Temperature and Driving Force Dependence of the Folding Rate of Reduced Horse Heart Cytochrome *c*[†]

Torbjörn Pascher*

Department of Chemical Physics, Lund University, Getingevägen 60, S-221 00 Lund, Sweden

Received November 14, 2000; Revised Manuscript Received February 2, 2001

ABSTRACT: Utilizing the stability difference between the ferro and ferri forms of horse heart cytochrome *c* (cyt *c*), folding of reduced cyt *c* was triggered by laser-induced reduction of unfolded oxidized cyt *c*. Measurements were made of the kinetics of the main folding phase (1 ms−10 s) in which collapsed reduced cyt *c* transforms to the native conformation. The folding rates were studied extensively as a function of temperature (5−75 °C) and guanidine hydrochloride (GdnHCl) concentration (1.6−4.9 M). At constant [GdnHCl], the Arrhenius plot of the folding rate constant (*k*) is nonlinear. At temperatures above 40 °C, the decrease in protein stability counteracts the expected increase in folding rate. Introducing free energy (ΔG), derived from protein stability data, into the Eyring and Arrhenius equations leads to: $\ln k = \ln(k_b T/h) + \Delta S^{\ddagger}/R - \Delta H^{\ddagger}/RT - \theta_m \Delta G/RT = \ln A - E_a/RT - \theta_m \Delta G/RT$, where θ_m is the ratio between the denaturant dependence of the folding rate and the stability. By using this equation at constant ΔG [or constant equilibrium constant (*K*)], linear Arrhenius plots are obtained. For the main folding phase of reduced cyt *c*, a positive $\Delta S^{\ddagger}$ is obtained indicating that the transition state is less ordered than the reactant. A model is proposed in which reduced cyt *c* first collapses into a compact intermediate, which needs to expand to reach the transition state of the rate-limiting folding reaction.

The kinetics of protein folding have been studied as a function of denaturant concentration (*1*−*10*), free energy (*4, 5*), equilibrium constant (*11*), and temperature (*1*−*3*). Denaturing agents, such as guanidine hydrochloride (GdnHCl)[1] or urea, generally decrease ln *k* (*k* is the rate of protein folding) linearly with increasing denaturant concentration. Also, a linear free energy relationship (linearity of the plot of ln *k* versus ln *K*, where *K* is the equilibrium constant between unfolded and folded protein) has been established for the folding of reduced horse heart cytochrome *c* (cyt *c*) (*4, 5*). Since protein stability decreases with increasing temperature and/or denaturant concentrations, the Gibbs free energy of folding, ΔG, has been modeled with a temperature dependence that contains the heat capacity at constant pressure, ΔC_p (*12*), and a linear denaturant concentration dependence, *m*, (*13*) according to Scalley et al. (*1*):

$$\Delta G(T,[\text{GdnHCl}]) = \Delta G(T,[0]) - m[\text{GdnHCl}] \quad (1)$$

$$\Delta G(T,[0]) = \Delta H(T) - T\Delta S(T) \quad (2)$$

$$\Delta H(T) = \Delta H(T_0) - \Delta C_p(T - T_0) \quad (3)$$

$$\Delta S(T) = \Delta S(T_0) - \Delta C_p \ln(T/T_0) \quad (4)$$

where T_0 is an arbitrary reference temperature set to 295 K. Note that ΔH, ΔS, and ΔC_p in eqs 2−4 have their reference point at [GdnHCl] = 0.

The temperature dependence of the protein folding rate has been found to exhibit non-Arrhenius characteristics, when measured at constant denaturant concentration. Plots of ln *k* versus 1/*T* markedly deviate from linearity at high temperatures. Indeed, at sufficiently high temperatures, folding rates may actually drop (*14, 15*). Two explanations have been given: one is based on transition state theory (*16*):

$$\ln k = \ln D - \frac{\Delta G^{\ddagger}}{RT} \quad (5)$$

where *D* is the prefactor of the Eyring equation and $\Delta G^{\ddagger}$ is the free energy of activation. Treating $\Delta G^{\ddagger}$ as ΔG in eqs 1−4, the temperature dependence of the activation parameters $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ accounts for the nonlinearity (*1, 2*). The second explanation is that protein folding exhibits super-Arrhenius behavior (*17*), where E_a/RT in the Arrhenius equation is squared. It has, however, also been shown that if the equilibrium constant (*K*) for the protein folding reaction is kept constant by varying the denaturant concentration, linear Arrhenius plots are obtained (*1, 2*).

In previous studies, we have examined the folding kinetics of reduced cyt *c* as a function of free energy (*4, 5*). The advantage of using reduced cyt *c* in folding studies is based on the higher stability toward unfolding of reduced (Fe^{2+}) cyt *c* compared to oxidized (Fe^{3+}) cyt *c*. This difference in stability creates conditions which make it possible to start with the unfolded oxidized state and to initiate folding of the reduced form by fast laser-induced electron transfer (*4, 5, 18*). Phototriggering the reaction permits investigation of folding events in the nanosecond to second time range.

[†] This work was supported by the Swedish Natural Science Research Council, the Knut and Alice Wallenberg Foundation, and the Crafoord Foundation.

* Correspondence should be addressed to this author. Phone: +46 46 222 47 39, fax: +46 46 222 41 19, email: Torbjorn.Pascher@chemphys.lu.se.

[1] Abbreviations: GdnHCl, guanidine hydrochloride; cyt *c*, horse heart cytochrome *c*; ln *k*, natural logarithm of the rate constant (unfolded to folded); ln *K*, natural logarithm of the equilibrium constant (unfolded to folded); U → F, from unfolded to folded; U → ‡, from unfolded to the transition state.

10.1021/bi0026223 CCC: $20.00
Published on Web 04/21/2001

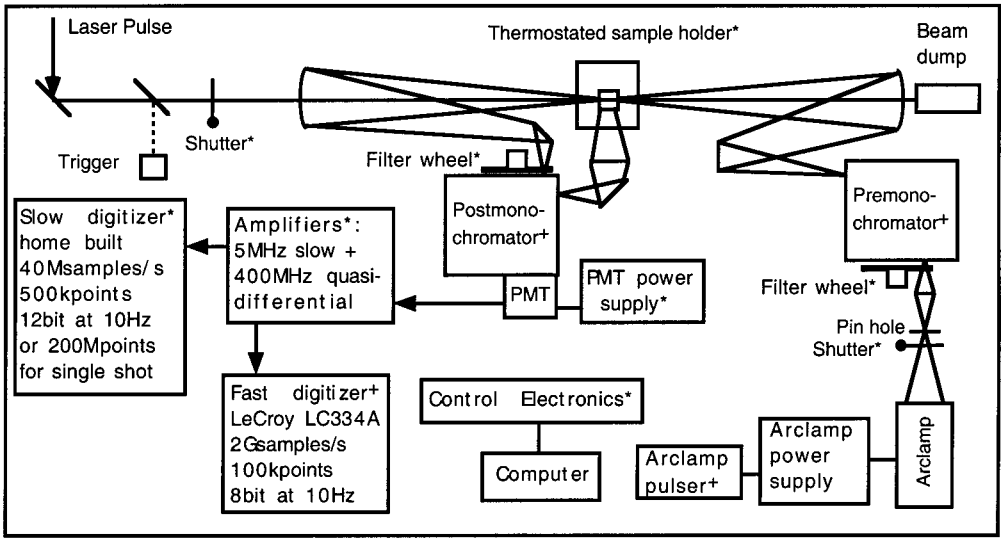

FIGURE 1: Laser flash photolysis detection system. For details, see text. + and * denote computer-controlled components, of which * are own constructions.

