Unfolding Mechanism of Rubredoxin from *Pyrococcus furiosus*†

Silvia Cavagnero,‡,§ Zhi H. Zhou,|| Michael W. W. Adams,|| and Sunney I. Chan*,‡

Arthur Amos Noyes Laboratories of Chemical Physics, California Institute of Technology, Pasadena, California 91125, and Department of Biochemistry and Center for Metalloenzyme Studies, University of Georgia, Athens, Georgia 30602

Received September 3, 1997

ABSTRACT: As part of our studies on the structural and dynamic properties of hyperthermostable proteins, we have investigated the unfolding pathways of the small iron−sulfur protein rubredoxin from *Pyrococcus furiosus* (RdPf) at pH 2. Unfolding has been initiated by temperature jump, triggered by manual mixing of a concentrated protein solution into a thermally preequilibrated buffer. The process has been followed in real time by absorption, tryptophan fluorescence emission, and far-UV circular dichroism. Unlike the case of the mesophilic rubredoxin from *Clostridium pasteurianum* (RdCp), RdPf displays a complex unfolding kinetics, pointing to the formation of at least three intermediates. All of the steps, including the one involving metal ion release, are extremely slow. However, hydrophobic core relaxation—not Fe^{3+} loss—is rate-determining for RdPf unfolding. This clearly rules out the fact that Fe^{3+} is solely responsible for the kinetic stability of RdPf. Results have been discussed in terms of sequential vs parallel pathways, and the possible role of irreversible phenomena has been taken into consideration. Aggregation does not appear to play a significant role in the observed kinetic complexities. According to a proposed sequential mechanism, partial release of secondary structure elements precedes iron loss, which is then followed by further loss of β-sheet content and, finally, by hydrophobic relaxation. Although the main features of the RdPf unfolding mechanism remain substantially unchanged over the experimentally accessible temperature range, final hydrophobic relaxation gets faster, relative to the other events, as the temperature is decreased. A qualitative assessment of the unfolding activation parameters suggests that this arises from the very low activation energies (E_a) that characterize this step.

Understanding protein-folding mechanisms (*1*−*3*) in depth is a formidably challenging task. It is a field of considerable activity since the original formulation of the problem (*4*). Despite the intellectual and experimental endeavors, many questions about several of the relevant issues need to be addressed. For instance, the folding pathways of proteins from hyperthermophilic microorganisms represent a particularly interesting aspect of the folding puzzle (*5*−*7*). Hyperthermophiles are a subclass of primitive microorganisms collectively known as extremophiles. Most of these belong to the *Archaea* (*8*) domain, and they have the particular ability to thrive under conditions that are not compatible with most other life forms. In particular, they can withstand high temperatures, pressures, and salt concentrations. Hyperthermophiles typically live in the proximity of volcanoes, hot springs, and marine hydrothermal vents. They grow optimally at temperatures ranging from 80 to 103 °C. Since most of these proteins have been isolated and characterized only within the past few years (*9*−*12*), thanks to the pioneering work of Karl O. Stetter, who first revealed the existence of hyperthermophilic species, there have been few reports of folding/unfolding studies on these unusual proteins. Specifically, the forces responsible for the exceptional thermal stability of proteins from hyperthermophiles are hard to pinpoint, and they have just begun to be elucidated (*13*−*16*). In particular, the characteristically slow unfolding kinetics of these proteins (*17*) render kinetic studies very attractive experimentally. Within this context, there are several mechanistic issues that are poised for systematic investigation. Among those are the specific folding pathways of hyperthermophilic proteins and their relationships to the pathways of proteins from mesophilic species. The present work represents a first step toward addressing these questions.

Rubredoxin from *Pyrococcus furiosus* (RdPf)[1] has been well characterized both biochemically (*18*) and structurally (*14*, *19*). It is, therefore, a very good model system for the study of hyperthermostable proteins. Previous studies have shown that electrostatic interactions slow the unfolding kinetic step involving iron loss (*17*). Nonetheless, several mechanistic aspects relative to the unfolding of this protein still need to be elucidated. For instance, it is not clear whether there is any hierarchy between secondary/tertiary

† This work was supported by NIH Grant Nos. GM 22432 (S.I.C.) and GM 50736 (M.W.W.A.) from the National Institute of General Medical Sciences, U.S. Public Health Service.

* To whom correspondence should be addressed.

‡ California Institute of Technology.

§ Present Address: Department of Molecular Biology MB-2, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037.

|| University of Georgia.

[1] Abbreviations used: RdPf, rubredoxin from *Pyrococcus furiosus*; RdCp, rubredoxin from *Clostridium pasteurianum*; OD, optical density; CD, circular dichroism; N, native state; I, intermediate state; U, unfolded state; *T*, absolute temperature; *k*, kinetic constant; *A*, Arrhenius preexponential factor; E_a, Arrhenius activation energy; *R*, universal gas constant; κ, transmission coefficient; k_B, Boltzmann constant; *h*, Planck's constant; $\Delta H^{\ddagger}$, activation enthalpy; $\Delta S^{\ddagger}$, activation entropy; $\Delta G^{\ddagger}$, activation free energy.

S0006-2960(97)02180-6 CCC: $15.00
Published on Web 02/18/1998

structure relaxation and iron loss. Likewise, it has not been established whether electrostatic forces play any role in the kinetics of unfolding steps other than the one involving metal release. The combined use of different spectroscopic techniques provides a particularly powerful tool to address the aforementioned issues. We have investigated the denaturation of RdPf in real time by steady-state absorption, circular dichroism (CD), and fluorescence emission. Each method has been selected for its ability to probe different molecular events such as changes in iron coordination environment (absorption spectroscopy), variations in overall secondary structure (far-UV circular dichroism spectroscopy), and exposure of the tryptophan (Trp)-sensitive portion of the hydrophobic core to the solvent (fluorescence emission spectroscopy). Unfolding turns out to be quite complex. Due to unusually slow observed rates, manual mixing was sufficient to ensure that no significant kinetic processes were missed during the dead time of the experiments. By studying unfolding as a function of temperature, we have also determined the specific contribution of Arrhenius preexponential factors and activation energies to the rates of the individual steps. All the data presented below have been derived at pH 2.

