependence experiments revealed that this phase consisted in a second-order process and a rate constant of about 0.66×10^6 M^{-1} s^{-1} was determined from fitting analysis according to the model 1. This value is similar to the values reported for association rate constants of several other dimeric proteins (*5*, *14*, *15*). Folding studies on small dimeric proteins have shown a variety of molecular mechanisms and a broad range of rate constants for the association step: (i) association via a collision of two partially folded monomers (P22 Arc repressor) (*15*), (ii) association occurring concomitantly with the hydrophobic collapse (Trp repressor) (*12*), and (iii) formation of stable monomer prior dimerization (λ repressor) (*47*). Three features seem to influence the association rate: (i) the number of residues involved at the interface (the association is faster when the number of residues relative to the whole sequence is high), (ii) the presence of hydrophobic clusters, and (iii) the possibility for the primary interaction to be non-native like. All these conditions are satisfied in the case of the Trp repressor where the association is nearly diffusion limited ($k_a = 10^8$ M^{-1} s^{-1}).

This is obviously not the case of R67 DHFR, where the contact area between the two monomers forming the native dimer is one of the smallest observed (535 Å^2). It consists of an intersubunit β-barrel formed by the interaction between the β-sheets from each monomer. Only three residues, G44, W45, and Y46, constitute the intersubunit hydrophobic core. A network of van der Waals interactions with C47, T48, and M26 from the pre-strand a loop stabilizes this interaction. The low number of residues involved in the interface imposes that residues should be all precisely adjusted to provide a stable interaction. This is evidenced by the fact that the association can be strongly affected even by a single point mutation (*48*). Taken together, these arguments explain why the monomer−monomer association rate constant of R67 DHFR is much slower than that expected for a diffusion-limited process, and why some local secondary and tertiary structuration is needed prior to association.

Formation of Native Tetramer. For some proteins (*12*, *49*) the dimerization step leads to the formation of a non-native dimeric intermediate and must be followed by further conformational rearrangements before the native state is reached. This seems not to be the case for R67 DHFR folding. Indeed, analysis of the R67 DHFR refolding kinetic with model 1 at pH 5.0 showed that adding a post-association

isomerization did not provide a better fit of the traces. Thus, if such an additional step did exist, it should be either very rapid (and hence appear coupled to the association) or spectroscopically silent. The kinetics obtained at pH 8.0 for the folding of denatured monomers U into native tetramers T could be well fit by just introducing one additional association step (tetramerization) to model 1 (see model 2). Furthermore, this fitting provided an association rate constant very close to that determined independently from experiments where R67 DHFR dimers were incubated for hours at pH 5.0 and thus had ample time to reach the "native" dimeric state before tetramerization was induced by a pH jump (*21*). That these native dimers associated at pH 8.0 with the same rate constant as that observed during the folding at that pH ruled out the possibility of an optically silent rate-limiting isomerization step between the dimer and the tetramer during R67 DHFR folding. Thus, the R67 DHFR dimer resulting from the association of two monomers is able to rapidly associate into the tetramer and therefore, should be either native or, at least, have the native loop connecting β-strands c and d, involved in the future dimer−dimer interaction.

CONCLUSION

The above considerations lead us to conclude the following:

(i) The presence of *cis*-proline in some unfolded R67 DHFR molecules is responsible for the slow phase. Isomerization of those prolines to the native trans configuration, occurring either in the unfolded state or within in an early monomeric intermediate, is needed for the folding to proceed.

(ii) At least two types of partially folded monomeric intermediate species, I and I′, have been shown to exist, while no evidence exists as to whether their secondary structure elements are nativelike. That I′ must be formed for association to take place strongly suggests that it should contain some nativelike tertiary structure features.

(iii) The dimerization is simultaneous with a kinetically indistinguishable completion of the secondary and tertiary structures of the dimer.

(iv) The dimers thus formed are native, or very close to native and associate into the functional tetramer.