Moreover, the covalently bound porphyrin in cyt *c* offers an excellent spectral probe to monitor structural changes in the heme region.

I have now extended previous studies (*4*, *5*) by investigating the temperature and denaturant dependencies of the folding kinetics of cyt *c*. In this connection, I have also examined alternative expressions for the preexponential factor D (*1*−*3*), since the standard Eyring form (k_bT/h) appears inappropriate for protein folding reactions because of the large size of protein molecules and the way folding reactions are believed to proceed (*1*, *2*). Furthermore, I have modeled the free energy of activation ($\Delta G^{\ddagger}$), both as a function of temperature and denaturant concentration, as well as a function of temperature and free energy.

By substituting denaturant concentration by free energy (ΔG), the reference point for $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ becomes $\Delta G = 0$, which is a characteristic of the protein, instead of [GdnHCl] = 0, which is a characteristic of the solvent system. This leads to elimination of the temperature dependence of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ and reduction of the number of unknown parameters from 4 to 3. This approach facilitates comparisons of the folding kinetics of different proteins, as it emphasizes protein-related rather than system-related protein properties.

MATERIALS AND METHODS

Protein Samples. Horse heart cyt *c* and NADH were obtained from Sigma and used without further purification. Guanidine hydrochloride (GdnHCl) was USB Ultrapure, and the other chemicals were of standard quality. The protein concentration was 15 μM both in kinetics and in stability experiments. The buffer contained, in addition to varying amounts of GdnHCl, 0.1 M sodium phosphate, pH 7.0 at 20 °C. All measurements were performed with 1.5 ml of protein solution in 1 cm fluorescence cuvettes.

Instrumentation. Kinetics were measured in a laser flash photolysis setup (Figure 1) based on the one described previously (*19*). The setup consists of a Spectra-Physics Quanta-Ray Pro230-10/MOPO-710 laser as pump source. The probe light comes from a 75 W xenon arc lamp and is focused on the entrance slit of a SPEX TRIAX-180 single monochromator (which minimizes photolysis of the sample by the probe light during absorbance measurements and selects the steady-state excitation wavelength for tryptophan quenching experiments). The light out from the monochromator is refocused at the sample with a spherical concave mirror. The sample is mounted in a holder equipped with a computerized Peltier temperature control system and a magnetic stirrer. In the absorbance mode, the probe light is again refocused with a spherical concave mirror on the axial entrance slit of a SPEX 270M monochromator/spectrograph and detected with a Hamamatsu R928 photomultiplier tube (PMT) mounted on a five stage base at the monochromator exit slit. The two spherical concave mirrors have holes through their centers to let the laser pulse pass through. In the tryptophan quenching mode, the fluorescence is collected at a 90° angle relative to the laser pulse and is imaged on the lateral entrance slit of the SPEX 270M after the image has been turned from horizontal to vertical. The light is collected at the spectrograph exit port, where a wavelength region of about 70 nm is focused on a Hamamatsu R928 tube mounted on a nine stage base. The signals from the PMTs are amplified with an overall bandwidth of >5 MHz (which can be limited to 1 MHz, 100 kHz, 10 kHz, or 1 kHz) and variable gain (56 kV/A to 3.2 MV/A). Signals were digitized by a transient recorder consisting of a 12 bit analog to digital converter (ADC) running at 40 Msamples/s, two 512 kword memory banks, and a summation unit. For repetitive signals, the ADC fills the first memory bank once for each laser shot, and the summation unit adds this trace to the sum of the previous ones stored in the second memory bank. For single event signals, the ADC fills the first memory bank in the same way as for repetitive signals, but the summation unit takes every second of 800 adjacent analog to digital conversions and sums them together. These sums are then stored in the second memory bank. The contents of these two memory banks are then compressed in a semi-logarithmic fashion by binning to a data set that covers 10

s and has a time resolution of 25 ns around the laser pulse. The binning improves the signal-to-noise ratio at longer times.

Protein Stability Measurements. The protein stability data were collected on the same equipment as the kinetics to ensure compatibility regarding bandwidth of the light, temperature, etc. The data consisting of absorbance spectra and tryptophan emission (excited at 285 nm and recorded around 360 nm) were collected at temperature intervals of about 5 °C over a temperature range of 5−75 °C and GdnHCl concentration intervals of 0.33 M over a suitable denaturant concentration range.

To the samples of oxidized protein was added 2 equiv of potassium hexacyanoferrate(III) to ensure that the protein was fully oxidized at all times. For data collection of reduced cyt *c*, the sample was kept in a sealed cuvette under an oxygen-free argon atmosphere. The problem of reducing and keeping cyt *c* completely reduced, despite its tendency to auto-oxidize, was solved by adding 15 μl of deoxygenated 2% sodium dithionite solution. In the presence of cyt *c*, dithionite is consumed slowly with a rate that increases with temperature. The added amount proved to be sufficient to keep cyt *c* reduced throughout a series of 15 successive recordings. Sodium dithionite contributed with about 1 absorbance unit at the excitation wavelength at the beginning, but was almost consumed toward the end of the 15 step temperature series. Due to internal absorbance, dithionite reduces the excitation (at 285 nm) and the observed emission (at 360 nm) of tryptophan fluorescence. This loss in tryptophan fluorescence intensity related to the amount of dithionite present was corrected using a calibration curve recorded at 40 °C. Examples of the absorbance spectra obtained are shown in Figure 2A.

Protein Folding Measurements. The samples used for kinetics measurements additionally contained 30 μM NADH as a photoreductant (*18*, *20*, *21*). The samples were deoxygenated and sealed in a cuvette. Before each laser shot, the sample was stirred for 12 s. A delay of an additional 12 s between the end of the stirring and the laser shot was found to be long enough for the convection of the sample to become insignificant. The probe spot had an area of 1 mm^2. The folding was initiated by a 7 mJ 355 nm laser pulse covering 5 mm^2 of the sample. This 5 time larger excitation volume was chosen to eliminate the effect of convection and diffusion during data collection. Each kinetics trace was initiated with a single laser pulse and collected for 10 s. The total time required for the collection of one kinetic trace, comprising stirring, delay time, laser pulse, and trace recording, was 35 s.

For transient absorbance data, one trace each was collected at 400, 415, 435, and 550 nm. These absorbances provide information on oxidation and environmental states of the heme group. At 400 nm, the disappearance of oxidized cyt *c* and at 415 nm the appearance of reduced cyt *c* can be observed. The absorbance at 550 nm together with that at 415 nm is used to distinguish between reduced unfolded and reduced folded cyt *c* (see Figure 2). The absorbance at 435 nm gives indications of heme environmental changes in early folding events. For tryptophan fluorescence quenching data, two traces (excited at 285 nm and collected at around 360 nm) were recorded for each temperature−[GdnHCl] pair. Because of their significantly higher noise levels, the tryptophan fluorescence quenching traces were only used to identify folding events.

