Apoflavodoxin Folding Mechanism: An α/β Protein with an Essentially Off-Pathway Intermediate†

J. Fernández-Recio,‡ C. G. Genzor,§ and J. Sancho*

Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Universidad de Zaragoza, 50009-Zaragoza, Spain

Received February 1, 2001; Revised Manuscript Received September 7, 2001

ABSTRACT: The folding reaction of *Anabæna* apoflavodoxin has been studied by stopped-flow kinetics and site-directed mutagenesis. Although the urea unfolding equilibrium is two-state, a transient intermediate accumulates during the folding reaction. The intermediate is monomeric, and it is not related to proline isomerization. Unlike many cases where the presence of an intermediate has been detected either by a burst phase or by the curvature, at low urea concentration, of the otherwise only observable kinetic phase, two kinetic phases are observed in apoflavodoxin folding whose total amplitude equals the amplitude of unfolding. To determine the role of the intermediate in the folding reaction, the apoflavodoxin kinetic data have been fitted to all conceivable three-species kinetic models (either linear or triangular). Using a stepwise fitting procedure, we find that the off-pathway mechanism explains most of the kinetic data (not a slow unfolding phase), the on-pathway mechanism being rejected. By using global analysis, good overall agreement between data and fit is found when a triangular mechanism is considered. The fitted values of the microscopic constants indicate that most of the unfolded molecules refold from the denatured state. Apoflavodoxin thus folds via a triangular, but essentially off-pathway, mechanism. We calculate that the retardation of the folding caused by the off-pathway intermediate is not large. Some unusual properties of the intermediate are discussed.

Many small proteins present two-state folding kinetics (*1*). Other proteins, however, even some whose conformational equilibrium is two-state, display more complex kinetics, with intermediates transiently accumulating during the reaction (*2*). When intermediate formation is fast, it is usually difficult to establish whether the intermediate appears on-pathway, as an obligatory conformation between the unfolded and the folded state, or off-pathway, as a trapped conformation from where the misfolded protein must escape through denaturation before entering the pathway leading to the folded state. In some cases, however, on-pathway intermediates have been described (*3*−*12*). When formation of the intermediate is slower and can be observed, the mechanism of the folding reaction can be assessed with greater confidence (*13*). The very role of on-pathway intermediates is, nevertheless, controversial as there is not overall agreement on whether they serve to accelerate the folding reaction or to slow it down (*2*, *14*).

We are studying protein stability and folding using the apoflavodoxin from *Anabæna* PCC 7119 as a model. The structure of the protein (*15*) is known (it is an α/β protein with an open central β-sheet), the gene has been cloned (*16*), the protein can be expressed in *E. coli*, and the conformational stability has been characterized (*17*). At neutral pH, the equilibrium urea denaturation conforms to a two-state model (*17*). Alternative equilibrium non-native conformations have been, however, observed when the native structure is destabilized by raising the temperature (*18*, *19*), by lowering the pH (*17*), or by truncation (*20*). Here we describe the folding and unfolding kinetics of apoflavodoxin at neutral pH. The unfolding kinetics are close to monophasic, but in the refolding reaction two major phases appear, indicating the presence of a folding intermediate. By global fitting of the macroscopic rate constants and amplitudes at different urea concentrations to different three-species folding mechanisms, the position of the intermediate in the folding reaction has been unambiguously established.

MATERIALS AND METHODS

Site-Directed Mutagenesis. Mutagenesis of the wild-type gene was performed as described before (*17*). The oligonucleotide 5′-ATATTCCAAGTAACACAGCCAATAA-3′ was used to mutate the only proline in flavodoxin (Pro55) to a valine. The mutation was identified by DNA sequencing (*21*).

Protein Expression and Purification. Apoflavodoxin Preparation and Quantitation. Protein expression and purification were performed as described (*17*). Apoflavodoxin was prepared from the holoflavodoxin, and its concentration was determined as described (*17*).

† This work has been supported by Grants PB97-1027 (DGES, Spain) and P15/97 (CONSI+ D, DGA, Spain). J.F.-R. is a recipient of a fellowship from the Spanish Ministerio de Educación y Cultura.

* To whom correspondence should be addressed. Phone: 34 976761286; Fax: 34 976762123; E-mail: jsancho@posta.unizar.es.

‡ Present address: Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037.

§ Present address: Operon S.A., Camino del Plano 19, Cuarte de Huerva, Zaragoza, Spain.

10.1021/bi010216t CCC: $20.00
Published on Web 11/21/2001

Unfolding and Refolding Kinetics. Unfolding and refolding kinetics were recorded using an Applied Photophysics stopped-flow apparatus (DX.17 MV model) with fluorescence detection. Unless indicated, the temperature was 25.0 ± 0.1 °C, the initial protein concentration 44 μM, and the buffer 50 mM MOPS,[1] pH 7.0, in all solutions. Ten volumes of the unfolding (or refolding) solution were typically mixed with 1 volume of folded (or unfolded) protein solution. The unfolded protein solution contained 4.0 M urea. The unfolding or refolding solutions contained different urea concentrations as required. Refolding to very low urea concentrations was studied by mixing 1 volume of apoflavodoxin in 10 mM NaOH (unfolded) with 10 volumes of 55 mM MOPS, pH 7 (containing 1 mM HCl and different urea concentrations).

Data Analysis. Experimental data (relaxation rates and amplitudes) were fitted to the equations obtained from analytical solution of the different kinetic models for a three-state system (see below). For each model, a global least-squares fit was performed, using a downhill simplex minimization method (*22*) implemented in the ICM software [Molsoft LLC; (*23*)] and the minimization routine implemented in Mathematica (Wolfram Research, Inc.). Standard errors for the parameters obtained in the global fitting were calculated by generating synthetic data sets from Gaussian distributions around the initial experimental data (*24*). After fitting different synthetic data sets, we obtained a probability distribution of estimated parameters. Standard deviations calculated from these distributions correspond very precisely to the standard errors for the actual parameters obtained in the global fitting of the experimental data.

Analytical Treatment of the Kinetic Mechanisms. Folding reactions involving three species give rise to two observable relaxations (λ_1, λ_2) with their corresponding amplitudes (*25*). Rigorous treatment of the different three-state kinetic mechanisms (Figure 1) gives analytical solutions that relate the observed rates and amplitudes with the microscopic kinetic rates.

Three-State Linear Mechanisms. In a linear mechanism (Figure 1a−g), the relaxations are related to the microscopic constants by the following equations:

$$\lambda_1 + \lambda_2 = k_1 + k_2 + k_3 + k_4 \tag{1}$$

$$\lambda_1\lambda_2 = k_1(k_3 + k_4) + k_2k_4 \tag{2}$$

from where analytical expressions for λ_1 and λ_2 can be derived:

$$\lambda_1 = \frac{k_1 + k_2 + k_3 + k_4 + \sqrt{(k_1 + k_2 + k_3 + k_4)^2 - 4(k_1k_3 + k_1k_4 + k_2k_4)}}{2} \tag{3a}$$

$$\lambda_2 = \frac{k_1 + k_2 + k_3 + k_4 - \sqrt{(k_1 + k_2 + k_3 + k_4)^2 - 4(k_1k_3 + k_1k_4 + k_2k_4)}}{2} \tag{3b}$$

On the other hand, the amplitudes associated with the observed kinetic phases can be solved as functions of the microscopic constants, the value of the observed property for each species, and the initial concentration of each of the three species (*25*−*27*). Importantly, unlike the apparent rate equations, the amplitude equations are different for each linear mechanism.

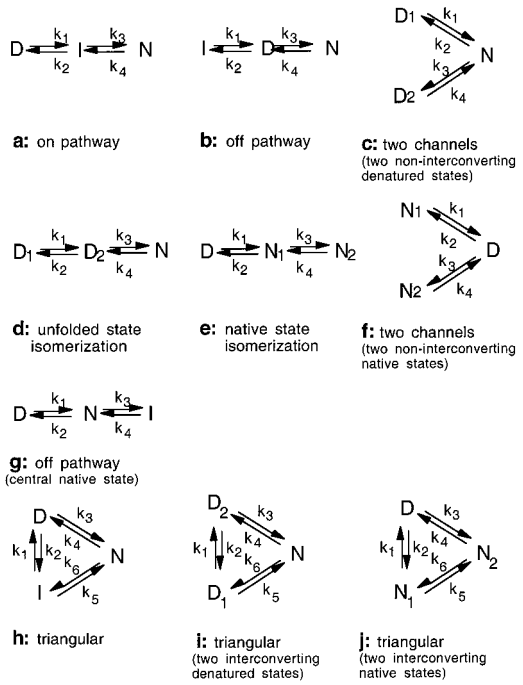

FIGURE 1: Folding mechanisms involving three species. Linear mechanisms: a−g. Triangular mechanisms: h−j.