MATERIALS AND METHODS

Sample Preparation. RdPf was isolated and purified according to known procedures (*18*). Experimental details for the absorption measurements have been described elsewhere (*17*). Protein concentrations ranged from 1 to 20 μM. These values are representative of the *in vivo* RdPf concentration (*18*). Buffers consisted of 40 mM glycine solutions at pH 2.0. No additional salts were added to the buffers (ionic strength 40 mM). Unfolding was irreversible, under our experimental conditions, as assessed by gradually lowering the sample temperature right after each thermal unfolding experiment. This behavior did not appear to depend on the specific cooling rates (ranging from approximately -1 to -10 °C/min) that were used to lower the sample temperature. No denaturing agents were employed. The kinetics of unfolding was monitored by UV−vis absorption, Trp fluorescence emission, and far-UV circular dichroism. All the spectra were baseline corrected with the software Lab Calc (Galactic Industries Corporation).

Fluorescence Spectroscopy. These experiments were performed with an SLM 4800 fluorometer, in a thermostated cuvette, under vigorous stirring. Samples were excited at 280 nm, and the time course of fluorescence emission was monitored at 340 nm. The excitation and emission bandwidths were 4 and 8 nm, respectively. Unfolding was initiated as previously described for the absorption spectroscopy experiments (*17*). The dead time of these kinetic experiments ranges between 1 and 2 s. The temperature was controlled via a thermostated water bath connected to the compartment surrounding the cuvette housing. Temperature was monitored using a thermocouple probe connected to a Fluke 80TK module plugged into a multimeter. Readings were accurate to ±0.5 °C.

Circular Dichroism Spectroscopy. Far-UV circular dichroism (CD) experiments were performed on a JASCO-600 instrument, in a water-jacketed thermostated 0.1 cm path cuvette. The time course of the spectroscopic signal was monitored at 225 nm. Unfolding was initiated as described for the absorption spectroscopy experiments. In this case, though, the addition of protein stock solution to the thermally preequilibrated buffer was rapidly followed by manual mixing. The dead time for the CD experiments was determined for each individual measurement, and it was generally found to be ~60 s. Temperature control for these experiments was achieved by connecting the cuvette directly to an external thermostated water bath. Temperatures were monitored in the water bath with a mercury thermometer (accuracy ±0.5 °C). No corrections for heat losses were necessary since temperatures in the water bath and the jacketed cell differed by only negligible amounts.

Kinetic Data Analysis. Kinetic parameters were derived by fitting the experimental kinetic traces to an expression for the observed spectroscopic signal of the form

$$f_{obs} = f_N x_N + f_I x_I + f_U x_U \quad (1)$$

where f_{OBS} represents the observed spectroscopic signal, f_N is the spectroscopic signal of the native state (N) at time zero, f_I is the hypothetical spectroscopic signal to be expected in a case where the whole protein is present as the intermediate (I), and f_U is the expected spectroscopic signal at the end of the reaction, when all the protein is in its unfolded state (U). Equation 1 reflects the fact that signals detected by optical spectroscopy are a linear combination of the signals due to each species that absorbs at the wavelength, or wavelength range, under examination. The symbols x_N, x_I, and x_U represent the mole fractions of N, I and U, respectively. These mole fractions are defined as follows

$$x_N = \frac{[N]}{[N_o]}, \quad x_I = \frac{[I]}{[N_o]}, \quad x_U = \frac{[U]}{[N_o]} \quad (2)$$

where $[N_o]$ is the initial concentration of the native state, [N] is the concentration of the native state at time *t*, [I] is the concentration of the intermediate at time *t*, and [U] is the concentration of denatured or unfolded state at time *t*.

Two different reaction schemes were considered when fitting the data. Kinetic analysis of the first scheme

$$N \rightarrow U \quad (3)$$

yielded an expression for the time course of the N and U concentrations. These were inserted into eq 1, leading to the following relation

$$f_{obs} = f_N e^{-kt} + f_U(1 - e^{-kt}) + a \quad (4)$$

where *t* is the time and *a* is a generic constant. The unfolding rate constant *k* and the constant *a*, which has been introduced to account for small baseline errors, have been treated as adjustable parameters. As mentioned, eq 1 has been used in the derivation of eq 4. The second term of eq 1 has been neglected, since the pertinent kinetic model does not predict the formation of any intermediate. In the specific case of the kinetic traces from absorption spectroscopy in the visible region, the optical density of the unfolded form is zero. Therefore, eq 4 reduces to

$$f_{obs} = f_N e^{-kt} + a \quad (5)$$

The second sequential reaction scheme is

$$N \rightarrow I \rightarrow U \quad (6)$$

Conventional kinetic analysis of eq 6 yielded an expression for the time course of the N, I, and U concentrations. No approximations were made. Substitution of these expressions into eq 1 gives

$$f_{obs} = f_N e^{-k_1 t} + f_I \frac{k_1}{(k_2 - k_1)}[e^{-k_1 t} - e^{-k_2 t}] + f_U\left\{1 + \frac{1}{(k_1 - k_2)}[k_2 e^{-k_1 t} - k_1 e^{-k_2 t}]\right\} \quad (7)$$

where k_1 and k_2 are the kinetic constants for the first and second step, respectively, of eq 6. The constants f_N and f_U were derived from the initial and final conditions, respectively, whereas f_I was regarded as an adjustable parameter in the fit. f_I should be, to a first approximation, temperature independent. In cases where f_I was observed to vary with temperature by more than about 20%, the value obtained from lowest temperature data in the series was substituted into eq 7. This modified equation, which no longer contains f_I as an adjustable parameter, was used for the fits at the higher temperatures. The rationale for selecting the lowest temperature f_I as the "good" value is that, in a variable-temperature series, it was found that the two kinetic phases are more clearly discriminated at the lowest temperature. As temperature increases, often the first phase becomes too fast to be observed. Therefore, data at lower temperature are expected to more accurately fit eq 7.