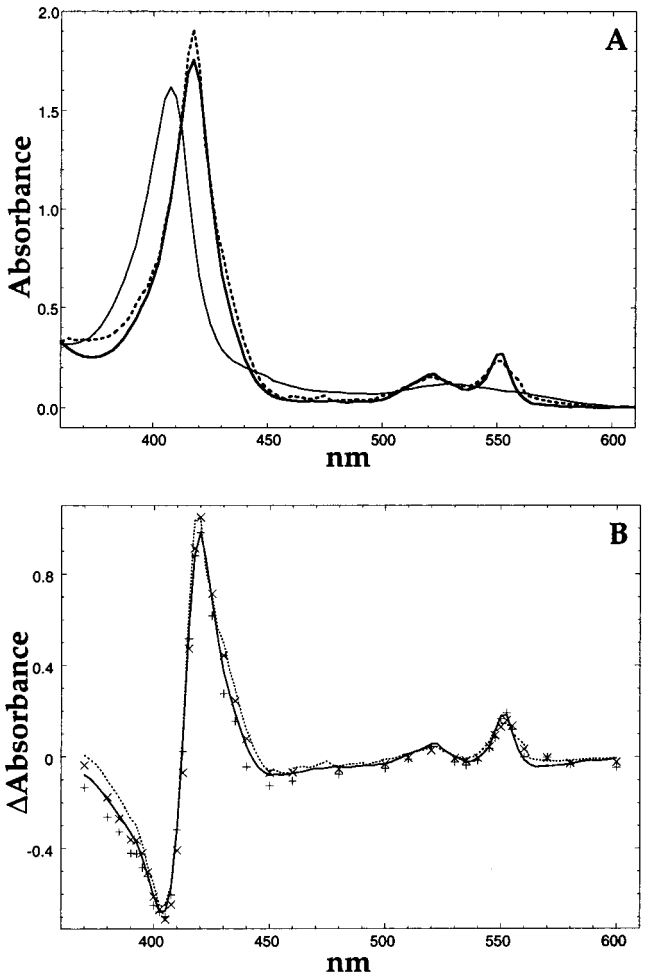

FIGURE 2: (A) Absorbance spectra of oxidized unfolded (thin), reduced folded (thick), and reduced unfolded (dashed) cyt *c* at 3.27 M GdnHCl and 314 K. The spectra were obtained from linear fits of the [GdnHCl] dependence of measured spectra. Since reduced unfolded cyt *c* does not exist at the given conditions, its spectrum has been extrapolated from a series of spectra measured at [GdnHCl] >5 M. This procedure gives spectra that look more similar to those obtained in the kinetics experiments than one measured directly at 7 M GdnHCl. Note that details in the shape of these three spectra can vary both with temperature and with [GdnHCl]. (B) Difference spectra of reduced cyt *c* minus oxidized unfolded cyt *c*. Dashed and solid lines represent unfolded reduced and folded reduced cyt *c*, respectively, calculated from the spectra shown in (A). The symbols × and + represent the spectra obtained from fits of the kinetics at different wavelengths, before (×) and after (+) the main folding phase, respectively. Both these spectra have been corrected for the consumption of NADH.

Data Analysis. The transient absorbance traces were fitted globally to a sum of five exponentials (*22*):

$$A_0 + \sum_{i=1}^{5} A_i e^{-k_i t} \qquad (6)$$

The fitting procedure was a combination of a Marquardt−Levenberg nonlinear least-squares routine used for the rate constants and general linear regression used for the amplitudes.

The thermodynamic models were fitted with the Nelder−Mead Simplex method, and the errors were estimated with *F*-statistics at a confidence limit of 95% (*22*, *23*).

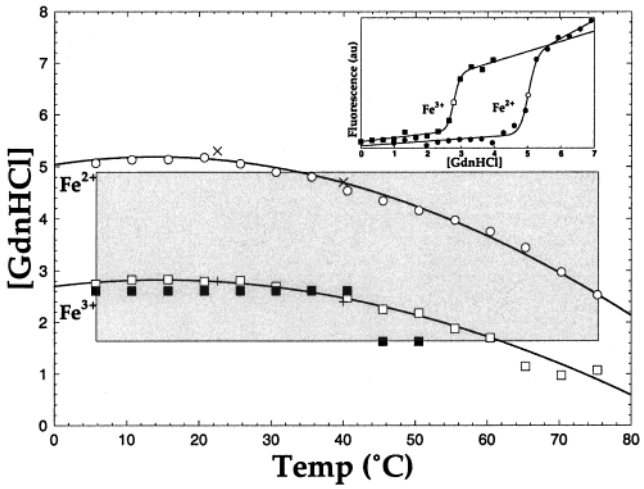

FIGURE 3: Relative stability of ferro- and ferri-cyt *c* as a function of temperature and denaturant concentration. The solid lines represent the midpoints of the denaturation curves as described by the thermodynamic parameters in Table 1. The open symbols (□, oxidized; ○, reduced) are the midpoints when fitting at each temperature individually. The shaded area represents the conditions for which folding experiments have been performed. The filled squares (■) indicate the breach points below which the ln *k* versus [GdnHCl] plot deviates from linearity (see Figure 6). For comparison, folding midpoints of oxidized (+) and reduced (×) cyt *c* from earlier work (*5*) have been included. Inset: Example of the tryptophan fluorescence curves for ferri- (■) and ferro-cyt *c* (●) as a function of [GdnHCl] at 299 K. The open symbols (□, oxidized; ○, reduced) show the midpoints used in the stability curves in the main figure.

RESULTS AND DISCUSSION

Protein Stability Data. The midpoints of the unfolding curves for oxidized and reduced cyt *c* are shown in Figure 3. The area between the curves shows conditions where reduction of unfolded ferri-cyt *c* yields at least a 50% conversion to folded ferro-cyt *c*. Folding experiments can also be performed outside this area, but the amplitudes of the folding phase become smaller and hence the measured folding rates become less accurate (compare legend of Table 2). Figure 3 shows the midpoints of the denaturation curves for both oxidized and reduced cyt *c* as determined from global fitting of the experimental data to eqs 1−4. The thermodynamic parameters resulting from the fits are shown in Table 1. For comparison, unfolding midpoints from a previous study (*5*), which show good agreement, have been added to Figure 3. When comparing ΔG of folding at zero denaturant concentration, relatively larger deviations (±10 kJ/mol) are observed, but are still within experimental errors.

Noteworthy is the similar shape of the stability curves of oxidized and reduced cyt *c*, indicating that the large free energy penalty (20 kJ/mol) for burying the extra charge of Fe^{3+} inside the protein is similar over the whole temperature range (*24*, *25*). The [GdnHCl]−temperature range within which the folding experiments have been performed is indicated by the shaded area in Figure 3.