Thus, as previously reported for many other small dimeric proteins (*14*, *49*, *50*), the folding of R67 DHFR appears to proceed via a step where secondary, tertiary, and quaternary structure formation are tightly coupled. However, unlike what was suggested by equilibrium studies, monomeric, partly folded intermediates do exist. One of them, I′, even appears to be an obligatory intermediate on the pathway to association-renaturation. It is noteworthy that R67 DHFR is an all β-sheet protein and that the existence of transient β-sheet folding intermediate is poorly documented.

Our study provides a striking example of the interest of a kinetic approach, as a complement to equilibrium folding/unfolding studies, in understanding complex biochemical processes such as the folding of a polypeptide chain. Indeed, equilibrium studies often fail to reveal the existence of transient, marginally stable intermediates, despite their essential role in the progress of the folding reaction. Complementary experiments using protein engineering are currently underway in the laboratory to gain a more precise description of the intermediate ensembles and to reach a better understanding of this "folding by association" process.

ACKNOWLEDGMENT

We thank Dr. Thierry Rose for initial genetic constructs and Dr. Alain Chaffotte for his helpful assistance with the Figure P software and his kinetic expertise.

REFERENCES

1. Eaton, W. A., Munoz, V., Thompson, P. A., Chan, C. K., and Hofrichter, J. (1997) *Curr. Opin. Struct. Biol. 7*, 10−14.
2. Nölting, B., Golbik, R., Neira, J. L., Sooler-Gonzales, A. S., Schreiber, G., and Fersht, A. R. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 826−830.
3. Kim, P. S., and Baldwin, R. L. (1982) *Annu. Rev. Biochem 51*, 459−489.
4. Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−658.
5. Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117−237.
6. Jennings, P. A., Finn, B. E., Jones, B. E., and Matthews, C. R. (1993) *Biochemistry 32*, 3783−3789.
7. Ziegler, M. M., Goldberg, M. E., Chafotte, A. F., and Baldwin, T. O. (1993) *J. Biol. Chem. 268*, 10760−10765.
8. Mann, C. J., Shao, X., and Matthews, C. R. (1995) *Biochemistry 34*, 14573−14580.
9. Waldburger, C. D., Jonsson, T., and Sauer, R. T. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2629−2634.
10. Wendt, H., Berger, C., Baici, A., Thomas, R. M., and Bosshard, H. R. (1995) *Biochemistry 34*, 4097−4107.
11. Nichtl, A., Buchner, J., Jaenicke, R., Rudolph, R., and Scheibel, T. (1998) *J. Mol. Biol. 282*, 1083−1091.
12. Gloss, L. M., and Matthews, R. C. (1998) *Biochemistry 37*, 15990−15999.
13. Doyle, S. M., Braswell, E. H., and Teschke, C. M. (2000) *Biochemistry 39*, 11667−11676.
14. Wallace, L. A., and Dirr, H. W. (1999) *Biochemistry 38*, 16686−16694.
15. Milla, M. E., and Sauer, R. T. (1994) *Biochemistry 33*, 1125−1133.
16. Srivastava, A. K., and Sauer, R. T. (2000) *Biochemistry 39*, 8308−8314.
17. Zetina, C. R., and Goldberg, M. E. (1982) *J. Mol. Biol. 157*, 133−148.
18. Planchenault, T., Navon, A., Schulze, A. J., and Goldberg, M. E. (1995) *Eur. J. Biochem. 240*, 615−621.
19. Goldberg, M. E., Semisotnov, G. V., Friguet, B., Kawajima, K., Ptitsyn, O. B., and Sugai, S. (1990) *FEBS Lett.* 263, 51−56.
20. Carrel, R. W., and Lomas, D. A. (1997) *Lancet 350*, 134−38.
21. Méjean, A., Bodenreider, C., Schuerer, K., and Goldberg, M. (2001) *Biochemistry* 40, 8169−8179.
22. Pattishall, K. H., Acar, J., Burchall, J. J., Goldstein, F. W., and Harvey, R. J. (1977*) J. Biol. Chem. 252*, 2319−2323.
23. Amyes, S. G. B., and Smith, J. T. (1976) *Eur. J. Biochem. 61*, 597−603.
24. Smith, S. L., Stone, D., Novak, P., Baccanari, D. P., and Burchall, J. J. (1979) *J. Biol. Chem. 254*, 6222−6225.
25. Stone, D., and Smith, S. L. (1979) *J. Biol. Chem. 245*, 6222−25.
26. Nichols, R., Weaver, C. D., Eisenstein, E., Blakley, R. L., Appleman, J., Huang, T. H., Huang, F. Y., and Howell, E. E. (1993) *Biochemistry 32*, 1695−706.
27. Matthews, D. A., Smith, S. L., Baccanari, D. P., Burchall, J. J., Oatley, S. J., and Kraut, J. (1986) *Biochemistry 25*, 4194−204.
28. Narayana, N., Matthews, D. A., Howell, E. E., and Nguyen-huu, X. (1995) *Nat. Struct. Biol. 2*, 1018−25.
29. Bradrick, T. D., Beechem, J. M., and Howell, E. E. (1996) *Biochemistry 35*, 11414−24.
30. Park, H., Bradrick, T. D., and Howell, E. E. (1997) *Protein Eng. 10*, 1415−24.
31. Reece, L. J., Nichols, R., Ogden, R. C., and Howell, E. E. (1991) *Biochemistry 30*, 10895−904.
32. Zhuang, P., Eisenstein, E., and Howell, E. E. (1994) *Biochemistry 33*, 4237−44.
33. Smith, S., and Burchall, J. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4619−4623.
34. Kuzmic, P. (1996) *Anal. Biochem. 237*, 260−273.
35. Cai, K., and Schirch, V. (1996) *J. Biol. Chem. 271*, 27311−27320.