Equations 4, 5, and 7 were used to fit the kinetic traces. Typically, curve fitting was done by first testing the experimental traces against eq 4 or 5. Whenever a satisfactory fit was obtained, the calculated first-order rate constant k was regarded as the true value. Otherwise, fitting to eq 7 was performed. In the latter case, the process was considered to take place according to eq 6, i.e., via formation of an intermediate. The software package KALEIDAGRAPH (Synergy Software) was used to implement all the curve-fitting routines.

Activation Parameters. The temperature dependence of unfolding rate constants has yielded activation parameters for each observed kinetic phase. The Arrhenius equation (*20*)

$$k = Ae^{-(E_a/RT)} \quad (8)$$

was used to obtain preexponential factors (A) and activation energies (E_a). T, k, and R represent the absolute temperature, the observed kinetic rate constant, and the universal gas constant, respectively. Linear regression analysis was applied to ln(k) vs 1/T plots. The slopes and y-axis intercepts of these plots yielded E_a and A, according to eq 8. $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ were obtained from the slope and y-axis intercept of ln(k/T) vs 1/T plots. Linear fits were analyzed according to the Eyring equation (*21*)

$$k = \frac{\kappa k_B T}{h} e^{(+\Delta S^{\ddagger}/R)} e^{(-\Delta H^{\ddagger}/RT)} \quad (9)$$

where κ is the transmission coefficient, k_B is the Boltzmann constant, and h is the Planck constant. Finally, activation free energies ($\Delta G^{\ddagger}$) were obtained by substituting $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ in the expression

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger} \quad (10)$$

RESULTS

All of the unfolding studies described here have been performed at pH 2. Under these conditions, the electrostatic component to the free energy of the protein is more unfavorable than at neutral pH (*17*). Furthermore, the salt bridges of RdPf are no longer present. However, RdPf still preserves pseudonative structural features at room temperature (*13*). Regardless of the spectroscopic probe used, RdPf unfolding rates are extremely slow in the temperature range that has been analyzed.

Absorption Spectroscopy. The temperature dependence of the charge-transfer absorption band kinetics of RdPf at pH 2 has been described previously (*17*). All the curves follow monoexponential kinetics. This is consistent with Fe^{3+} release occurring in a single step upon unfolding.

Fluorescence Spectroscopy. RdPf fluorescence emission kinetic traces displayed in Figure 1a clearly indicate the formation of an unfolding intermediate on a slow time scale. This is unusual since unfolding of most proteins studied so far follows monophasic kinetics (*22*). The kinetic traces fit well to a biexponential expression, derived from the expected observed fluorescence trace for the simple two-step mechanism N → I → U, where N is the native state, I is a generic intermediate, and U is the unfolded state. Figure 1a displays the curve fits to the fluorescence emission traces, superimposed on the original data. The fits at 72−78 °C are quite satisfactory. Slightly poorer results have been obtained for the traces recorded at 63.3 and 68.3 °C. These fits apparently overestimate the rate of the first kinetic phase, and they underestimate the rate of the second kinetic phase. This suggests that the backrate for the first step of mechanism 6 should not be neglected at lower temperatures. In any case, we regard the fits based on eq 7 as reasonably satisfactory. Figure 2 illustrates the Eyring plot for the first and second RdPf fluorescence kinetic phases. The fit for the first step (k_1) is very good, whereas the fit for the second phase (k_2) is affected by a larger error. This may be explained in terms of our inability to detect some mechanistic subtleties or, as discussed in the next section, in terms of a temperature-dependent heat capacity of activation.

To provide a direct comparison between the kinetic unfolding behavior of hyperthermophilic and mesophilic species, the unfolding of the mesophilic rubredoxin from *Clostridium pasteurianum* (RdCp) has also been followed, at pH 2.0 and 45.0 °C (Figure 1b). RdCp unfolding is clearly monoexponential, with no observable formation of kinetic intermediates. This behavior is not surprising, being in line with that of most other mesophilic proteins. The intensity of the RdCp fluorescence emission trace increases with time, leveling off after the protein has unfolded. This indicates that the increase in quantum yield caused by Fe^{3+} release dominates over quenching due to solvent exposure, as the protein unfolds.

In principle, it is conceivable that some of the observed kinetic complexities displayed by RdPf may be due to aggregation phenomena, rather than to intrinsic intramolecular processes. To analyze this possibility, we have studied

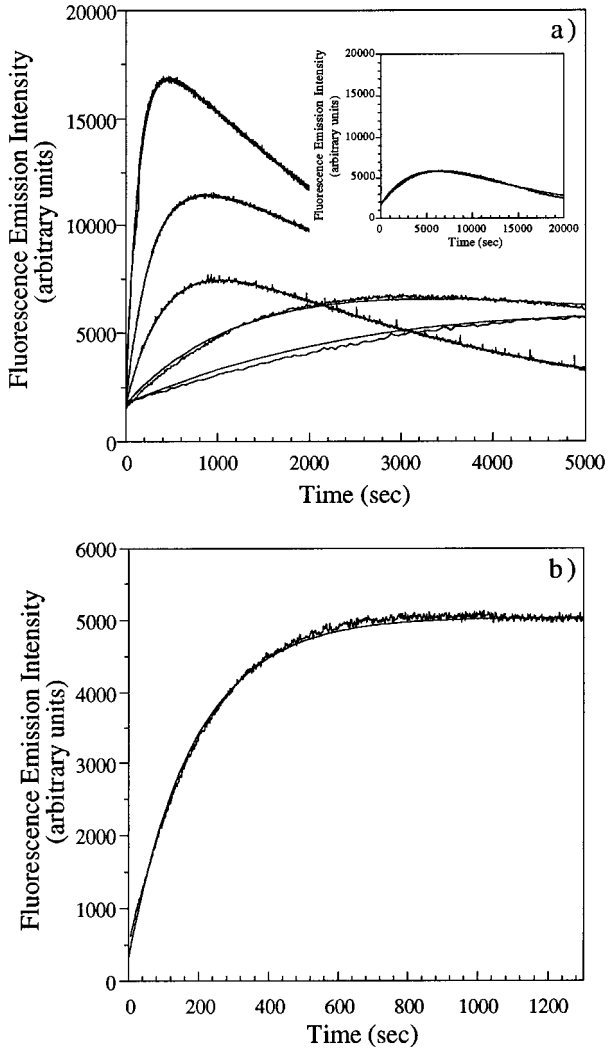


FIGURE 1: RdPf (a) and RdCp (b) unfolding kinetics as followed by Trp fluorescence emission spectroscopy. The experimental pH is 2.0. The excitation and emission wavelengths are 280 and 340 nm, respectively. Experimental temperatures for RdPf unfolding are as follows, going from the lower to the upper trace: 63.3, 68.3, 72.3, 74.6, and 77.8 °C. Nonlinear regression curve fitting for RdPf was performed according to a three-state model, and data were fit to eq 7. The insert shows the complete RdPf unfolding time course at 63.3 °C. RdCp unfolding was followed at 45.0 °C. Curve fitting for the RdCp trace was based on a two-state model, as described by eq 4.