Analysis of the Kinetics. An example of the kinetics traces recorded for each [GdnHCl]−temperature combination is shown in Figure 4. Within the recorded time range, five phases are distinguished. These five phases are fitted by the five exponentials of eq 6.

Phase **1** originates from the laser-induced fluorescence of NADH. The rate of this phase is affected by the recovery time of the amplifier. Similarly, the tryptophan fluorescence quenching trace cannot be followed during the first 200 μs, as the photomultiplier tube and the amplifier need considerable time to recover from the large amount of scattered laser light.

Phases **2** and **3** primarily comprise the reduction of cyt *c*. Laser photolysis of NADH has been shown to result in two reductants, a solvated electron and a NADH radical cation (*20*), which reduce folded ferri-cyt *c* in two steps (*21*). Presuming that the unfolded protein behaves in the same way, then the bleaching of the 400 nm absorbance and the corresponding increase of the 415 nm absorbance during phases **2** and **3** represent the reduction of cyt *c*, which is completed within 1 ms. The difference in the relative amplitudes as a function of time of the traces at 415 and 435 nm during phases **2** and **3**, however, shows that events other than reduction also occur during phase **3**. Most probably, early folding events, like those observed by Chen et al. (*26*) and Jones et al. (*27*), are hidden beneath these cyt *c* reduction phases. Since the present experiments had been designed with focus on the millisecond to second time frame (phase **4**), this question cannot be resolved from the present data set. The exact nature of the events in phase **3** will be the subject of future studies. Phase **5** represents the reoxidation of cyt *c*, presumably by trace amounts of oxygen (*26*).

Interpretation of the Folding Kinetics. Phase **4**, proceeding over a time frame of 1 ms to 10 s, represents the main folding event. The kinetics of this phase are analyzed in detail in this report. Phase **4** coincides with a decrease in tryptophan fluorescence, indicating that tryptophan-59 moves closer to the heme. The heme difference absorbance spectra at the beginning of phase **4** resemble those of reduced unfolded minus oxidized unfolded cyt *c* (*4*) (see Figure 2) and of reduced minus oxidized states of a bis-His model complex (*18*). At the end of phase **4**, the spectra resemble those of the folded reduced minus the unfolded oxidized protein (*4*, *18*). Earlier CD measurements within this time frame (*26*) had shown a change in the amplitude of the α-helix CD signal from 20% to about 95% of that of the native protein. In this CD study, three phases with increasing CD signal had been observed with rates of 2.9×10^5, 170, and 5.6 s^{-1} (or 9.1 s^{-1} depending on the fitting procedure used). (Additionally, also a phase with decreasing CD signal with a rate of 5.6×10^3 s^{-1} had been observed.) The question is, whether these three phases represent three parallel folding pathways, starting from three distinct subpopulations of unfolded cyt *c* with, e.g., different states of axial heme ligation, or reflect a series of consecutive steps in α-helix formation. A calculation of the folding rate for phase **4** for the conditions of the CD experiment (24 °C and 3.5 M GdnHCl) yields 3.1 s^{-1}. This rate is close enough to that of the slowest CD phase (5.6 s^{-1}), with the largest α-helix increase, to show that this CD phase and phase **4** represent the same process. A phase corresponding to the CD middle phase, with the smallest amplitude, could not be identified in the absorbance traces. This could be due to the fact that this phase represents a folding step which produces only a small change in secondary structure but no changes of the heme environment. The latter hypothesis is supported by the work by Jones et al. (*27*) which shows that the slowest of the reactions in the equilibrium scheme between species with different heme ligation is more than 10 times faster than the

Table 1: Thermodynamic and Apparent Activation Parameters for the Main Folding Phase of Cytochrome *c* at 295 K[a]

parameter	units	U → F Fe^{3+}	U → F Fe^{2+}	U → ‡ $D = k_bT/h \cdot Fe^{2+}$	U → ‡ $D = 10^{10} \cdot Fe^{2+}$	U → ‡ $D = cT/\eta \cdot Fe^{2+}$
ΔH	kJ/mol	-93 ± 21	-120 ± 20	53 ± 9	56 ± 9	41 ± 9
ΔS	J mol^{-1} K^{-1}	-170 ± 50	-190 ± 40	12 ± 30	74 ± 30	51 ± 30
ΔC_p	kJ mol^{-1} K^{-1}	-4.9 ± 1.6	-5.1 ± 1.1	-2.1 ± 0.6	-2.1 ± 0.6	-2.0 ± 0.6
m	kJ mol^{-1} M^{-1}	-15 ± 3	-12 ± 2	-5.9 ± 0.5	-5.9 ± 0.5	-5.6 ± 0.5

[a] F = folded, U = unfolded, ‡ = transition state, Fe^{2+} = reduced cyt *c*, Fe^{3+} = oxidized cyt *c*. Parameters for the unfolded-to-transition state reaction (U → ‡) have been calculated with three different expressions for *D* according to eqs 7, 8, and 9, respectively. The sum of squared differences between measured and calculated folding rates is the same for $D = k_bT/h$ and $D = 10^{10}$, but is 0.5% smaller for $D = cT/\eta$.

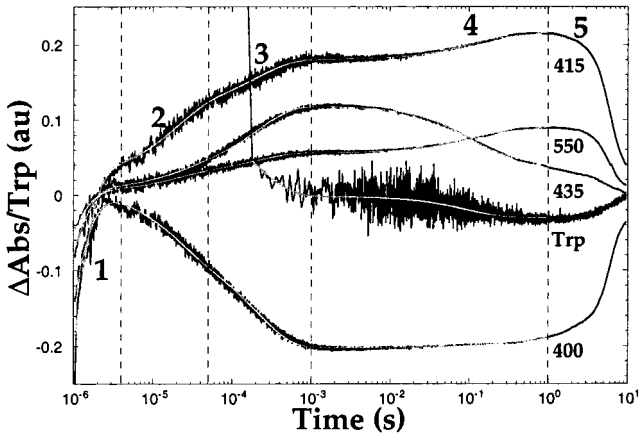

FIGURE 4: Kinetics measurements: Example of kinetics traces at four different wavelengths (in nm) and of tryptophan quenching (Trp) recorded at 324 K and 3.59 M GdnHCl. The measured data are shown in black and the lines of the global fits according to eq 6 in gray. The improvement in signal-to-noise ratio for longer times is due to heavier binning. The time resolution of the tryptophan fluorescence quenching trace has been reduced for clarity. Five different phases, **1**−**5**, are distinguished. Their ranges are indicated by broken lines. For the nature of these phases, see text. Phase **4** is the main folding event. In all, similar kinetics recordings have been performed for 165 different *T*−[GdnHCl] combinations (Table 2).

CD middle phase. Hence, it follows that if the parallel pathway theory is correct, all the protein molecules would have folded with the faster rate. From the present work, no conclusions can be drawn if the fastest of the CD phases represents a separate fast path to folded cyt *c* or is part of a consecutive folding scheme. To summarize, the spectral changes recorded in the present study indicate that cyt *c* during phase **4** changes from a collapsed, mainly His-His ligated state with low α-helix content to a correctly His-Met ligated fully folded state.