36. Svensson, M., Jonasson, P., Jonsson, B. H., Lindgren, M., Martensson, L. G., Gentile, M., Boren, K., and Carlsson, U. (1995) *Biochemistry 34*, 8606−8620.
37. Brandts, J. F., Halovorson, H. R., and Brennan, M. (1975) *Biochemistry 14*, 4953−4963.
38. Lang, K., Schmid, F. X., and Fisher, G. (1987) *Nature 329*, 268−270.
39. Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry 33*, 2761−2772.
40. Georgescu, R. E., Li, J.-H, Goldberg, M. E., Tasayco, M., and Chaffotte, A. F. (1998) *Biochemistry 37*, 10286−10297.
41. Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10436−10443.
42. Dam, J., Rose, T., Goldberg, M. E., and Blondel, A. (2000) *J. Mol. Biol. 302*, 235−250.
43. Woody, R. W. (1995) *Methods Enzymol. 246*, 34−70.
44. West, F. W., Seo, H.-S, Bradrick, T. D., and Howell, E. E. (2000) *Biochemistry 39*, 3678−3689.
45. Chaffotte, A. F., Guillou, Y., and Goldberg, M. E. (1992) *Biochemistry 31*, 9694−9702.
46. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) *Nat. Struct. Biol. 2*, 663−73.
47. Banik, U., Saha, R., Mandal, N. C., Bhattacharyya, B., and Roy, S. (1992) *Eur. J. Biochem. 206*, 15−21.
48. Martinez, M. A., Pezo, V., Marliere, P., and Wain-Hobson, S. (1996) *Embo J. 15*, 1203−10.
49. Kim, D.-H, Jang, D. S., Nam, G. H., and Choi, K. Y. (2001) *Biochemistry 40*, 5011−5017.
50. Rietveld, A. W. M., and Ferreira, S. (1998) *Biochemistry 37*, 933−937.

BI020453B