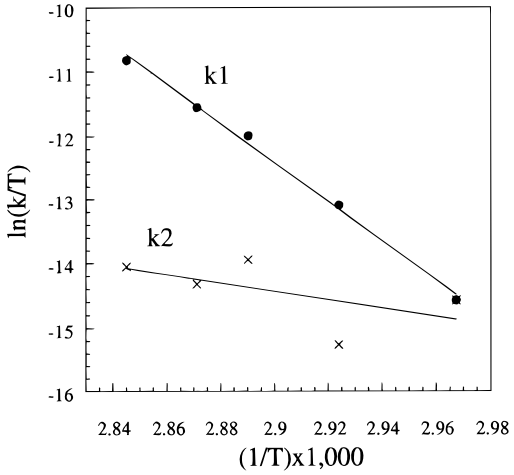


FIGURE 2: Eyring plot for the two RdPf unfolding kinetic phases (k_1, faster kinetic phase; k_2, slower kinetic phase) as detected by Trp fluorescence emission, at pH 2.0. Linear least-squares curve fitting was performed on both traces.

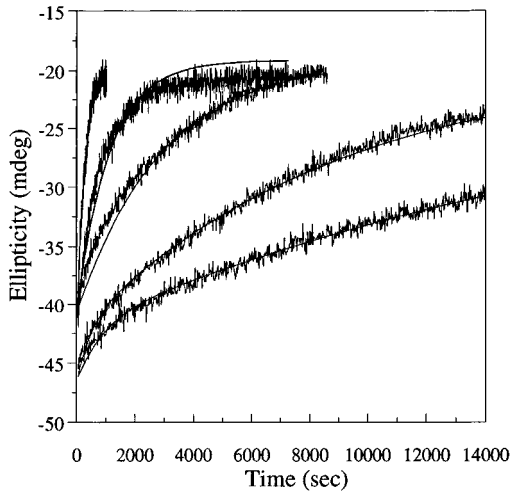


FIGURE 3: Kinetics of RdPf thermal unfolding as followed by circular dichroism spectroscopy. The experimental pH is 2.0. The experimental temperatures are as follows: (a) 63.6 °C, (b) 66.3 °C, (c) 70.1 °C, (d) 75.7 °C, and (e) 78.5 °C. Nonlinear least-squares curve fitting was performed. A three-state kinetic model was used for curve fitting. Data were fit to eq 7.

the concentration dependence of the RdPf unfolding kinetics over the 1−100 μM concentration range (data not shown). Apparent rate constants do not appreciably vary between 1 and 18 μM concentration, suggesting that one of the following two situations applies: (a) either aggregation does not play a role in the concentration range relevant to the experiments performed in this work or (b) aggregation is fast compared to unfolding, and it occurs only after the unfolded state has formed. Case b is kinetically undetectable, and it does not interfere to any extent with the ability to follow intrinsic unfolding events. However, as RdPf concentration approaches 100 μM, significantly slower unfolding rates have been observed. This phenomenon, whose origin is not entirely clear at present, is not immediately suggestive of aggregation, since aggregation is generally expected to give rise to faster apparent unfolding rate constants as protein concentration is increased. One viable possibility is that, at concentrations close to 100 μM, intermolecular association of either the native or one of the RdPf intermediate states promotes formation of an unfolding-competent more stable species. This is expected to denature more slowly than its nonaggregated counterpart, giving rise to slower observed kinetics at increasing initial protein concentrations. In any case, this behavior is not directly relevant to the present work, since we have employed RdPf concentrations considerably lower than 100 μM to follow unfolding.

Circular Dichroism Spectroscopy. Experimental far-UV circular dichroism (CD) spectra are shown in Figure 3. Data are best fit by curves derived from the biexponential expression eq 7. Excellent fits are obtained at all experimental temperatures. Over the course of data fitting, we observed that the values obtained for the first kinetic constant (k_1) tend to become unreasonably high as higher temperatures are approached. This can be explained since, according to the kinetic trend established by the k_1 values at lower temperature, the expected lifetimes of the native state at higher temperatures range between 90 s (at 78.5 °C) and

Table 1: Activation Parameters for the Thermal Unfolding of RdPf at pH 2[a]

technique	E_a (kcal mol^{-1})	A (s^{-1})	$\Delta H^{\ddagger}$ (kcal mol^{-1})	$\Delta S^{\ddagger}$ (cal mol^{-1} deg^{-1})	$\Delta G^{\ddagger}$ [c] (kcal mol^{-1})
UV−vis[b]	61 ± 2	$4 \times 10^{36} \pm 2 \times 10^{36}$	60 ± 2	101 ± 4	25 ± 3
fluorescence (first kinetic phase)	62 ± 1	$2 \times 10^{36} \pm 3 \times 10^{36}$	61 ± 1	105 ± 3	24 ± 2
fluorescence (second kinetic phase)	13 ± 5	$6 \times 10^{4} \pm 4 \times 10^{5}$	13 ± 5	−39 ± 15	27 ± 10
CD (first kinetic phase)	59 ± 1	$2 \times 10^{35} \pm 2 \times 10^{35}$	58 ± 1	101 ± 2	23 ± 2
CD (second kinetic phase)	62 ± 2	$2 \times 10^{36} \pm 7 \times 10^{36}$	62 ± 2	105 ± 7	25 ± 4

[a] The process has been followed by UV−vis absorption, fluorescence, and circular dichroism spectroscopies. Activation energies (E_a), Arrhenius preexponential factors (A), activation enthalpies ($\Delta H^{\ddagger}$), activation entropies ($\Delta S^{\ddagger}$), and activation free energies ($\Delta G^{\ddagger}$) were individually derived for each of the experimentally observed kinetic phases. Unless otherwise stated, uncertainties have been derived by applying error propagation routines (*47*) to individual errors on the slopes and intercepts of Arrhenius and Eyring plots. [b] Activation parameters and uncertainties taken from previously published absorption data (*17*). [c] These values are for a temperature of 75 °C.