Nonlinear Arrhenius Plot. The rate constants (*k*) for the folding reaction of phase **4**, as a function of temperature and [GdnHCl], are compiled in Table 2. An example of the nonlinearity of the Arrhenius plot for the folding reaction at constant [GdnHCl] is shown in Figure 5. Note that the plot is almost linear from 5 to 40 °C. Within this temperature region, the midpoints of the denaturation curve in Figure 3 are roughly constant. This indicates that the deviation of the Arrhenius plot from linearity arises from a slower folding rate due to decreasing thermal stability of the folded protein.

Nonlinear ln k Dependence on [GdnHCl]. Deviations from linearity also are observed for the dependence of ln *k* on [GdnHCl] at low temperatures (Figure 6). This deviation at low [GdnHCl], where ln *k* is almost independent of [GdnHCl], could mean that most of the ferri-cyt *c* is already folded and that only the reduction of folded ferric-cyt *c* occurs. However, the tryptophan quenching traces show a decrease in fluorescence, indicating that protein folding does take place. Furthermore, the amplitude of the "folding" signal gets smaller as the [GdnHCl] is lowered and not the reverse, as would be expected if the observed signal comes from the already folded protein. Moreover, the start of the deviation from linear [GdnHCl] dependence (shown as squares in Figure 3) does not follow the stability curve of the oxidized state. In folding studies of oxidized cyt *c*, a similar deviation from linearity has been found in the ln *k*−[GdnHCl] plot (*6*, *7*).

The nonlinear dependence of ln *k* on [GdnHCl], which also has been observed for other proteins, has been interpreted in different ways. For a barnase mutant (*8*) and a domain of phosphoglycerate kinase (*9*), the nonlinearity has been taken as an indication of intermediates in the reaction pathway. For human U1A protein (*10*), depending on experimental conditions, it has been attributed either to an aggregation of the unfolded protein or to a shift of the transition state to more denatured structures. The aggregation had occurred at protein concentrations above 3 μM and [GdnHCl] <1.5 M and appeared as an increasing reduction of the folding rate as the [GdnHCl] was lowered. In the present experiment, the total cyt *c* concentration was 5 times higher than that for which aggregation of human U1A had been observed. However, at the conditions of lowest [GdnHCl], at which proteins are most sensitive for aggregation, the actual concentration of oxidized *unfolded* cyt *c* was significantly lower, because the folded form is more stable under these conditions. Moreover, the drop in the folding rate reported for humanU1A protein is not observed for cyt *c* (Figure 6). These circumstances together with the comparatively high [GdnHCl] at which the bend in the [GdnHCl] dependence occurs disfavor the aggregation hypothesis in the case of cyt *c*. Thus, the observed change of $m_{U\rightarrow\ddagger}$ as a function of [GdnHCl] indicates a shift in the folding mechanism between the low and the high [GdnHCl] environment. This can be caused either by a shift in the position of the transition state along the reaction coordinate (*10*), in which case the slope $m_{U\rightarrow\ddagger}$ in the ln *k* vs [GdnHCl] plot will change smoothly with [GdnHCl], or by a population of intermediates (*9*), in which case two linear segments of the plots ought to be separated by a noticeable bend. The shape of the curves in Figure 6 favors the second alternative.

The Preexponential Factor of the Eyring Equation. To account for nonlinear Arrhenius behavior at high [GdnHCl], the protein folding rates have been globally fitted according to transition state theory (eqs 1−5). The validity of the expression for the preexponential factor *D* in eq 5 has been discussed in the literature (*1*, *2*, *17*, *28*). Two different expressions have been used in studies of the folding rate de-

Table 2: Natural Logarithm of the Rate Constants of the Folding Reactions in Phase **4** as a Function of Temperature (K) and [GdnHCl][a]

temperature		[GdnHCl] (M)										
label	K	1.63	1.96	2.29	2.61	2.94	3.27	3.59	3.92	4.25	4.58	4.90
5	278.9	1.39[C]	2.40[C]	1.11[C]	0.98[C]	0.16	−1.31	−3.27[A]	−2.86[A]	−4.03[A]	−6.14[A]	−4.19[A]
10	283.9	1.94[C]	2.11[C]	1.58[C]	1.32[C]	0.68	−0.67	0.21	−2.34	−3.81[A]	−4.72[A]	−3.30[A]
15	288.9	2.40[C]	2.06[C]	2.13[C]	1.87[C]	1.57	0.92	−0.32	−1.09	−1.14	−4.98[A]	−2.81[A]
20	293.9	2.52[C]	3.00[C]	2.65[C]	2.63[C]	2.03	1.26	0.38	0.26	−0.52	−2.37	−1.94
25	298.9	2.83[C]	3.22[C]	3.07[C]	2.73[C]	2.53	1.74	1.12	0.99	−0.59	−1.38	−1.70
30	303.8	3.33[C]	3.46[C]	3.38[C]	3.24[C]	2.96	2.12	1.53	0.66	0.59	−1.27	−1.95
35	308.8	3.95[C]	3.87[C]	3.83[C]	3.77[C]	3.46	2.83	2.00	0.84	0.59	−1.15	−1.84[B]
40	313.7	4.37[C]	4.41[C]	4.22[C]	4.20[C]	3.82	3.04	2.18	0.90	0.80	0.10	−2.17[B]
45	318.7	4.78[C]	6.13	5.10	4.51	4.09	3.43	2.32	1.32	1.41	0.26[B]	−2.39[B]
50	323.7	5.23[C]	6.21	5.25	4.63	4.26	3.56	2.35	1.61	1.61[B]	3.98[B]	−2.48[B]
55	328.7	6.74	6.21	5.53	4.84	4.30	3.51	2.59	1.91	2.29[B]	4.26[B]	−3.04[B]
60	333.6	6.75	6.22	5.58	4.97	4.51	3.45	2.74	2.74[B]	3.26[B]	3.90[B]	−3.39[B]
65	338.5	6.79	6.36	5.59	4.66	4.63	3.72	3.38[B]	4.52[B]	4.03[B]	3.94[B]	−2.37[B]
70	343.5	6.74	ND	5.20	4.51	4.60[B]	4.15[B]	4.35[B]	3.62[B]	3.79[B]	3.87[B]	−2.84[B]
75	348.5	6.44	5.81	4.70	4.78[B]	5.04[B]	4.19[B]	3.73[B]	3.18[B]	4.01[B]	4.13[B]	−1.86[B]

[a] The column "label" indicates the approximate temperatures in degrees centigrade used as labels in the figures. ND, not determined. The rates marked with superscripts A, B, and C are not used in calculation of the apparent activation parameters. [A], rate too slow to be determined accurately ($t_{1/2}$ > 9.5 s). [B], unfolding dominates under these conditions. [C], deviation from linear [GdnHCl]-dependent response at low [GdnHCl] and low temperature. At these conditions, only a limited fraction (1−50%) of unfolded oxidized cyt *c* coexists with folded oxidized cyt *c* (cf. Figure 3). By laser-triggered reduction with NADH, this unfolded fraction is converted to unfolded reduced cyt *c* and will take part in the folding reaction. This affects the amplitude of the absorbance changes that record the folding event. However, the low amplitudes are still large enough to yield kinetic rates with reasonable accuracy. The temperature was measured by an HD-9214, Delta Ohm Pt-100 digital thermometer with temperature errors of less than 0.5 K.