For the *on-pathway* model (Figure 1a), the equations that relate amplitudes and microscopic rates for the unfolding reaction ($[I]_0 = [D]_0 = 0$) are

$$A_1^{unf} = \frac{F_D k_4k_2 - F_I k_4(\lambda_1 - k_1) + F_N k_4(\lambda_1 - k_1 - k_2)}{\lambda_1(\lambda_1 - \lambda_2)} \tag{4}$$

$$A_2^{unf} = \frac{-F_D k_4k_2 + F_I k_4(\lambda_2 - k_1) - F_N k_4(\lambda_2 - k_1 - k_2)}{\lambda_2(\lambda_1 - \lambda_2)} \tag{5}$$

and for refolding ($[N]_0 = [I]_0 = 0$):

$$A_1^{ref} = \frac{F_D k_1(\lambda_1 - k_3 - k_4) - F_I k_1(\lambda_1 - k_4) + F_N k_1k_3}{\lambda_1(\lambda_1 - \lambda_2)} \tag{6}$$

$$A_2^{ref} = \frac{-F_D k_1(\lambda_2 - k_3 - k_4) + F_I k_1(\lambda_2 - k_4) - F_N k_1k_3}{\lambda_2(\lambda_1 - \lambda_2)} \tag{7}$$

where F_N, F_I, and F_D are the fluorescence of the native, intermediate, and unfolded states, respectively, when they are at maximum concentration (equal to the sum of all initial kinetic species).

For each denaturant concentration, unfolding (or refolding) amplitudes can be normalized by dividing the experimental values by the maximum change of fluorescence expected during unfolding (or refolding) at that denaturant concentration. In complex mechanisms, the maximum fluorescence change is the difference between the fluorescence of all initial kinetic species existing in native conditions and the fluo-

[1] Abbreviations: MOPS, 3-(*N*-morpholino)propanesulfonic acid; DTT, dithiothreitol.

rescence of all initial kinetic species existing in denatured conditions. However, when there is only one populated state at either native or denaturant conditions (or when the populated states have the same fluorescence), reduced amplitudes can be calculated by dividing experimental amplitudes by $F_D - F_N$ (refolding) or by $F_N - F_D$ (unfolding). The sum of all reduced amplitudes (refolding and unfolding), calculated in this way, for each denaturant concentration, is always equal to 1.

Thus, for the on-pathway model, the reduced amplitudes for the unfolding reaction ($[I]_0 = [D]_0 = 0$) follow:

$$A_1^{\mathrm{RED,unf}} = \frac{k_4(\lambda_1 - k_1 - k_2) + F_I^{\mathrm{rel}} k_4(\lambda_1 - k_1)}{\lambda_1(\lambda_1 - \lambda_2)} \quad (8)$$

$$A_2^{\mathrm{RED,unf}} = -\frac{k_4(\lambda_2 - k_1 - k_2) - F_I^{\mathrm{rel}} k_4(\lambda_2 - k_1)}{\lambda_2(\lambda_1 - \lambda_2)} \quad (9)$$

and for refolding ($[N]_0 = [I]_0 = 0$):

$$A_1^{\mathrm{RED,ref}} = \frac{F_I^{\mathrm{rel}} k_1(\lambda_1 - k_4) - k_1 k_3}{\lambda_1(\lambda_1 - \lambda_2)} \quad (10)$$

$$A_2^{\mathrm{RED,ref}} = -\frac{F_I^{\mathrm{rel}} k_1(\lambda_2 - k_4) - k_1 k_3}{\lambda_2(\lambda_1 - \lambda_2)} \quad (11)$$

where

$$F_I^{\mathrm{rel}} = \frac{F_I - F_D}{F_N - F_D} \quad (12)$$

is the relative fluorescence of the intermediate, with a value of 0 if $F_I = F_D$, and of 1 if $F_I = F_N$. If the fluorescence of each species is assumed to vary linearly with urea concentration (*28*), the following relationship between relative fluorescence of the intermediate and urea concentration applies:

$$F_I^{\mathrm{rel}} = \frac{F_I^\circ + m_I[\mathrm{urea}] - F_D^\circ - m_D[\mathrm{urea}]}{F_N^\circ = m_N[\mathrm{urea}] - F_D^\circ - m_D[\mathrm{urea}]} \quad (13)$$

where F_N°, F_D°, and F_I° are the fluorescence values for the native, unfolded, and intermediate states in the absence of urea.

Microscopic rate constants depend on urea concentration according to eq 32 (*29*−*31*). Thus, observed relaxations (λ_1 and λ_2) and amplitudes (A_1 and A_2) will also depend on urea concentration.

For the *off-pathway* model (Figure 1b), similar equations can be derived that relate reduced amplitudes and microscopic rates for the unfolding reaction ($[I]_0 = [U]_0 = 0$):

$$A_1^{\mathrm{RED,unf}} = \frac{F_I^{\mathrm{rel}} k_4 k_2 + k_4(\lambda_1 - k_1 - k_2)}{\lambda_1(\lambda_1 - \lambda_2)} \quad (14)$$

$$A_2^{\mathrm{RED,unf}} = -\frac{F_I^{\mathrm{rel}} k_4 k_2 + k_4(\lambda_2 - k_1 - k_2)}{\lambda_1(\lambda_2 - \lambda_2)} \quad (15)$$

and for refolding ($[N]_0 = [I]_0 = 0$):

$$A_1^{\mathrm{RED,ref}} = \frac{F_I^{\mathrm{rel}} k_2(\lambda_1 - k_4) + k_3(\lambda_1 - k_1)}{\lambda_1(\lambda_1 - \lambda_2)} \quad (16)$$

$$A_2^{\mathrm{RED,ref}} = -\frac{F_I^{\mathrm{rel}} k_2(\lambda_2 - k_4) + k_3(\lambda_2 - k_1)}{\lambda_2(\lambda_1 - \lambda_2)} \quad (17)$$

For a *two-channel mechanism* (Figure 1c), where two different folding channels are considered, two relaxations (λ_1 and λ_2) are observed in the folding reactions, whose analytical expressions are also identical to those of the previously described models (eqs 3a, 3b). On the other hand, the reduced amplitudes for the unfolding reaction ($[D_1]_0 = [D_2]_0 = 0$) will be related to the microscopic rate constants by the following expressions:

$$A_1^{\mathrm{RED,unf}} = \frac{-F_{D_2}^{\mathrm{rel}}(k_3(\lambda_1 - k_1)) + (k_2(\lambda_1 - k_4) + k_3(\lambda_1 - k_1))}{\lambda_1(\lambda_1 - \lambda_2)} \quad (18)$$

$$A_2^{\mathrm{RED,unf}} = -\frac{-F_{D_2}^{\mathrm{rel}}(k_3(\lambda_2 - k_1)) + (k_2(\lambda_2 - k_4) + k_3(\lambda_2 - k_1))}{\lambda_2(\lambda_1 - \lambda_2)} \quad (19)$$

And the reduced amplitudes for the refolding reaction ($[N]_0 = 0$) will be as follows:

$$A_1^{\mathrm{RED,ref}} = \frac{[D_1^{4M}(k_1(\lambda_1 - k_4)) + D_2^{4M}(k_4(\lambda_1 - k_1))] - F_{D_2}^{\mathrm{rel}}[D_1^{4M}(k_1 k_3) + D_2^{4M}(k_4(\lambda_1 - k_1 - k_2))]}{(D_1^{4M} + D_2^{4M})\lambda_1(\lambda_1 - \lambda_2)} \quad (20)$$

$$A_2^{\mathrm{RED,ref}} = -\frac{[D_1^{4M}(k_1(\lambda_2 - k_4)) + D_2^{4M}(k_4(\lambda_2 - k_1))] - F_{D_2}^{\mathrm{rel}}[D_1^{4M}(k_1 k_3) + D_2^{4M}(k_4(\lambda_2 - k_1 - k_2))]}{(D_1^{4M} + D_2^{4M})\lambda_2(\lambda_1 - \lambda_2)} \quad (21)$$

According to this model, both denatured states, D_1 and D_2, can be populated at high denaturant conditions, so the terms D_1^{4M} and D_2^{4M} in eqs 20 and 21 are the concentrations of these denatured states, D_1 and D_2, at 4 M urea (initial conditions in refolding). D_1^{4M} and D_2^{4M} can be calculated from the microscopic rate constants and their dependence on urea concentration (eq 32), using eq 22:

$$\frac{D_2^{4M}}{D_1^{4M}} = \frac{k_1 k_3}{k_2 k_4} = \frac{k_1^0 e^{m_1 4.0/RT} k_3^0 e^{m_3 4.0/RT}}{k_2^0 e^{m_2 4.0/RT} k_4^0 e^{m_4 4.0/RT}} \quad (22)$$

$F_{D_2}^{\mathrm{rel}}$ is the relative fluorescence of the denatured state D_2 according to

$$F_{D_2}^{\mathrm{rel}} = \frac{F_{D_2} - F_{D_1}}{F_N - F_{D_1}} \quad (23)$$

If the two denatured states (D_1 and D_2) display the same fluorescence, $F_{D_2}^{\mathrm{rel}} = 0$. Accordingly, the dependence of $F_{D_2}^{\mathrm{rel}}$ on urea concentration is equivalent to that of F_I^{rel} in eq 13.