170 s (70.1 °C). These times are definitely comparable with the dead time of the CD measurements (about 60 s). Therefore, the corresponding kinetic values cannot be accurately defined at temperatures higher than 70 °C. Having taken this into consideration, we have not included the k_1 values obtained at 70.1 and 75.7 °C in the Arrhenius curve fit for this kinetic phase. However, we have retained the k_1 value at 78.5 °C (dead time, 32 s), simply because it produces consistent results with the data obtained at the two lowest temperatures examined and it provides a rough estimate of the expected high-temperature behavior. It is, nonetheless, important to remember that the Arrhenius curve for this CD phase is affected by a considerably large error, given the limited number of available experimental points and the unavoidable low accuracy associated with the data acquired at 78.5 °C. As a consequence, it is necessary to bear in mind that the activation parameters for this kinetic step are affected by an even larger experimental error than quoted in Table 1.

Activation Parameters. A complete compilation of unfolding activation parameters, according to both Arrhenius and Eyring theories, is presented in Table 1. With the exception of the second fluorescence kinetic phase, all the values referring to the different kinetic steps are very similar to each other. The activation parameters for the second fluorescence kinetic phase are affected by quite a large error, due to considerable scattering of the experimental rates values in the corresponding Arrhenius plot (Figure 4). Therefore, the resulting data should be regarded as valid only from a qualitative point of view to establish broad comparisons with the other kinetic phases. As a consequence, we have kept our data evaluation below strictly within these limits. Additionally, all Arrhenius preexponential factors A are also affected by a large experimental error. This is unavoidable, given the relatively narrow unfolding temperature range that is experimentally accessible for proteins. Again, A values should only be considered to establish broad qualitative comparisons. Having taken this into consideration, we observe that all the activation parameters for the last unfolding kinetic step are considerably lower than those for the other RdPf unfolding steps.

A comparative Arrhenius plot for RdPf unfolding as followed by absorption, Trp fluorescence emission and CD spectroscopies is presented in Figure 4. A complex five-state four-step unfolding sequential mechanism can be immediately proposed by inspection of these data. This mechanism comprises initial partial loss of secondary structure (fast CD phase), followed by Fe^{3+} release (fast fluorescence phase, absorption data) and by loss of further

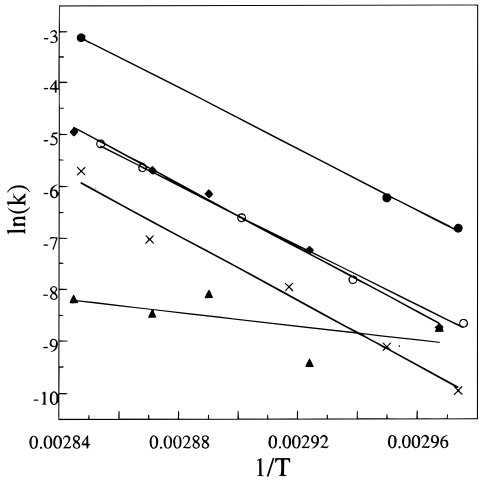


FIGURE 4: Comparative Arrhenius plots for RdPf thermal unfolding at pH 2, monitored by different spectroscopic techniques: ●, circular dichroism data, first phase; ○, absorption data (taken from Cavagnero et al. (1997)); ◆, fluorescence data, first phase; × circular dichroism data, second phase; ▲, fluorescence data, second phase.

secondary structure (slow CD phase). Extensive solvent exposure of the hydrophobic core (slow fluorescence phase) finally leads to the unfolded state formation. The broad features of this proposed mechanism do not undergo any significant changes over the experimentally tested temperature range.

DISCUSSION

Extension of previous absorption spectroscopy kinetic studies (*17*) to Trp fluorescence emission and CD has been instrumental to reveal that RdPf unfolding takes place via at least four kinetic steps. Absorption measurements at wavelengths corresponding to the ligand to metal RdPf charge-transfer transitions, i.e., in the visible region of the spectrum, are a good probe for the Fe^{3+} environment. Trp fluorescence emission is exquisitely sensitive to the chemical environment surrounding Trp(s) in a protein, while far-UV CD is a good probe of its overall degree of secondary structure.

Multistep Unfolding. RdPf unfolds, i.e., it loses its secondary and tertiary structures, on time scales comparable to those associated with Fe^{3+} release. More importantly, distinct kinetic events have been identified along the unfolding pathway. While the absorption data, which display monophasic kinetics, suggest that Fe^{3+} loss takes place in a single event, the fluorescence and CD denaturation profiles are biphasic, clearly pointing to intermediate formation

(Figures 1 and 3). This is particularly interesting since single-domain protein unfolding usually takes place as an all-or-none event (*22*), with no detectable kinetic intermediates. As a result, single exponential unfolding kinetic traces, bearing no lag phase, are normally observed (*23*−*29*). Furthermore, probes of both secondary and tertiary structures usually give rise to very similar rate constants (*27*, *30*), supporting a cooperative process. An exception is given by ribonuclease A (RNase A), which presents a lag phase preceding the detectable unfolding time course, as observed by NMR (*31*). However, the unfolding of RNase A appears entirely monophasic when followed by far- and near-UV CD spectroscopy (*31*). More recently, NMR has been able to reveal similar unfolding features for lysozyme (*32*) and dihydrofolate reductase (DHFR) (*33*). A different behavior has been observed for multidomain proteins such as phosphoglycerate kinase, whose unfolding proceeds via distinct kinetic phases, each corresponding to the independent unfolding of one of its two domains (*34*). This general trend for unfolding should be contrasted with the behavior of typical spectroscopic probes in refolding processes. These are known to be often characterized by several kinetic intermediates (*35*).