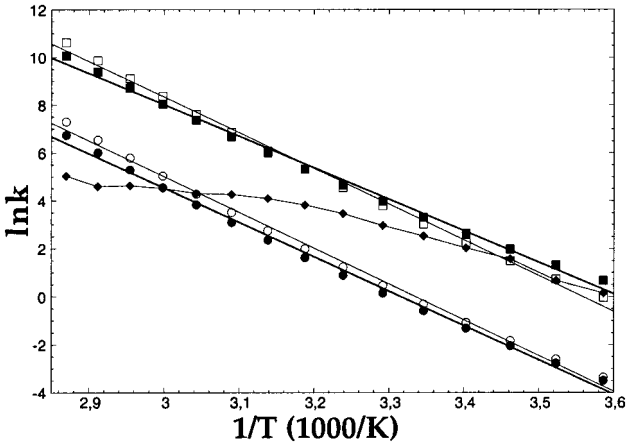

FIGURE 5: Arrhenius plots for the folding rate of reduced cyt *c* at constant [GdnHCl], *K*, and Δ*G*, respectively. (◆) The non-Arrhenius temperature dependence of the main folding phase at [GdnHCl] = 2.94 M. The plot is roughly linear up until 40 °C (0.00315 K^{-1}). The deviation of the 75 °C (0.00285 K^{-1}) point is attributable to the predominance of the unfolding reaction. The open symbols (□) and (○) show Arrhenius plots at constant equilibrium constant, $K = 100\,000$ and $K = 100$, respectively. The filled symbols (■) and (●) show Arrhenius plots at constant free energy, $\Delta G = -30$ kJ/mol and $\Delta G = -3$ kJ/mol, respectively. Data at constant *K* and Δ*G* were calculated from eqs 1−5 and 7.

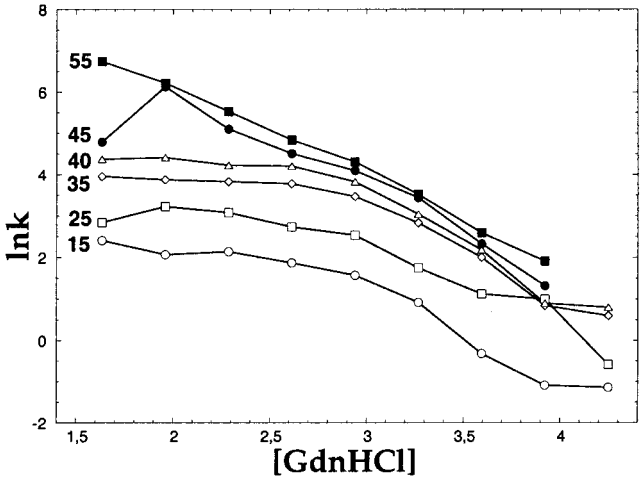

FIGURE 6: Nonlinear [GdnHCl] dependence of ln *k* at temperatures below 55 °C. The labels indicate the approximate temperature in degrees centigrade. For the exact conditions, see Table 2. The data indicate a shift in a folding mechanism that is [GdnHCl]-dependent at high GdnHCl concentrations to one that is not [GdnHCl]-dependent at low GdnHCl concentrations and low temperatures.

pendence on temperature and denaturant: one of these is the standard Eyring preexponential factor (*3*):

$$D = \frac{k_b T}{h} \tag{7}$$

where k_b is Boltzmann's constant and *h* is Planck's constant.

Since this prefactor has been derived for reactions of small molecules, in which a bond is made or broken, it may not be suitable for protein folding. The mechanism by which a protein molecule leaves the transition state is believed to be the addition of an amino acid to the native-like part of the transition state by chain diffusion. Hence, the following preexponential factor has been considered more appropriate (*1*, *2*):

$$D = 10^{10}\ \mathrm{s}^{-1} \tag{8}$$

This factor has been derived from theoretical estimates of the rate constant for a random coil-to-helix transition (where one amino acid in a random coil conformation is added to a preexisting helix stretch) and is of the order of $10^{10}-10^{11}$ s^{-1} (*29*). Equations 7 and 8 in connection with eqs 1−5 give equally good fits of the data, however, different values for $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$, and $\Delta C_p^{\ddagger}$. This makes it impossible to evaluate experimentally whether eq 7 or eq 8 is the more suitable, and the differences in the resulting parameters make it

difficult to compare data from different studies. Therefore, results from fits using both of the preexponential factors are given in Table 1. Equation 8 does not consider expected environmental dependencies for a process in which an unstructured amino acid is included in a structured segment in a diffusion-like process. It appeared reasonable to expect this chain diffusion to have the same dependencies as the Stokes−Einstein equation for calculating the diffusion coefficient. I have therefore tried a new way of describing D as a function of temperature and solvent viscosity (η):

$$D = \frac{cT}{\eta} \tag{9}$$

where c is a constant.

This expression is in agreement with Kramers' theory for diffusion-controlled reactions (*30*). Schmid and co-workers (*31*) have studied the effect of viscosity on protein folding reactions by adding viscosogens such as ethylene glycol and sucrose. Taking into account the stabilizing effect of the viscosogen, they found that for a small rapidly folding protein, like the cold-shock protein CspB from *Bacillus subtilis*, the folding and unfolding time constants, τ, are proportional to the change in relative viscosity for viscosity changes up to at least a factor of 2. Similar results have been obtained for protein L (*32*). On the other hand, if the protein quickly collapses to a molten globule conformation, no viscosity dependence is observed for the following folding step (*31*). Measurements of the actual rate constant for the addition of an amino acid to a helix segment at one end of a peptide yielded a rate constant of $3 \times 10^8 \times T/273\ s^{-1}$ (*33*). By using this rate constant taken from the middle of the peptide's unfolding curve (300 K) and the viscosity for water (*34*), the constant c of eq 9 is estimated to be roughly 9×10^2 kg m^{-1} s^{-2} K^{-1} (9×10^5 cP K^{-1} s^{-1}). The viscosity of the GdnHCl solutions and its dependence on temperature (*34*−*36*) and concentration (*37*) were estimated using data given in the literature.

Activation parameters calculated using the preexponential factors from eqs 7−9 are summarized in Table 1. The fits with prefactors according to eqs 7 and 8 are equally good; the fit with the prefactor from eq 9 is slightly better. The fact that no significant improvements are obtained may be due to either the crude way in which the viscosity was extrapolated or the possibility that phase **4** represents a molten globule-to-folded protein transition which should not exhibit any significant viscosity dependence (*31*).

Introduction of Free Energy into the Eyring Equation. The denaturant concentration can be substituted by ln K or ΔG (*4*, *5*, *11*) as the dependent variable for protein folding. The correlation shown in Figure 7 indicates a linear dependence of ln k on ln K. This observation can be verified by calculating [GdnHCl] from ΔG (applying eqs 1−4) and using this concentration in the analogous expression for $\Delta G^{\ddagger}$:

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - \theta_m \Delta H - T(\Delta S^{\ddagger} - \theta_m \Delta S) + (\Delta C_p{}^{\ddagger} - \theta_m \Delta C_p)\left(T - T_0 - T \ln \frac{T}{T_0}\right) + \theta_m \Delta G \tag{10}$$

where θ_m is $m_{U\rightarrow\ddagger}/m_{U\rightarrow F}$ (the ratio of the denaturant dependence of $\Delta G^{\ddagger}$ and ΔG according to eqs 1−4).