Triangular Mechanism. In a triangular mechanism (Figure 1h−j), folding reactions also give rise to two relaxation constants (λ_1 and λ_2) with their corresponding amplitudes (A_1 and A_2). These relaxations are related to the microscopic constants by the following equations:

$$\lambda_1 + \lambda_2 = k_1 + k_2 + k_3 + k_4 + k_5 + k_6 \quad (24)$$

$$\lambda_1\lambda_2 = k_1(k_3 + k_4 + k_6) + k_2(k_4 + k_5 + k_6) + k_3(k_5 + k_6) + k_4k_5 \quad (25)$$

from where general analytical expressions for λ_1 and λ_2 can be derived:

$$\lambda_1 = \frac{k_1 + k_2 + k_3 + k_4 + k_5 + k_6}{2} + \frac{\sqrt{(k_1 + k_2 + k_3 + k_4 + k_5 + k_6)^2 - 4(k_1(k_3 + k_4 + k_6) + k_2(k_4 + k_5 + k_6) + k_3(k_5 + k_6) + k_4k_5)}}{2} \quad (26a)$$

$$\lambda_2 = \frac{k_1 + k_2 + k_3 + k_4 + k_5 + k_6}{2} - \frac{\sqrt{(k_1 + k_2 + k_3 + k_4 + k_5 + k_6)^2 - 4(k_1(k_3 + k_4 + k_6) + k_2(k_4 + k_5 + k_6) + k_3(k_5 + k_6) + k_4k_5)}}{2} \quad (26b)$$

In a triangular model such as that in Figure 1h, the reduced amplitudes associated with the observed exponential phases are given by the following analytical expressions:

$$A_1^{\mathrm{RED,unf}} = \frac{-F_I^{\mathrm{rel}}(k_6(\lambda_1 - k_2 - k_3) - k_2k_4) + (k_4(\lambda_1 - k_1 - k_2 - k_5) + k_6(\lambda_1 - k_1 - k_2 - k_3))}{\lambda_1(\lambda_1 - \lambda_2)} \quad (27)$$

$$A_2^{\mathrm{RED,unf}} = -\frac{-F_I^{\mathrm{rel}}(k_6(\lambda_2 - k_2 - k_3) - k_2k_4) + (k_4(\lambda_2 - k_1 - k_2 - k_5) + k_6(\lambda_2 - k_1 - k_2 - k_3))}{\lambda_2(\lambda_1 - \lambda_2)} \quad (28)$$

$$A_1^{\mathrm{RED,ref}} = \frac{F_I^{\mathrm{rel}}(k_2(\lambda_1 - k_4 - k_6) - k_3k_6) + (k_3(\lambda_1 - k_1 - k_5) - k_2k_5)}{\lambda_1(\lambda_1 - \lambda_2)} \quad (29)$$

$$A_2^{\mathrm{RED,ref}} = -\frac{F_I^{\mathrm{rel}}(k_2(\lambda_2 - k_4 - k_6) - k_3k_6) + (k_3(\lambda_2 - k_1 - k_5) - k_2k_5)}{\lambda_2(\lambda_1 - \lambda_2)} \quad (30)$$

In triangular mechanisms (Figure 1h−j), eq 31 applies at any denaturant concentration:

$$k_1k_3k_6 = k_2k_4k_5 \quad (31)$$

Therefore, k_6 can be defined as a function of the other constants in the analytical expressions of the relaxation constants and amplitudes (eqs 26−30). Fitting of these equations to experimental data will yield k_1, k_2, k_3, k_4, and k_5 values, from which that of k_6 (and of its dependence on urea concentration according to eq 32) can be derived (eq 31).

RESULTS

Folding and Unfolding Kinetics of Wild-Type Recombinant Apoflavodoxin. The unfolding of recombinant wild-type apoflavodoxin can be fitted to a double exponential (Figure 2a). At high urea concentration, the faster phase accounts for about 95% of the total fluorescence change, and the second, minor, phase for the remaining 5% (Figure 2c). The observed rate constant of the major phase follows a typical urea dependence (Figure 2b) with a minimum at the reported equilibrium denaturation midpoint of around 2 M (*17*). The fluorescence decay observed in the unfolding kinetics equals the difference in fluorescence between the fully folded and fully unfolded states in equilibrium (not shown). The relative amplitude and the rate constant of the small, slow, unfolding phase are independent of protein concentration (from 1 to 18 μM; data not shown). They are also independent of the presence of 1 mM DTT (that would reduce putative dimers arising from oxidation of the single, buried, cysteine residue in apoflavodoxin, Cys54).

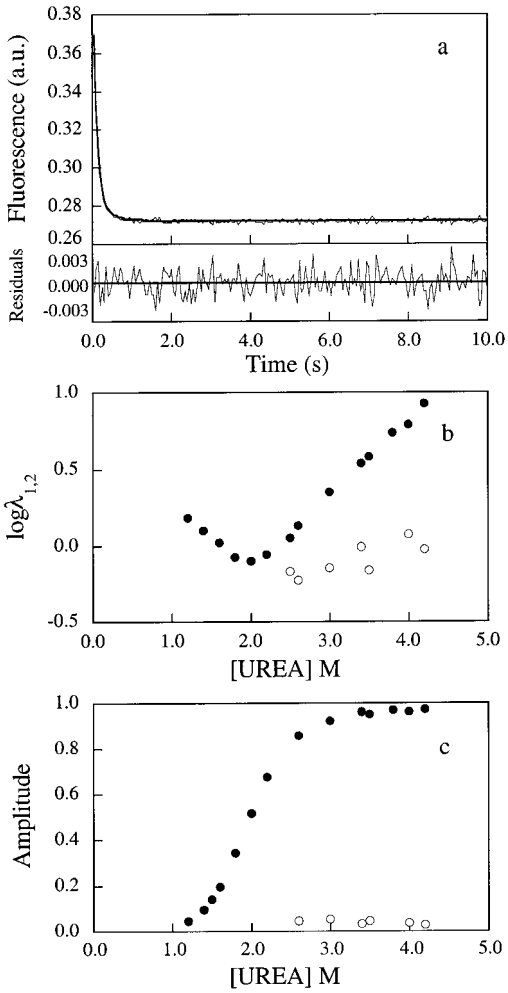

FIGURE 2: (a) Kinetic trace of apoflavodoxin unfolding in 4 M urea at pH 7.0 (25 °C) fitted to a double exponential. (b) Dependence of observed relaxations (in units of s^{-1}) on urea concentration (filled circles, fast phase; open circles, slow phase). (c) Reduced amplitudes (filled circles, fast phase; open circles, slow phase).

The refolding of urea-unfolded recombinant wild-type apoflavodoxin can also be fitted to a double exponential (Figure 3a). In this case, however, the amplitudes of the two phases are comparable, and their relative values vary with urea concentration (Figure 3c). The faster phase displays a typical 'V' shape, with the minimum around the mid-denaturation urea concentration, while the slower phase is markedly curved at low urea concentrations (Figure 3b).

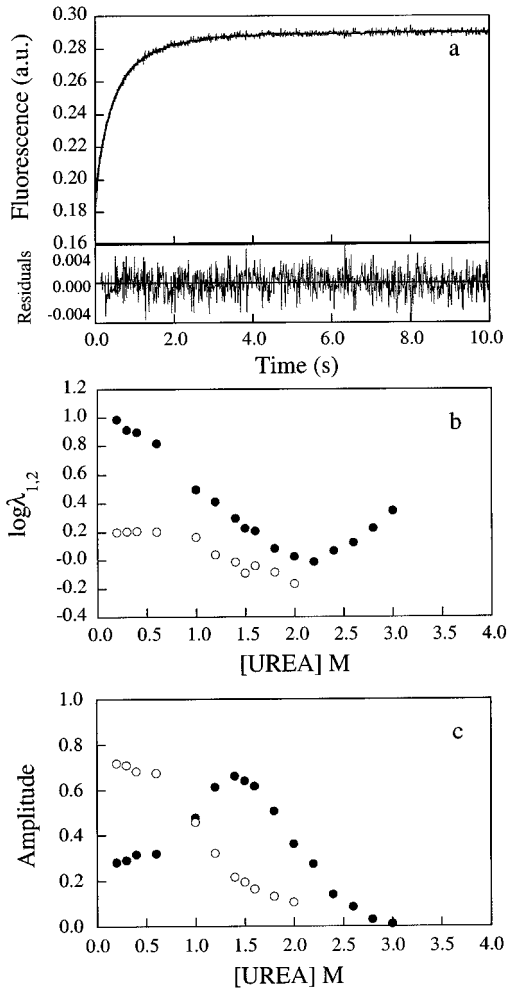

FIGURE 3: (a) Kinetic trace of apoflavodoxin refolding in 0.4 M urea at pH 7.0 (25 °C) fitted to a double exponential. (b) Dependence of observed relaxations (in units of s^{-1}) on urea concentration (filled circles, fast phase; open circles, slow phase). (c) Reduced amplitudes (filled circles, fast phase; open circles, slow phase).