By simple comparison of reaction rates at a similar temperature, it is apparent that the first fluorescence phase for RdPf unfolding corresponds to the same event as the one detected by UV−vis. Accordingly, the relative Arrhenius plots are, within experimental error, virtually superimposable (Figure 4). Activation parameters from UV−vis are also very similar to those from fluorescence (Table 1). The presence of distinct unfolding kinetic phases may well be related to the fact that RdPf is a metalloprotein. The cysteine−tetracoordinate Fe^{3+} core is kinetically very stable, given the high unfolding $\Delta G^{\ddagger}$ derived from absorption measurements in the visible region (*17*). Moreover, iron release only occurs after partial secondary structure relaxation has taken place (Figure 4). However, this step is not rate-determining, since the unfolding events that follow Fe^{3+} release are even slower. This rules out the fact that Fe^{3+} is solely responsible for the kinetic stability of RdPf. Furthermore, it is possible that kinetic complexities in RdPf unfolding arise, at least partially, because Fe^{3+} coordination bonds are released in sequential, slow steps. This would not disagree with any of our data, particularly the observed monoexponential decay of the UV−vis absorption signal. This hypothesis will be tested in future experiments. A clear-cut detection of unfolding intermediates in metalloproteins has generally not been achieved, so far. These have only been detected indirectly from equilibrium thermal unfolding measurements, as in the case of bacterial cytochrome P-450 (*36*) or, more recently, from the guanidine hydrochloride dependence of unfolding rates, as for cytochrome c_2 (*37*). The unfolding kinetics of the mesophilic RdCp presented here (Figure 1b) provides an excellent comparison to RdPf, since these two proteins share several structural features, including the presence of a four-Cys one-Fe iron−sulfur center. The monoexponential unfolding behavior of RdCp suggests that the kinetic complexities observed in RdPf are an unusual and peculiar property of this hyperthermophilic protein, rather than a mere consequence of the presence of an iron−sulfur center.

Detection of unfolding intermediates before the unfolding transition state is important because it enables one to gain important insight into the RdPf energy landscape in the configurational space region shared between the unfolded form and the transition state or ensemble of transition states. This has generally not been possible before with mesophilic proteins since most of their unfolding intermediates, if present, are not sufficiently populated to be detected. The reason this has been possible for a hyperthermophilic rubredoxin lies in the fact that all the RdPf unfolding steps are slow. This naturally allows sufficient populations of intermediates to build up, and consequently be observed, in real time.

Fluorescence Quenching Events. Kinetic traces of typical fluorescence protein unfolding processes are characterized by a decrease in the observed signal with time. This is usually interpreted in terms of solvent quenching due to increased exposure of the Trp-containing hydrophobic probe to the solvent, in the unfolded form. In our case, the observed fluorescence quantum yield increases as the intermediate is formed. This is due to Fe^{3+}, a strong fluorescence quencher, being released during this step. The second fluorescence phase corresponds to an increased exposure of the two RdPf Trp to the solvent, which is reflected in a gradual decrease in fluorescence quantum yield. This hypothesis is supported by the CD experiments, which also show biphasic kinetics with different kinetic constants (Figure 3).

Activation Parameters. The activation parameters derived from the absorption, CD, and fluorescence data are presented in Table 1. Both preexponential factors (A) and activation energies (E_a) are very high for all the kinetic phases, except for the slower fluorescence kinetic step. Furthermore, the activation parameters for UV−vis and fluorescence (first phase) suggest that the two probes describe the same process. Additionally, both activation entropy and enthalpy due to partial release of secondary structure (first and second CD phases) are similar to those for the step involving Fe^{3+} loss. It follows that these unfolding events, despite being distinctly different in nature, are characterized by comparable free energy barriers, and they therefore have to meet similar kinetic demands. A rigorous comparison of activation parameters is not possible because of the associated experimental error. However, in qualitative terms, it is possible to say that the slow rates observed for the second fluorescence phase, the rate-determining step for unfolding, do not arise from high E_a, but from a particularly small A, which is in fact much lower than for the other kinetic steps. Considered from another point of view, this step, which accounts for final hydrophobic relaxation before reaching the unfolded state, is characterized by a very low $\Delta S^{\ddagger}$, implying that the protein already has a degree of disorder comparable to that of the unfolded state before going through this step. The calculated $\Delta S^{\ddagger}$ value for the second fluorescence phase is actually negative. Although very surprising, this finding should not be taken literally given the quality of the data that we have in hand. More accurate experiments, including a higher number of measured kinetic rates, are necessary to evaluate whether these results hold valid. Finally, regardless of the individual values of A or E_a, an important difference between the unfolding of RdPf and that of RdCp (*17*) is that the overall unfolding rates of RdPf are

very slow. These slow rates arise as a result of a combination of smaller A and/or larger E_a, compared to RdCp.

Unfolding Mechanism. Comparative Arrhenius plots for RdPf unfolding monitored by different techniques (Figure 4) reveal that this is a complex kinetic process. As a result of the analysis in Figure 4, the simplest scheme consistent with the experimental data comprises the following sequential model:

$$N \rightleftharpoons I_1 \rightleftharpoons I_2 \rightleftharpoons I_3 \rightleftharpoons U \ (\rightarrow A) \qquad (11)$$