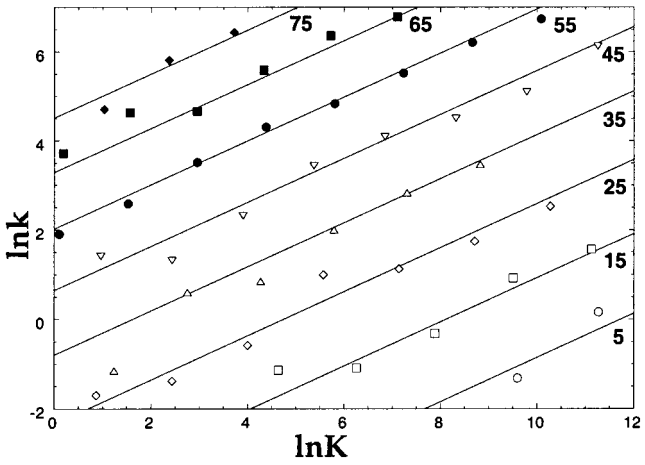

FIGURE 7: Correlation between the folding rate (ln k) and the equilibrium constant (ln K). The symbols show the folding rates, taken from every second temperature series, of the measurements compiled in Table 2. The labels indicate the approximate temperatures in degrees centigrade. The solid lines represent folding rates as predicted by eqs 7 and 11−12 (Table 3).

It has been shown that if ln K is kept constant a linear Arrhenius plot is obtained (*1*, *2*). This indicates that $\Delta C_p{}^{\ddagger} - \theta_m \Delta C_p$ is 0, which is correct within experimental error. By eliminating this term from eq 10, the following expression is obtained:

$$\ln k = \ln D + \frac{\Delta S^{\ddagger}_{\Delta G}}{R} - \frac{\Delta H^{\ddagger}_{\Delta G}}{RT} - \theta_m \frac{\Delta G}{RT} = \ln A - \frac{E_a}{RT} - \theta_m \frac{\Delta G}{RT} \tag{11}$$

where $\Delta S^{\ddagger}_{\Delta G}$ is $\Delta S^{\ddagger} - \theta_m \Delta S$, $\Delta H^{\ddagger}_{\Delta G}$ is $\Delta H^{\ddagger} - \theta_m \Delta H$, A is the frequency factor, and E_a is the activation energy, all with $\Delta G = 0$ as reference point. Note that normally the reference point for the thermodynamic parameters is zero denaturant concentration (*1*−*3*).

$\Delta G/RT$ can be further substituted by $-\ln K$, yielding

$$\ln k = \ln D + \frac{\Delta S^{\ddagger}_{\Delta G}}{R} - \frac{\Delta H^{\ddagger}_{\Delta G}}{RT} + \theta_m \ln K = \ln A - \frac{E_a}{RT} + \theta_m \ln K \tag{12}$$

This equation states that if the Eyring and Arrhenius plots are made for different ln K's then the slopes will be constant and the intercepts will vary. This yields a model for which the increase in "driving force" (ln K) leads to an increase in the *apparent* entropy of activation or, in the case of the Arrhenius equation, to an increase in the *apparent* frequency factor. For eq 11, on the other hand, the Eyring and Arrhenius plots at different ΔG have different slopes, but the same intercept. This results in a model in which the *apparent* entropy of activation is independent of the driving force and the height of the activation barrier is reduced by the free energy scaled by θ_m (or, in the case of the Arrhenius equation, the *apparent* activation energy is reduced similarly). This approach is more appealing because it consists solely of energy terms. These relationships for eqs 11 and 12 are illustrated in Figure 5. The slight curvature for the calculated points is the result of an imperfect cancellation of $\Delta C_p{}^{\ddagger} -$

Table 3: Eyring Activation Parameters for the Gibbs Free Energy Dependence Test[a]

parameter	units	cyt *c* Fe^{2+} $D = k_bT/h$	cyt *c* Fe^{2+} $D = 10^{10}$	cyt *c* Fe^{2+} $D = cT/\eta$	protein L[A] $D = k_bT/h$	protein L9[B] $D = k_bT/h$	CspB[C] $D = k_bT/h$
$\Delta H^{\ddagger}_G$	kJ/mol	117 ± 5	120 ± 5	102 ± 5	90	73	90
$\Delta S^{\ddagger}_G$	J mol^{-1} K^{-1}	127 ± 15	189 ± 16	160 ± 15	35	17	81
θ_m	10^{-3}	494 ± 35	494 ± 35	474 ± 35	578	580	937

[a] The sum of squared differences between measured and calculated rates is about 7.5% greater than the corresponding fit in Table 1. This is, however, still lower than the 12.2% used to calculate the 95% confidence limit of Table 1. [A]Ref (*1*). [B]Ref (*2*). [C]Ref (*3*).

$\theta_m\Delta C_p$, probably caused by difficulties in determining the stability of the protein at high temperature, owing to the diminishing tryptophan fluorescence.

To check the validity of the approach described by eqs 11 and 12, data for ΔG and $\Delta G^{\ddagger}$ as a function of temperature and denaturant concentration were needed for other proteins. In the literature such data were found for three proteins: protein L (*1*), protein L9 (*2*), and CspB (*3*). For protein L and protein L9, ΔG and $\Delta G^{\ddagger}$ had been described by eqs 1−4. For CspB, however, ΔH, ΔS, and ΔC_p were all denaturant concentration dependent and, moreover, urea had been used instead of GdnHCl. From the thermodynamic parameters of these three proteins, the folding rate and the free energy of folding were calculated for a temperature−denaturant data grid similar to the experimental grid used for cyt *c*. These data sets were then fitted according to eq 11. The results are given in Table 3. It should be noted that these results are based on fits of other fits and no errors of the parameters can be calculated in a normal sense. The maximum divergence of the folding rate between the published and the recalculated fits is a factor of 1.40, 1.00, and 1.13 for protein L, protein L9, and CspB, respectively. This means, that for protein L9, eq 11 gives an equally good fit as eqs 1−5. For protein L and CspB, there is a small discrepancy between the predicted folding rates for the model described by eq 11 and the corresponding published models (*1, 3*). A definite endorsement in these two latter cases is not possible without access to original experimental data. Nevertheless, the results appear promising and indicate that eqs 11 and 12 might be generally applicable to protein folding problems.