Expression, Equilibrium Characterization, and Folding Kinetics of the Pro55Val Mutant. The Pro55Val mutant was expressed like wild type, but the yield was lower. Part of the mutant protein was recovered as apoprotein, indicating that the binding of the FMN cofactor was weakened. Since the conformational stability of wild-type apoflavodoxin at pH 7.0 is low (4.1 kcal mol^{-1}) and the mutation introduced is not conservative, there was a chance that the mutant was severely destabilized. To test this possibility, we analyzed the urea unfolding curve (data not shown). According to our data, the stability of the mutant is of 4.3 kcal mol^{-1}, which suggests Pro55Val constitutes a valid model to investigate any possible influence of the only proline residue of apoflavodoxin on its folding mechanism.

To this end, we measured the folding and unfolding kinetics of the mutant. Both the major and minor unfolding phases were observed, and the two refolding phases as well (Figure 4). The overall behavior of the proline-lacking mutant is, thus, very similar to that of wild type, ruling out any involvement of Pro55 in the observed complexities of the folding reaction.

Double-Jump Interrupted Unfolding Kinetics. Non-proline cis peptide bonds could also, in principle, lead to folding

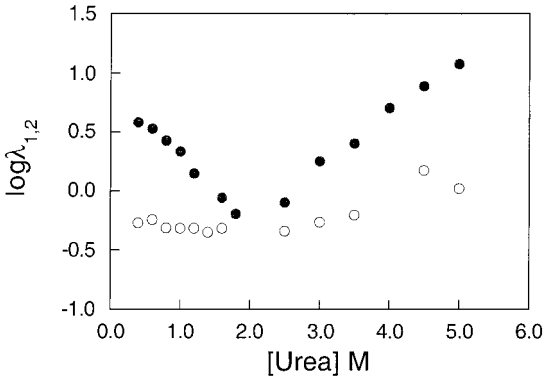

FIGURE 4: Dependence of the observed folding and unfolding relaxation constants of Pro55Val apoflavodoxin (filled circles, fast phase; open circles, slow phase).

intermediates (*32*, *33*). Although there is no crystallographic evidence of any such bond in apoflavodoxin (*15*), it has been pointed out that non-proline cis peptide bonds may sometimes pass undetected (*34*). The concentration of slow folding species arising from isomerization in the unfolded state can be markedly reduced by performing double-jump kinetics where the unfolded species is generated and, shortly afterward, subjected to refolding. We have performed double-jump interrupted/unfolding−refolding experiments at different unfolding times from 0.1 to 0.6 s. As typical cis−trans isomerization processes in proteins occur slowly [generally over a time scale of 10−100 s; (*35*)], they should not take place within the 0.1−0.6 s incubation time used for the experiment. In the double-jump experiments, two kinetics phases are observed (not shown) whose relative amplitudes and rate constants are independent of the time of unfolding and identical to those obtained from aged unfolded protein solutions. This rules out that the two observed refolding kinetic phases arise from heterogeneity in the unfolded state due to a cis−trans equilibrium.

Refolding Kinetics of Recombinant Wild-Type Apoflavodoxin as a Function of Protein Concentration. Molten globule equilibrium intermediates often tend to aggregate, and kinetic intermediates are known to resemble molten globules (*36*−*38*). On the other hand, it has been shown that transient aggregates may accumulate in the folding of proteins (*39*, *40*) and that they can be easily mistaken for folding intermediates (*41*). To investigate if the observed biphasic refolding of apoflavodoxin reflects the accumulation of a dimeric intermediate, we studied the influence of apoflavodoxin concentration on the observed refolding kinetic constants and relative amplitudes. No concentration effects were observed from 0.5 to 14.0 μM final apoflavodoxin concentration (Figure 5), which indicates that the observed intermediate is monomeric.

Kinetics of Nonrecombinant Wild-Type Apoflavodoxin, Directly Purified from Cyanobacteria. Throughout this work we use recombinant flavodoxin that, in principle, only differs from the original cyanobacterial flavodoxin in the N-terminal residue (*16*). To be sure that expression of the flavodoxin gene in *E. coli* does not introduce any subtle heterogeneity that could cause the observed complexity of the folding mechanism, we performed unfolding and refolding kinetics on nonrecombinant apoflavodoxin directly prepared from cyanobacterial *Anabaena* flavodoxin. The kinetic behavior (not shown) was the same as that of the recombinant protein.

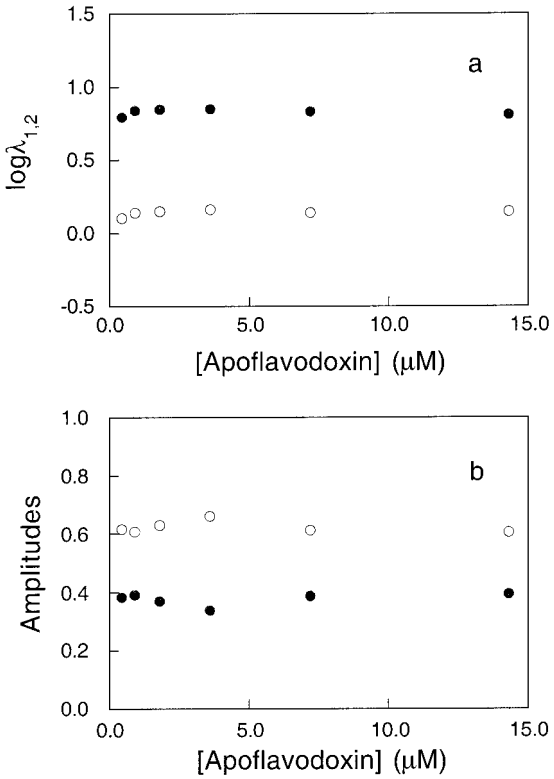

FIGURE 5: Dependence of the observed relaxations (a) and amplitudes (b) of apoflavodoxin refolding at 0.4 M urea on protein concentration. Filled circles, fast phase; open circles, slow phase.

Preliminary Analysis: Discrimination between On-Pathway and Off-Pathway Models. To analyze the observed apoflavodoxin kinetics, we first considered the simplest three-species kinetic model: a linear mechanism where the intermediate can be located either on-pathway or off-pathway (Figure 1a,b). In either linear mechanism, the experimental data (apparent rates and reduced amplitudes) can be expressed as a function of four microscopic rate constants plus the relative value of the spectroscopic property of the intermediate (see Materials and Methods). The value of these unknowns can be analytically calculated, at each urea concentration, by solving the theoretical equations of the particular mechanism. For any urea concentration in the transition region, where both folding and unfolding rate constants can be measured, six equations are available (eqs 3a, 3b, 8, 9, 10, and 11 for the *on-pathway* model, and eqs 3a, 3b, 14, 15, 16, and 17 for the *off-pathway* model; see Materials and Methods) that can be used to calculate the four microscopic constants at that urea concentration.

In practice, we calculated the individual microscopic rate constants (k_1, k_2, k_3, and k_4), at different urea concentrations, from the observed values of λ_1 and λ_2 and their associated folding and unfolding reduced amplitudes, using eqs 3a and 3b, together with every possible combination of two equations among eqs 8, 9, 10, and 11 for the on-pathway model, or among eqs 14, 15, 16, and 17 for the off-pathway model. This was done systematically while keeping the relative fluorescence of the intermediate (F_I^{rel}) at 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 ($F_I^{rel} = 0$ when $F_I = F_D$, and $F_I^{rel} = 1$ when $F_I = F_N$; see eq 12). Values of F_I^{rel} higher than 1.0 were not considered because the intermediate was expected to display lower fluorescence than the folded protein (*17, 18*). To discriminate between the on-pathway and off-pathway models, the sets of equations (plus a given value of F_I^{rel}) that yielded, at some urea concentrations, absurd (negative) values for the microscopic constants were discarded. In fact, most of the sets derived from the on-pathway equations were discarded at this stage.

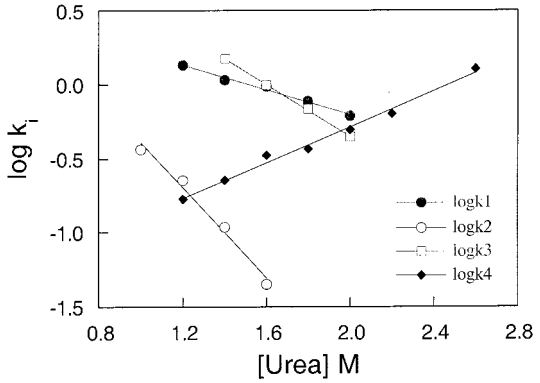

FIGURE 6: Microscopic rate constants calculated from experimental data and eqs 3a, 3b, 14, 15, 16, and 17 (off-pathway model) at different urea concentrations (with $F_N = 1.0$; $F_I = 0.0$; $F_U = 0.0$).