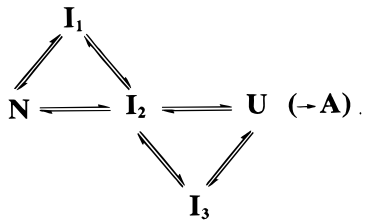
The native state partially loses some secondary structure (I_1, first CD phase), Fe^{3+} is then released (I_2, first fluorescence phase, absorption data), more secondary structure is then lost (I_3, second CD phase), followed by formation of the unfolded state (U). Irreversible aggregation and/or degradation either coincide with the last kinetic step or they immediately follow the formation of the denatured species. This kinetic scheme is entirely consistent with the concentration dependence studies described in the Results. According to Figure 4, the transition state for unfolding is crossed as Trp gets more exposed to the solvent, i.e., during the last step of eq 11. This is what is commonly called a late transition state along the unfolding pathway, since most structural elements (i.e., secondary and tertiary structure loss) are released during later stages of unfolding. One would be tempted to say that this corresponds to an early folding transition state, assuming that folding and unfolding pathways are essentially identical, except for the direction in which they are traveled. The principle of microscopic reversibility suggests that this is the case if folding/unfolding takes place at equilibrium. However, this does not generally apply to kinetic experiments. Additionally, we are unable to experimentally test the reversibility of the process, as discussed in the next section. Therefore, a late unfolding transition state does not necessarily imply an early folding transition state, in our case. The existence of a kinetic step involving partial loss of secondary structure prior to Fe^{3+} release is fully consistent with equilibrium data, which identify structural fluctuations whose free energy for hydrogen−deuterium exchange is considerably smaller than that for the transition between the native and an Fe^{3+}-depleted state (*16*). Although the main features of the proposed mechanism apply to the entire experimental temperature range, the data indicate that the second kinetic fluorescence phase, corresponding to hydrophobic relaxation, becomes relatively faster than the other events, toward lower temperatures. In other words, hydrophobic relaxation is no longer rate-limiting under these conditions. This is an important implication of the very low E_a that has been observed for this step.

Unfolding vs Aggregation and Self-Degradation. A word of caution must be introduced at this point. The thermal unfolding of RdPf is irreversible in our hands. The present data do not allow an assessment of when the irreversible step(s) take place. The simplest-case scenario is that this is a fast phenomenon, which is triggered by hydrophobic exposure upon unfolding. If this were the case, these irreversible event(s) would take place right after unfolding. Moreover, they would be kinetically undetectable. Alternatively, it is possible that these side reactions take place on the same time scale as the observed secondary and tertiary structure loss by RdPf. In this regard, the cysteinyl thiols that are freed upon Fe^{3+} release are potential cross-linking agents. However, thiol-coupling processes are normally facile at pH 8−8.4 in solution, whereas our working conditions are at pH 2. Nonetheless, the high temperatures involved in our experiments may still favor chemical cross-linking. To test this hypothesis, we have performed control experiments in the presence of a strong excess of DTT (data not shown). We have not detected any differences in the observed unfolding kinetics. As shown in the previous section, the negligible concentration dependence of RdPf unfolding rates in the concentration range relevant to our experiments suggests that either aggregation does not take place or it is a very fast event taking place only after the unfolded state has formed. An additional side process that may compete with unfolding is acid-catalyzed peptide backbone hydrolysis (*38*). Whether and at what stage of RdPf denaturation this reaction takes place needs to be experimentally assessed in future studies. For the current study, the following considerations suffice to suggest that hydrolytic degradation is more likely to follow than to be concurrent with RdPf unfolding. Amide bond hydrolysis usually requires more extreme conditions than those employed in this study (*39*), but it is greatly facilitated by the presence of aspartic acid (*40*), particularly if followed by proline (*41*) and asparagine. These residues are all present and highly conserved in most rubredoxins from different species, including RdPf and RdCp. On the other hand, hyperthermophilic proteins, particularly when nearly conformationally intact, are well-known for being extremely resistant to hydrolysis and other self-degradative reactions relative to their mesophilic couterparts, at pH lower than neutral (*42*, *43*). Additionally, hyperthermophilic proteins have generally been shown to lose activity faster than they become hydrolyzed. This suggests that degradation usually follows unfolding, and not vice versa (*42*). Finally, water-assisted peptide backbone cleavage at low pH exhibits low activation energies and, at the same time, relatively high positive activation entropies (*44*). None of the kinetic steps detected in this study satisfy this criterion. Therefore, given all the above considerations, the observed kinetic complexities are likely to be caused by intramolecular events peculiar to the RdPf unfolding process. Furthermore, regardless of the nature and the stage at which the irreversible steps take place, it appears unlikely that these are responsible for most of the significant secondary structure relaxation that takes place upon RdPf unfolding. As a result, our studies can be regarded as relevant to the understanding of the actual unfolding pathway of RdPf.

Features of the Unfolding Transition State. In thermodynamic terms, the folding/unfolding heat capacity change is commonly believed to be proportional to the difference in solvent exposure of hydrophobic groups between native and denatured states (*45*). Similarly, the hydrophobic character of a folding/unfolding transition state can be qualitatively evaluated, with respect to the starting species, by the magnitude of the activation heat capacity ($\Delta C_p^{\ddagger}$) for the rate-determining step (*46*). This, in turn, is proportional to the curvature of the Arrhenius plot for the rate-determining step. It is clear from Figure 4 that, in general, the unfolding transition states for the various kinetic steps do not involve significant changes in hydrophobicity (i.e., the Arrhenius plots are linear) with respect to the ground state or the initial

state for that step. However, since the experimental data points for the Arrhenius plot of the rate-determining step are more scattered than usual, we cannot draw definitive conclusions on this point with respect to this step on the basis of our data. Since the rate-determining step involves exposure of the hydrophobic core to the solvent, it may well be that $\Delta C_p^{\ddagger}$ is different from zero for this step.

Sequential vs Parallel Pathways. Folding pathways are potentially extremely complex. A more elaborate model takes into account the fact that some of the pathways may take place in parallel with others. We constrained this alternative model to account for the requirement that I_2 has to be formed along the path, as shown below:



(12)

Both models 11 and 12 are consistent with the observed kinetic data. It is not possible to definitely say which of the two models is correct. However, in light of the available experimental evidence, we favor the first one on the grounds that the CD data satisfactorily fit a biexponential expression. In case the second model were operative, the CD kinetic trace would, in principle, fit best to a quadriexponential expression. Nonetheless, in the case where the kinetic rates for the different steps were relatively similar, it would be virtually impossible to discriminate between the two mechanisms.

The slow thermal denaturation rates of RdPf have allowed direct detection of kinetic complexities taking place upon unfolding. The present studies suggest that the process takes place via a complex multistep pathway that starts with partial loss of secondary structure followed by Fe^{3+} release. In a further step, more secondary structure elements are relaxed. The process ends with exposure of the Trp environment (i.e., the hydrophobic core) to the solvent, followed by, or concomitant with, irreversible step(s). It is hoped that these experiments will prompt more unfolding and refolding kinetic investigations on hyperthermophilic proteins soon.