This new approach (eq 11) has both disadvantages and advantages compared to eqs 1−5. A disadvantage is that protein stability data are needed to interpret the kinetics data, but since these are usually collected anyway, this is not a major problem. On the positive side is that the expression becomes simpler. ΔC_p only shows up in the stability equations and does not ruin the fits of the kinetics by coupling the errors for $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ with each other. Only about one-third as many data points are needed to determine $\Delta H^{\ddagger}_{\Delta G}$ and $\Delta S^{\ddagger}_{\Delta G}$ with the same precision as $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ from eqs 1−4. Furthermore, because the reference point for $\Delta H^{\ddagger}_{\Delta G}$ and $\Delta S^{\ddagger}_{\Delta G}$ is zero free energy ($K = 1$), instead of zero denaturant concentration, the comparison between different proteins becomes protein-related, with the protein itself, and not the surrounding system, as the reference point. This means that if the folding rates of two proteins are compared differences in protein stability have been factored out, because the driving force for the two reactions is the same. If, instead, the comparison is made at a fixed denaturant concentration, the two proteins will, most likely, have different stability, and the reason for their relative folding rates becomes obscured due to different driving forces. A further benefit of the new approach is the possibility of comparing data from studies that have used different denaturants, such as GdnHCl and urea.

In comparing the folding of oxidized and reduced cyt *c*, it is worth noting that the curvature in the GdnHCl dependence shown in Figure 6 also is observed for the oxidized form of the protein (*6, 7*). The oxidized protein folds more than 10 times faster at zero driving force under comparable circumstances (pH 6.5 instead of 7.0), but has a slightly lower driving force dependence (*6*). Roder (*38*) and Englander (*7*) have shown that the folding mechanism for oxidized cyt *c* is more complex than a simple two-state model. If the reduced form of the protein behaves in the same way, then θ_m cannot be used to determine the position of the transition state along the reaction coordinate. The temperature dependence data in Table 1 confirm that a more complex model is needed. The positive $\Delta S^{\ddagger}$ indicates that the protein gets less ordered as it reaches the transition state. For a simple two-state mechanism, the activated complex should be more ordered than the unfolded protein and hence have a negative $\Delta S^{\ddagger}$ (*11*). The experimentally derived positive $\Delta S^{\ddagger}$ thus supports a mechanism in which reduced cyt *c* first forms a compact misfolded form which then expands to form the transition state of the rate-limiting folding reaction. The α-helix content of this misfolded intermediate is about 20% of the folded form (*26*). Of the three other proteins listed in Table 3, protein L9 also has a positive activation entropy at zero denaturant concentration, indicating that despite an *apparent* two-state behavior its mechanism might be more complex. At zero driving force, $\Delta S^{\ddagger}_{\Delta G}$ is greater than 0 for all four proteins (Table 3), emphasizing the importance of thorough investigations of early events of protein folding.

ACKNOWLEDGMENT

Thanks are extended to Drs. Harry Gray, Villy Sundström, and Jay Winkler for stimulating discussions and for critical comments on the manuscript.

REFERENCES

1. Scalley, M. L., and Baker, D. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 10636−10640.
2. Kuhlman, B., Luisi, D. L., Evans, P. A., and Raleigh, D. P. (1998) *J. Mol. Biol. 284*, 1661−1670.
3. Schindler, T., and Schmid, F. X. (1996) *Biochemistry 35*, 16833−16842.
4. Pascher, T., Chesick, J. P., Winkler, J. R., and Gray, H. B. (1996) *Science 271*, 1558−1560.
5. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R., and Gray, H. B. (1996) *Chem. Biol. 3*, 491−497.
6. Ikai, A., Fish, W. W., and Tanford, C. (1973) *J. Mol. Biol. 73*, 165−184.

7. Englander, S. W., Sosnick, T. R., Mayne, L. C., Shtilerman, M., Qi, P. X., and Bai, Y. (1998) *Acc. Chem. Res. 31*, 737−744.
8. Matouschek, A., Kellis, J. T., Serrano, L., Bycroft, M., and Fersht, A. R. (1990) *Nature 346*, 440−445.
9. Parker, M. J., Spencer, J., and Clarke, A. R. (1995) *J. Mol. Biol. 253*, 771−786.
10. Silow, M., and Oliveberg, M. (1997) *Biochemistry 36*, 7633−7637.
11. Onuchic, J. N., Socci, N. D., Luthey-Schulten, Z., and Wolynes, P. G. (1996) *Folding Des. 1*, 441−450.
12. Elwell, M. L., and Schellman, J. A. (1977) *Biochim. Biophys. Acta 494*, 367−383.
13. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
14. Tan, Y.-J., Oliveberg, M., and Fersht, A. R. (1996) *J. Mol. Biol. 264*, 377−389.
15. Socci, N. D., Onuchic, J. N., and Wolynes, P. G. (1996) *J. Chem. Phys. 104*, 5860−5868.
16. Eyring, H. (1935) *J. Chem. Phys. 3*, 107−115.
17. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) *Proteins: Struct., Funct., Genet. 21*, 167−195.
18. Telford, J. R., Tezcan, F. A., Gray, H. B., and Winkler, J. R. (1999) *Biochemistry 38*, 1944−1949.
19. Stowell, M. H. B., Larsen, R. W., Winkler, J. R., Rees, D. C., and Chan, S. I. (1993) *J. Phys. Chem. 97*, 3054−3057.
20. Lindqvist, L., Czochralska, B., and Grigorov, I. (1985) *Chem. Phys. Lett. 119*, 494−498.
21. Orii, Y. (1993) *Biochemistry 32*, 11910−11914.
22. Beechem, J. M. (1992) *Methods Enzymol. 210*, 37−54.
23. Johnson, M. L., and Faunt, L. M. (1992) *Methods Enzymol. 210*, 1−37.
24. Churg, A. K., and Warshel, A. (1986) *Biochemistry 25*, 1675−1681.
25. Bixler, J., Bakker, G., and McLendon, G. (1992) *J. Am. Chem. Soc. 114*, 6938−6939.
26. Chen, E., Wittung-Stafshede, P., and Kliger, D. S. (1999) *J. Am. Chem. Soc. 121*, 3811−3817.
27. Jones, C. M., Henry, E. R., Hu, Y., Chan, C.-K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11860−11864.
28. Chan, H. S., and Dill, K. A. (1998) *Proteins: Struct., Funct., Genet. 30*, 2−33.
29. Schwarz, G. (1965) *J. Mol. Biol. 11*, 64−77.
30. Kramers, H. A. (1940) *Physica 7*, 284−304.
31. Jacob, M., Schindler, T., Balbach, J., and Schmid, F. X. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5622−5627.
32. Plaxco, K. W., and Baker D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 13591−13596.
33. Thompson, P. A., Eaton, W. A., and Hofrichter, J. (1997) *Biochemistry 36*, 9200−9210.
34. Weast, R. C. (1985) *CRC Handbook of Chemistry and Physics*, 65th ed., p F-37, CRC Press, Boca Raton, FL.
35. Banerjee, R., Frilley, B., and Guissani, A. (1999) *Indian J. Biochem. Biophys. 36*, 107−117.
36. Falkenhagen, H., and Vernon E. L. (1932) *Phys. Z. 33*, 140.
37. Kawahara, K., and Tanford, C. (1966) *J. Biol. Chem. 241*, 3228−3232.
38. Shastry, M. C. R., Sauder, J. M., and Roder H. (1998) *Acc. Chem. Res. 31*, 717−725.

BI0026223