For the off-pathway model, the results were more promising, and several sets of equations yielded reasonable results. In Figure 6, the four microscopic constants (k_i), calculated with a particular set of equations, have been plotted against urea concentration. With the model applied in the whole urea concentration range analyzed, a linear relationship between the logarithm of the microscopic constants and urea concentration would be expected, according to eq 32 (*29*−*31*):

$$\ln k_i = \ln k_i^0 + (m_i/RT)[\text{urea}] \qquad (32)$$

where k_i^0 is the value of the microscopic constant in the absence of denaturant, m_i a proportionality constant, R the constant of the perfect gases, and T the absolute temperature. As shown in Figure 6, the off-pathway model seems to fit well the data in the transition region.

Since several sets of equations (most derived from the off-pathway model) gave us reasonable sets of microscopic constants, we performed the following consistency test. The microscopic constants (k_i) and m_i values were used to calculate the concentrations of D, I, and N at high urea concentration. The k_i and m_i values should predict that, at high urea concentration, the dominant species is the denatured state. This was the case for the fits to the off-pathway model. However, the k_i and m_i values obtained using the few sets of equations that had not produced absurd results for the on-pathway model wrongly predicted, in all cases, that at high urea concentration the dominant species was the intermediate.

Global Fitting of the Experimental Data to the Different Three-State Linear Kinetic Models. The previous partial analysis (where not all the available kinetic information was used simultaneously) indicates that the apoflavodoxin folding intermediate is not on-pathway and provides a reasonable preliminary agreement to the linear off-pathway mechanism. We have extended the analysis to any possible three-species linear model by globally fitting the whole data (relaxation constants and amplitudes at several urea concentrations) to the equations that relate them to the microscopic rate constants in the different linear models (see Materials and Methods). Only the slow unfolding phase (not very accurately

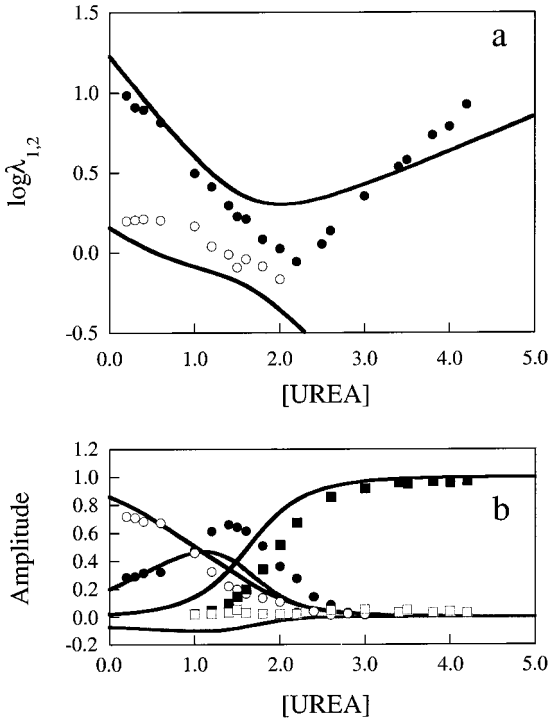

FIGURE 7: Global fit of observed relaxation constants (a: filled circles, fast phase; open circles, slow phase) and their corresponding amplitudes (b: filled circles, fast refolding phase; open circles, slow refolding phase; filled squares, fast unfolding phase; open squares, slow unfolding phase) of apoflavodoxin folding to the equations of the on-pathway mechanism.

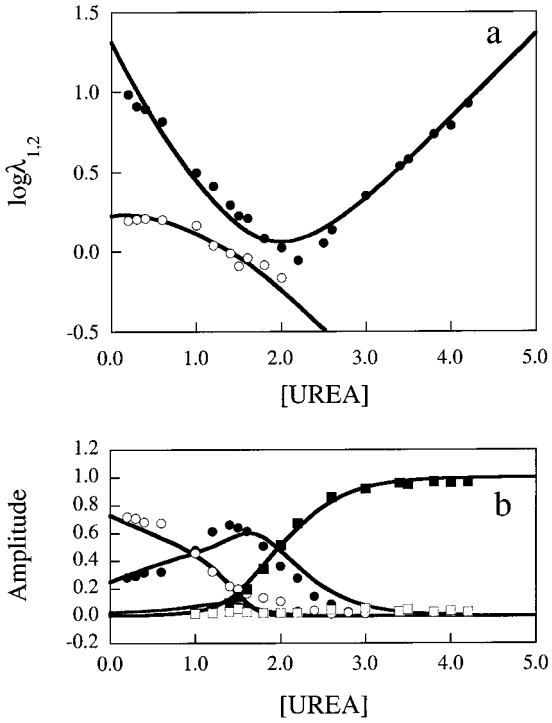

FIGURE 8: Global fit of observed relaxation constants (a: filled circles, fast phase; open circles, slow phase) and their corresponding amplitudes (b: filled circles, fast refolding phase; open circles, slow refolding phase; filled squares, fast unfolding phase; open squares, slow unfolding phase) of apoflavodoxin folding to the equations of the off-pathway mechanism.

determined because of its small amplitude) has not been included in the analysis, since its consideration precluded finding reasonable global fits (see next section for a fit to a triangular model, that includes this phase). Convergence has been tested by running several minimization processes with different, randomly generated, initial estimates. We have typically considered that convergence was good when several runs, starting from different values, achieved the same results (±5%).

The first model considered was the on-pathway mechanism (Figure 1a). Although this model had been invalidated in the preliminary analysis (see above section), we, nevertheless, attempted a global fitting of the experimental data to the pertinent equations (eqs 3a, 3b, 8, 9, 10, and 11), in the whole range of urea concentration (0−4 M). No reasonable global fit was obtained using F_I^{rel} values between 0 and 1, and consideration of the slow unfolding phase did not improve the fit (not shown). This definitively rules out that the apoflavodoxin folding intermediate is on-pathway. The best possible fit to the on-pathway model is shown in Figure 7.

The next model analyzed was the off-pathway mechanism. The best fit was obtained with $F_I^{rel} = 0$. As shown in Figure 8, the off-pathway model describes well the experimental data ($R^2 = 0.97$). The initial estimates of k_i^0 and m_i were randomly varied, and, in all cases, the minimization converged to very similar values: $k_1^0 = 5.347$ (±0.356) s^{-1}; $m_1 = -0.66$ (±0.03) kcal mol^{-1} M^{-1}; $k_2^0 = 10.429$ (±2.304) s^{-1}; $m_2 = -2.08$ (±0.12) kcal mol^{-1} M^{-1}; $k_3^0 = 6.250$ (±0.517) s^{-1}; $m_3 = -0.72$ (±0.04) kcal mol^{-1} M^{-1}; $k_4^0 = 0.052$ (±0.008) s^{-1}; $m_4 = 0.72$ (±0.03) kcal mol^{-1} M^{-1}. When a value of $F_I^{rel} = 1$ was used, the global fit obtained was worse ($R^2 = 0.95$), and the curvature of the slow apparent rate at low urea concentration was not predicted. Intermediate F_I^{rel} values did not improve the fit (not shown). We also tried to fit the data to the off-pathway model considering that the urea dependence of the relative fluorescence of the intermediate could be independent of that of the native and unfolded states. In such a general case, F_I^{rel} is described by eqs 12 and 13 (see Materials and Methods). To avoid convergence problems, derived from the increased number of parameters, a robust minimization technique, based on a Lorenztian distribution, was applied. The fit obtained was only slightly better ($R^2 = 0.98$) than when $F_I^{rel} = 0$, and the final values for k_i^0 and m_i were very similar (data not shown).

The third model evaluated was a two-channel mechanism where two different denatured states can evolve to the native state through parallel routes (Figure 1c). Although it is possible to find a reasonable global fit to the experimental data ($R^2 = 0.96$), it does not predict the curvature of the slow refolding relaxation with urea concentration (not shown).