ACKNOWLEDGMENT

We thank Prof. Doug Rees, Dr. Reginald Waldeck, Dr. Brian Schultz and Ms. Lisa Bibbs for interesting and pertinent discussions.

REFERENCES

1. Creighton, T. E. (1990) *Biochem. J. 270*, 1−16.
2. Lecomte, J. T. J., and Matthews, C. R. (1993) *Protein Eng. 6*, 1−10.
3. Matthews, C. R. (1993) *Annu. Rev. Bioch. 62*, 653−683.
4. Anfinsen, C. B. (1973) *Science 181*, 223−230.
5. Jaenicke, R. (1996) *FASEB J. 10*, 84−92.
6. Rehaber, V., and Jaenicke, R. (1993) *FEBS Lett. 317*, 163−166.
7. Schultes, V., and Jaenicke, R. (1991) *FEBS Lett. 290*, 235−238.
8. Woese, C. R., Magrum, L. J., and Fox, G. E. (1978) *J. Mol. Evol. 11*, 245−252.
9. Adams, M. W. W. (1993) *Annu. Rev. Microbiol. 47*, 627−658.
10. Adams, M. W. W., and Kelly, R. M. (1995) *Chem. Eng. News* Dec. 18, 32−42.
11. Flam, F. (1994) *Science 265*, 471−472.
12. Stetter, K. O., Fiala, G., Huber, G., Huber, R., and Segerer, A. (1990) *FEMS Microbiol. Rev. 75*, 117−124.
13. Cavagnero, S., Zhou, Z.-H., Adams, M. W. W., and Chan, S. I. (1995) *Biochemistry 34*, 9865−9873.
14. Day, M. W., Hsu, B. T., Joshua-Tor, L., Park, J.-B., Zhou, Z. H., Adams, M. W. W., and Rees, D. C. (1992) *Protein Sci. 1*, 1494−1507.
15. Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., and Ladenstein, R. (1996) *J. Mol. Biol. 264*, 1132−1144.
16. Hiller, R., Zhou, Z. H., Adams, M. W. W., and Englander, S. W. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 11329−11332.
17. Cavagnero, S., Debe, D. A., Zhou, Z. H., Adams, M. W. W., and Chan, S. I. (1998) *Biochemistry 37*, 3369−3376.
18. Blake, P. R., Park, J.-B., Bryant, F. O., Aono, S., Magnuson, J. K., Eccleston, E., Howard, J. B., Summers, M. F., and Adams, M. W. W. (1991) *Biochemistry 30*, 10885−10895.
19. Blake, P. R., Park, J.-B., Zhou, Z. H., Hare, D. R., Adams, M. W. W., and Summers, M. F. (1992) *Protein Sci. 1*, 1508−1521.
20. Arrhenius, S. (1889) *Z. Phys. Chem. 4*, 226−248.
21. Eyring, H. (1935) *Chem. Rev. 17*, 226−248.
22. Creighton, T. E. (1993) in *Proteins—Structures and Molecular Properties* pp 310−311, Freeman, New York.
23. Chen, B., Baase, W. A., and Schellman, J. A. (1989) *Biochemistry 28*, 691−699.
24. Chen, X., and Matthews, C. R. (1994) *Biochemistry 33*, 6356−6362.
25. Kuwajima, K., Mitani, M., and Sugai, S. (1989) *J. Mol. Biol. 206*, 547−561.
26. Mücke, M., and Schmid, F. X. (1994) *Biochemistry 33*, 12930−12935.
27. Yamasaki, K., Ogasahara, K., Yutani, K., Oobatake, M., and Kanaja, S. (1995) *Biochemistry 34*, 16552−16562.
28. Segawa, S.-I., and Sugihara, M. (1984) *Biopolymers 23*, 2473−2488.
29. Serrano, L., Matouschek, A., and Fersht, A. R. (1992) *J. Mol. Biol. 224*, 805−818.
30. Schmid, F. X. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 197−242, Freeman, New York.
31. Kiefhaber, T., Labhardt, A. M., and Baldwin, R. L. (1995) *Nature 375*, 513−515.
32. Laurents, D. V., and Baldwin, R. L. (1997) *Biochemistry 36*, 1496−1504.
33. Hoeltzli, S. D., and Frieden, C. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9318−9322.
34. Beechem, J. M., Sherman, M. A., and Mas, M. T. (1995) *Biochemistry 34*, 13943−13948.
35. Dobson, C. M., Evans, P. A., and Radford, S. E. (1994) *Trends Biochem. Sci. 19*, 31−37.
36. Nöltig, B., Jung, C., and Snatzke, G. (1992) *Biochim. Biophys. Acta 1100*, 171−176.
37. Sauder, M. J., MacKenzie, N. E., and Roder, H. (1996) *Biochemistry 35*, 16852−16862.
38. Brown, R. S., Bennet, A. J., and Slebocka-Tilk, H. (1992) *Acc. Chem. Res. 25*, 481−488.
39. Bodanzsky, M. (1988) *Peptide Chemistry, A Practical Textbook*, Springer-Verlag, New York.
40. Schultz, J. (1967) *Methods Enzymol. 11*, 255−263.
41. Piszkiewicz, D., Landon, M., and Smith, E. L. (1970) *Biochim. Biophys. Res. Commun. 40*, 1173−1178.
42. Daniel, R. M., Dines, M., and Petach, H. H. (1996) *Biochem. J. 317*, 1−11.
43. Ahern, T. J., and Klibanov, A. M. (1985) *Science 228*, 1280−1284.
44. Antonczak, S., Ruiz-Lopez, M. F., and Rivail, J. L. (1994) *J. Am. Chem. Soc. 116*, 3912−3921.

45. Privalov, P. L., and Makhatadze, G. I. (1990) *J. Mol. Biol. 213*, 385−391.

46. Oliveberg, M., Tan, Y.-J., and Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8926−8929.

47. Bevington, P. R., and Robinson, D. K. (1992) in *Data Reduction & Error Analysis for the Physical Sciences*, 2nd ed., McGraw-Hill, New York.

BI9721804