The presence of two slowly interconverting denatured states related by cis−trans isomerization seemed very unlikely, as jugded from the similarity of the Pro55Val mutant and wild-type kinetics, and from the double-jump experiments (see above). We, nevertheless, attempted a global fit of the experimental data to a fourth kinetic model where two denatured states interconvert before evolving to the native state (Figure 1d). We assumed that the two denatured states (D_1 and D_2) displayed the same fluorescence value, but we did not make any assumption about their relative populations at high urea concentration. Global fit

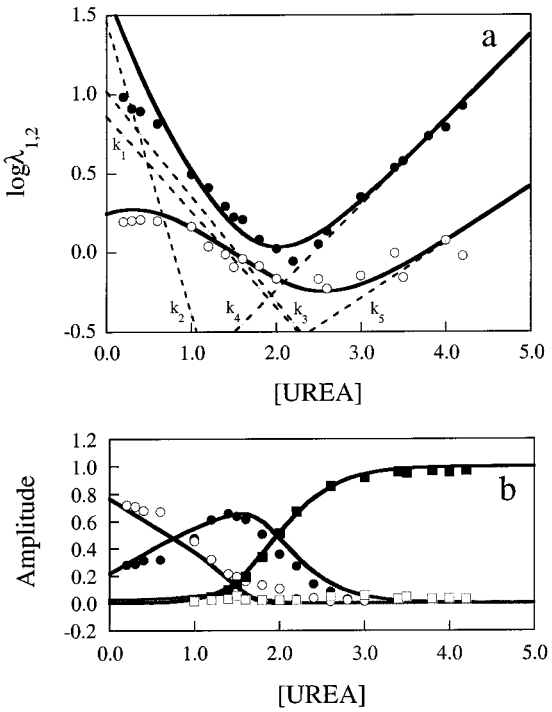

FIGURE 9: Global fit of observed relaxation constants (a: filled circles, fast phase; open circles, slow phase) and their corresponding amplitudes (b: filled circles, fast refolding phase; open circles, slow refolding phase; filled squares, fast unfolding phase; open squares, slow unfolding phase) of apoflavodoxin folding to the equations of a triangular mechanism. The calculated linear dependence of the microscopic constants on urea concentration is shown as dashed lines.

of the experimental data to this kinetic model (equations not shown) gives exactly the same k_i and m_i values as those obtained for the off-pathway model (and the same global correlation coefficient: $R^2 = 0.97$). Moreover, the values obtained for the microscopic rate constants imply that only D_2 is populated (>98%) at high denaturant conditions ([urea], 4 M), which reduces this model to the off-pathway one.

In addition, any other possible three-state linear model was tested. This includes the following: two interconverting native states (Figure 1e), two noninterconverting native states that fold through parallel routes (Figure 1f), and one intermediate state that has to fold completely before unfolding (Figure 1g). None of them led to a satisfactory fit (not shown).

Our previous analyses thus indicate that, of all possible linear three-state models, only the off-pathway one fits well the experimental data. Although a complex dependence of the intermediate fluorescence yielded the best fit, a much more simplified model, that assumes the intermediate fluorescence is like that of the denatured state, sufficed to obtain a good fit.

Global Fitting of the Experimental Data to a Triangular Kinetic Model. We note that although the off-pathway model describes most of the apoflavodoxin observed kinetic behavior (Figure 8), it does not predict the slow unfolding phase (see Figure 2). To get a good fit of this phase, a triangular mechanism (Figure 1h; see equations under Materials and Methods) has to be invoked. The best global fit to the triangular mechanism (with $F_I^{rel} = 0$) is shown in Figure 9 ($R^2 = 0.95$). Consideration of a rigorous [urea] dependence of F_I^{rel} (see Materials and Methods) neither improved the fit nor changed the results (not shown). The values obtained for the microscopic rate constants [$k_1^0 = 7.263$ ($\pm$0.989) s^{-1}; $m_1 = -0.82$ ($\pm$0.04) kcal mol^{-1} M^{-1}; $k_2^0 = 28.983$ ($\pm$9.400) s^{-1}; $m_2 = -2.51$ ($\pm$0.29) kcal mol^{-1} M^{-1}; $k_3^0 = 10.403$ ($\pm$1.186) s^{-1}; $m_3 = -0.90$ ($\pm$0.03) kcal mol^{-1} M^{-1}; $k_4^0 = 0.049$ ($\pm$0.008) s^{-1}; $m_4 = 0.73$ ($\pm$0.04) kcal mol^{-1} M^{-1}; $k_5^0 = 0.045$ ($\pm$0.023) s^{-1}; $m_5 = 0.48$ ($\pm$0.07) kcal mol^{-1} M^{-1}; $k_6^0 = 0.00085$ ($\pm$0.00041) s^{-1}; $m_6 = 0.42$ ($\pm$0.35) kcal mol^{-1} M^{-1}] make this mechanism close to the off-pathway model because the four microscopic constants shared by the two models have similar values, and the two additional constants in the triangular mechanism are very small. According to the values obtained for the six microscopic constants of the triangular mechanism, most of the unfolded molecules that fold to the intermediate conformation must unfold back before reaching the native conformation, and only a small percentage evolves directly from the intermediate to the native state. In practice, the apoflavodoxin intermediate acts as a kinetic trap that slows down the folding reaction.

For the sake of completeness, we have finally considered alternative triangular mechanisms: first, a more general triangular mechanism where, even at high denaturant concentration, two different unfolded states may be significantly populated (Figure 1i). For this model, we have generalized the amplitude equations of the triangular mechanism (not shown) to allow the concentration of the intermediate not to be zero at high urea concentration (that is, at the beginning of the folding reaction). We have then fitted all data to this generalized triangular model ($R^2 = 0.94$) and obtained the corresponding microscopic constants. These constants set the concentration of the intermediate at high urea concentration and, therefore, unbiasedly decide whether the mechanism involves a single denatured state (conventional triangular mechanism) or if two different species populate the unfolded state. The values of the constants (not shown) were almost the same as those obtained in the fit to the conventional triangular mechanism (Figure 1h), and indicate that only one denatured state is populated at 4 M urea (>97%).

Last, a triangular model with two interconverting native states populated at low denaturant concentration (Figure 1j) and displaying the same fluorescence was considered. The fit was clearly worse (not shown).

DISCUSSION

A Monomeric, Non-Proline-Related Intermediate in the Folding of Apoflavodoxin. The folding kinetics of apoflavodoxin are biphasic, indicating that an intermediate accumulates during the reaction. Intermediates in folding reactions may be of different types. In some cases, they are related to the presence of a cis proline residue in the folded protein that is mainly in the trans conformation in the unfolded protein (*32*). The only proline residue in apoflavodoxin (Pro55) is trans and thus unlikely to give rise to the observed intermediate. In addition, the apparent constant values are clearly higher than those of typical cis−trans prolyl isomerization processes [below 0.01 s^{-1} at 25 °C; (*35*)]. We have, nevertheless, mutated the proline residue to valine, and found that the mutant protein displays biphasic refolding kinetics very similar to the wild-type protein, which

rules out that the proline in wild type is related to the intermediate. Besides, the persistence of the two refolding phases in double-jump refolding experiments excludes non-proline cis peptide bonds (that might have passed unnoticed in the crystallographic modeling) as the origin of the intermediate.

Another possible type of folding intermediate are non-monomeric species. Since intermediates are, by definition, partly folded, they are likely to expose hydrophobic surfaces that could facilitate the association of monomers. Nonmonomeric intermediates have been described in the folding of a number of proteins (*41*−*43*). One interesting case is that of Che Y (a structural homologue of apoflavodoxin) in whose folding kinetics an intermediate with close-to-native fluorescence has been described (*44*). Based on the fact that a dimeric equilibrium intermediate has also been observed for this protein (*45*), the possibility that its folding intermediate could be dimeric was tested (*44*). To test if the apoflavodoxin intermediate is dimeric, we have measured the folding of apoflavodoxin at different protein concentrations from 0.5 to 14 μM. The observed rate constants and the relative amplitudes of both phases are concentration-independent, which proves the apoflavodoxin intermediate is a monomeric species. Disulfide bond formation during apoflavodoxing refolding can also be ruled out since the presence of DTT does not change the kinetics (not shown).

The Apoflavodoxin Intermediate Is Off-Pathway. The monomeric apoflavodoxin intermediate thus probably represents a compact conformation that is partially folded. The role of these 'conventional' intermediates in protein folding is intensely questioned [see (*2*) and references cited therein]. On one hand, there is an important difference between being on the folding pathway (as an obligatory species from the denatured to the native state) and being off the folding pathway (as a side reaction product that must unfold to undertake the correct folding pathway). Quite often the intermediates detected in the folding of proteins are in rapid equilibrium with the unfolded protein, and this makes it very difficult to distinguish between the on-pathway and off-pathway mechanisms (*3*, *46*). In apoflavodoxin, it is possible to distinguish between these two mechanisms because we can observe the two refolding phases and not just a burst [the relevance of burst phases has been recently questioned; (*47*)] or a curvature in the otherwise only measurable refolding phase (*48*). When the entire refolding process can be observed, the amplitudes and relaxations obtained at different urea concentrations contain enough information to calculate the four microscopic constants that define the equilibria, and to discriminate between the two mechanisms.

Although this can be directly done by fitting all the kinetic data to the equations that apply to each mechanism to see which one explains the data, we feared that the multidimensional space formed by the pertinent equations could not be properly searched in this way. Therefore, we first performed a stepwise analysis using the data at each urea concentration in the transition region to analytically solve the equations of the on-pathway and off-pathway mechanisms and get the values of the microscopic constants. Only the off-pathway mechanism gave nonabsurd values for the microscopic constants, and they displayed the expected urea concentration dependence (eq 32). The on-pathway mechanism was thus discarded.

The same conclusion was reached from a global analysis of all the available data (excluding the small, slow, unfolding phase) at all urea concentrations to eqs 3a, 3b, 8, 9, 10, and 11 (on-pathway model) or to eqs 14, 15, 16, and 17 (off-pathway model), where the microscopic constants were expressed as functions of urea concentration (as in eq 32). In the global fit, the four microscopic constants at 0 M urea and their urea dependencies were floating parameters. Figure 8 shows that the agreement between the experimental data and the global fit to an off-pathway model is good. Convergence is also good, and the values obtained for the parameters are independent of the initial estimates (not shown). In contrast, no reasonable fit was ever found to the equations describing the on-pathway model (Figure 7).

A Triangular Mechanism (Essentially Off-Pathway) Explains the Small, Slow, Unfolding Phase. Most of the apoflavodoxin kinetic data is thus explained by the off-pathway mechanism (Figure 8). The small, slow unfolding phase is, however, not captured by this model, and, in fact, the data corresponding to this phase (that accounts for around 5% of the unfolding amplitude) were excluded from the fitting to either one of the linear models (inclusion of this phase never improved the fits; not shown). The phase seems real as it is also present in the apoflavodoxin directly obtained from *Anabæna* cells, and in several apoflavodoxin variants (for example, in the Pro55Val mutant). Moreover, the phase does not disappear in the presence of DTT, and it is protein concentration independent. It seems, thus, that this phase should also be explained.

The phase can be easily explained if a triangular mechanism is assumed. Figure 9 compares all kinetic data with the best fit to the equations of a triangular mechanism. The fit captures the slow phase. Interestingly, the values of the four microscopic constants that are common to the off-pathway and triangular models are quite similar in the two mechanisms. The microscopic constants calculated for the triangular mechanism indicate that most apoflavodoxin molecules fold from the denatured state. This includes the molecules that transiently populate the folding intermediate, that in most cases become denatured before refolding to the native state. The folding mechanism of apoflavodoxin is thus triangular but essentially off-pathway.

Some rollover seems to appear in the folding arm of the fast chevron. Although this could reflect early folding events (not changing the initial fluorescence intensity of the denatured state), it could also be simply due to a lower accuracy of the stopped-flow data, in this region of higher rate constants, for the fast refolding phase whose relative amplitude at low urea concentration is small.

The global fit of the kinetic data to the triangular mechanism is consistent with the equilibrium urea denaturation data (*17*). Using the values of k_i and m_i obtained in the global fit, the relative concentration of each of the three species can be calculated as a function of urea concentration and, from that, the equilibrium unfolding curve can be predicted using eq 33:

$$F = F_N[N] + F_I[I] + F_D[D] \qquad (33)$$

where F is the predicted fluorescence at each urea concentration and F_N, F_I, and F_D are the normalized fluorescence values of the native, intermediate, and unfolded state,

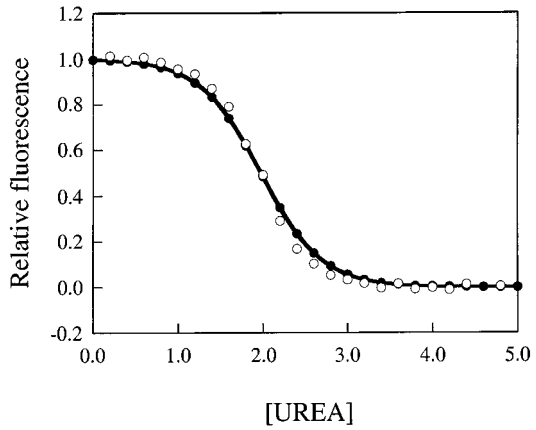

FIGURE 10: Theoretical equilibrium urea unfolding curve calculated from the microscopic kinetic rate constants obtained after global fitting to a triangular mechanism (line, filled circles) superimposed to the experimental one [open circles (*17*)].

respectively. We show in Figure 10 that the theoretical curve is, for all practical purposes, a two-state curve that essentially coincides with the experimental one (*17*). From the fitted values of the microscopic rate constants, we calculate that, in our standard buffer (50 mM MOPS, pH 7, 25 °C), the highest intermediate population (achieved at 0 M urea) is only 1.6%. The equilibrium is thus, in practice, two-state. The ΔG and m values calculated for the two-state N−D equilibrium (3.2 kcal mol^{-1} and 1.63 kcal mol^{-1} M^{-1}) are not far from the experimentally determined values (4.1 and 2.09, respectively). Importantly, the urea concentration of middenaturation (that can be more accurately determined from the equilibrium unfolding curves than either ΔG or m) is 1.96 (±0.03) M (*17*, *49*), and can be predicted as 1.94 (±0.01) kcal mol^{-1} from the fitted values of the microscopic rate constants. The overall good agreement of the experimental two-state equilibrium unfolding curve and that calculated from the fit can be appreciated in Figure 10.

On the Nature of the Intermediate State. The apoflavodoxin intermediate, essentially off-pathway, is one of the first documented off-pathway folding intermediates and, therefore, may provide interesting insight into the nature of these dead-end conformations. Indeed, some of its properties seem unusual when compared with those of other folding intermediates described as on-pathway.

The m-values associated with the different kinetic transitions can provide a qualitative indication of the average compactness of partially or totally unfolded states (*2*, *29*). The relative compactness of the intermediate state (α_I) with respect to the unfolded state (considering $\alpha_D = 0$ and $\alpha_N = 1$) can be derived from the m-values obtained after fitting experimental data to the off-pathway or triangular models: $\alpha_I = (m_2 - m_1)/(m_3 - m_4)$. From our analysis, the intermediate state seems to be as compact as the native state (off-pathway model: $\alpha_I = 0.99 \pm 0.04$; triangular model: $\alpha_I = 1.04 \pm 0.18$), which emphasizes its off-pathway character as a collapsed kinetic trap. A similar off-pathway over-compact state has been recently described in the ribosomal protein S6 (*50*), so it seems that compactness could be an intrinsic characteristic of kinetically trapped off-pathway intermediates.

More intriguing is the negative m-value for the unfolding reaction of the intermediate, obtained from either the off-pathway model ($m_1 = -0.66 \pm 0.03$ kcal mol^{-1}) or the triangular model ($m_1 = -0.82 \pm 0.04$ kcal mol^{-1}). The associated errors are small, so we should rely on the actual figures that suggest the transition state between I and D (TS_{ID}) is more compact than the intermediate (itself as compact as the native state). In this respect, we notice that, although the wild-type apoflavodoxin X-ray structure (*15*) is that of an ordinarily compact protein (the crevice left by the removed FMN cofactor is filled by an aromatic side chain: Trp57), the structure in solution seems to be substantially more open in the FMN binding region (*51*). This means that apoflavodoxin conformations more compact than the folded state in solution are indeed possible. The reason the ID transition state is more compact than the intermediate is, however, not obvious. At this moment, we can only speculate that compaction of a region of this intermediate is frustrated by the presence of other stabilizing interactions elsewhere in the intermediate. A debilitation of these interactions would then lead the intermediate into a more compact ID transition state.

The apoflavodoxin intermediate also departs from the behavior reported for on-pathway intermediates (see ref *12* for a recent example) in that, according to the positive value of *m5* (associated with the I to N transition), it is more compact than the IN transition state. An unusual behavior in the IN transition has also been found for the ribosomal protein S6 (*50*). In that case, the kinetic m-value associated with the IN conversion was very close to 0, suggesting "glassy" downhill folding. In the apoflavodoxin intermediate, however, the positive m-value for the IN transition suggests that the over-compact collapsed intermediate must initially expand in order to achieve the native contacts.

Last, despite its compactness, our global analysis is consistent with an intermediate fluorescence intensity similar to that of the denatured state. In this respect, it should be noted that, although the effect of polypeptide compaction on its wavelength of maximal emission is rather predictable (a blue-shift), the effect on the intensity depends heavily on packing details. Incidentally, the apoflavodoxin equilibrium intermediates characterized so far (*17*−*18*) display, despite their well-characterized compactness, fluorescence intensities close to that of the unfolded state.

It seems thus that, in the apoflavodoxin folding reaction, some molecules follow a simple two-state mechanism while others get trapped in a compact intermediate conformation, probably with some non-native interactions. This intermediate can then fold slowly, through a less structured transition state, or, alternatively, it can unfold through a peculiarly compact transition state, and then try again. Further experimental characterization of these compact intermediate and transition states can contribute to analyze the role of collapsed species on protein folding.

Does Intermediate Accumulation Significantly Retard Apoflavodoxin Folding? One pertinent question related to the role of intermediates in protein folding is whether they speed up or slow down protein folding. For on-pathway intermediates, the question is under debate, but off-pathway intermediates must always slow the reaction. Our system provides an opportunity to calculate the influence of the intermediate on the speed of the reaction. We have simulated the apoflavodoxin unfolding curve that would be observed if the off-pathway intermediate did not occur at all (other

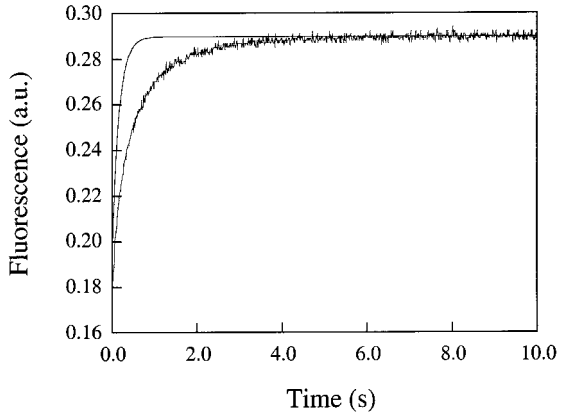

FIGURE 11: Simulation (smooth line) of the kinetic trace of refolding of apoflavodoxin in 0.4 M urea if it followed a two-state model with the same microscopic rate constants (k_{UN} and k_{NU}) that have been obtained after global fitting of the experimental data to the triangular model ($k^0_{UN} = 10.403\ s^{-1}$; $m_{UN} = -0.90$ kcal mol^{-1} M^{-1}; $k^0_{NU} = 0.049\ s^{-1}$; $m_{NU} = 0.73$ kcal mol^{-1} M^{-1}). The experimental kinetic trace is also shown for comparison.

things being equal), and it is compared in Figure 11 to the observed kinetic trace. As the figure shows, the difference is not large, and it is doubtful that, for this particular protein, it would be advantageous from a physiological point of view (avoidance of potential aggregation problems, speed of holoflavodoxin formation) to evolve toward intermediate destabilization.

The view that emerges on the folding pathway of apoflavodoxin, based on the results presented here, is that the denatured protein may give rise to two alternative compact conformations of different stability (the intermediate and the native state) that are in equilibrium both directly and through the denatured state. Although protein molecules in the intermediate conformation may directly become native, the energetics of the equilibria determine that most of them unfold before becoming native. The intermediate thus slows down the folding, but not much. According to this view, the intermediate is useless but probably harmless. Interestingly, despite the fact that apoflavodoxin (169 aa) is twice as big as some small model proteins, there is a direct path from the denatured to the native state, which means that proteins of this size can, in principle, fold in a two-state fashion.

ACKNOWLEDGMENT

We are grateful to R. Abagyan and M. Totrov for their assistance in calculations with ICM.

REFERENCES

1. Jackson, S. E. (1998) *Folding Des. 3*, R81−R91.
2. Roder, H., and Colon, W. (1997) *Curr. Opin. Struct. Biol. 7*, 15−28.
3. Khorasanizadeh, S., Peters, I. D., Butt, T. R., and Roder, H. (1996) *Nat. Struct. Biol. 3*, 193−205.
4. Tan, Y. J., Oliveberg, M., and Fersht, A. R. (1996) *J. Mol. Biol. 264*, 377−389.
5. Hosszu, L. L. P., Craven, C. J., Parker, M. J., Lorch, M., Spencer, J., Clarke, A. R., and Waltho, J. P. (1997) *Nat. Struct. Biol. 4*, 801−804.
6. Shastry, M. C. R., and Roder, H. (1998) *Nat. Struct. Biol. 5*, 385−392.
7. Farooq, A. (1998) *Biochemistry 37*, 15170−15176.
8. Gloss, L. M., and Matthews, C. R. (1998) *Biochemistry 37*, 15990−15999.
9. Tsui, V., Garcia, C., Cavagnero, S., Siuzdak, G., Dyson, H. J., and Wright, P. E. (1999) *Protein Sci. 8*, 45−49.
10. Capaldi, A. P., Ferguson, S. J., and Radford, S. E. (1999) *J. Mol. Biol. 286*, 1621−1632.
11. Bai, Y. (2000) *Protein Sci. 9*, 194−196.
12. Capaldi, A. P., Shastry, M. C. R., Kleanthous, C., Roder, H., and Radford, S. E. (2001) *Nat. Struct. Biol. 8*, 68−72.
13. Heidary, D. K., Gross, L. A., Roy, M., and Jennings, P. A. (1997) *Nat. Struct. Biol. 4*, 725−731.
14. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) *Science 267*, 1619−1620.
15. Genzor, C. G., Perales-Alcón, A., Sancho, J., and Romero, A. (1996) *Nat. Struct. Biol. 3*, 329−332.
16. Fillat, M. F., Borrias, W. E., and Weisbeerk, P. J. (1991) *Biochem. J. 280*, 187−191.
17. Genzor, C. G., Beldarraín, A., Gómez-Moreno, C., López-Lacomba, J. L., Cortijo, M., and Sancho, J. (1996) *Protein Sci. 5*, 1376−1388.
18. Irún, M. P., Garcia-Mira, M. M., Sánchez-Ruiz, J. M., and Sancho, J. (2001) *J. Mol. Biol. 306*, 877−888.
19. Irún, M. P., Maldonado, S., and Sancho, J. (2001) *Protein Eng. 14*, 173−181.
20. Maldonado, S., Jiménez, M. A., Langdon, G. M., and Sancho, J. (1998) *Biochemistry 37*, 10589−10596.
21. Sanger, F., Niklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5465.
22. Nelder, J. A., and Mead, R. (1965) *Comput. J. 7*, 308−313.
23. Abagyan, R., Totrov, M., and Kuznetsov, D. (1994) *J. Comput. Chem. 15*, 488−506.
24. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992) in *C. The art of scientific computing*, 2nd ed., Cambridge University Press, New York.
25. Ikai, A., and Tanford, C. (1973) *J. Mol. Biol. 73*, 145−163.
26. Hagerman, P. J., and Baldwin, R. L. (1976) *Biochemistry 7*, 1462−1473.
27. Benson, S. W. (1960) in *The Foundations of Chemical Kinetics*, McGraw-Hill, New York.
28. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
29. Tanford, C. (1970) *Adv. Protein Chem. 24*, 1−95.
30. Chen, B., Baase, W. A., and Schellman, J. A. (1989) *Biochemistry 28*, 691−699.
31. Schellman, J. A. (1987) *Biopolymers 26*, 549−559.
32. Schmid, F. X. (1992) in *Protein Folding* (Creigthon, T. E., Ed.) W. H. Freeman, New York.
33. Odefey, C., Mayr, L. M., and Schmid, F. X. (1995) *J. Mol. Biol. 245*, 69−78.
34. Jabs, A., Weiss, M. S., and Hilgenfeld, R. (1998) *J. Mol. Biol. 286*, 291−304.
35. Kiefhaber, T., and Schmid, F. X. (1992) *J. Mol. Biol. 224*, 231−240.
36. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892−896.
37. Balbach, J., Forge, V., van Nuland, N. A. J., Winder, S. L., Hore, P. J., and Dobson, C. M. (1995) *Nat. Struct. Biol. 2*, 866−870.
38. Raschke, T. M., and Marqusee, S. (1997) *Nat. Struct. Biol. 4*, 298−304.
39. Brems, D. N., Plaisted, S. M., Havel, H. A., and Tomich, C. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 27*, 4541−4546.
40. Goldberg, M. E., Rudolph, R., and Jaenicke, R. (1991) *Biochemistry 30*, 49−56.
41. Silow, M., and Oliveberg, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 6084−6086.
42. Segel, D. J., Eliezer, D., Uversky, V., Fink, A. L., Hodgson, K. O., and Doniach, S. (1999) *Biochemistry 38*, 15352−15359.
43. Doyle, S. M., Braswell, E. H., and Teschke, C. M. (2000) *Biochemistry 39*, 11667−11676.
44. Muñoz, V., López, E. M., Jager, M., and Serrano, L. (1994) *Biochemistry 33*, 5858−5866.
45. Filimonov, V. V., Prieto, J., Martínez, J. C., Bruix, M., Mateo, P. L., and Serrano, L. (1993) *Biochemistry 32*, 12906−12921.

46. Baldwin, R. L. (1996) *Folding Des. 1*, R1−R8.
47. Qi, P. X., Sosnick, T. R., and Englander, S. W. (1998) *Nat. Struct. Biol. 5*, 882−884.
48. Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M., and Fersht, A. R. (1990) *Nature 346*, 440−445.
49. Fernández-Recio, J., Romero, A., and Sancho, J. (1999) *J. Mol. Biol. 290*, 319−330.
50. Otzen, D. E., and Oliveberg, M. (1999) *Proc. Natl. Acad. Sci. U.S.A. 21*, 11746−11751.
51. Langdon, G. M., Jimenez, M. A., Genzor, C. G., Maldonado, S., Sancho, J., and Rico, M. (2001) *Proteins: Struct., Funct., Genet. 43*, 476−488.

BI010216T