

Rate of Intrachain Contact Formation in an Unfolded Protein: Temperature and Denaturant Effects

Stephen J. Hagen*, Caleb W. Carswell and Erik M. Sjolander

Physics Department, University
of Florida, PO Box 118440
Gainesville, FL 32611, USA

We have measured the effect of temperature and denaturant concentration on the rate of intrachain diffusion in an unfolded protein. After photodissociating a ligand from the heme iron of unfolded horse cytochrome *c*, we use transient optical absorption spectroscopy to measure the time scale of the diffusive motions that bring the heme, located at His18, into contact with its native ligand, Met80. Measuring the rate at which this 62 residue intrachain loop forms under both folding and unfolding conditions, we find a significant effect of denaturant on the chain dynamics. The diffusion of the chain accelerates as denaturant concentration decreases, with the contact formation rate approaching a value near $\sim 6 \times 10^5 \text{ s}^{-1}$ in the absence of denaturant. This result agrees well with an extrapolation from recent loop formation measurements in short synthetic peptides. The temperature dependence of the rate of contact formation indicates an Arrhenius activation barrier, $E_a \sim 20 \text{ kJ/mol}$, at high denaturant concentrations, comparable to what is expected from solvent viscosity effects alone. Although E_a increases by several $k_B T$ as denaturant concentration decreases, the overall rate of diffusion nevertheless increases. These results indicate that inter-residue energetic interactions do not control conformational diffusion in unfolded states, even under folding conditions.

© 2001 Academic Press

Keywords: protein folding; cytochrome *c*; energy landscape; denatured state; kinetics

*Corresponding author

Introduction

New spectroscopic techniques for studying protein folding have clearly demonstrated that the most rapid stages of *in vitro* folding occur on microsecond, or even nanosecond, time-scales.¹ The observation of such rapid dynamics, and the finding that some proteins can actually attain a native structure in less than a millisecond, raises the question of what ultimately limits the speed of protein folding. The rate of intramolecular diffusion defines one important time-scale: if a polypeptide can only fold as fast as it can form its shortest internal loops, then the rate of intrachain contact formation in a randomly coiled molecule can present an upper limit to the speed of protein folding.²

A study of unfolded cytochrome *c* provided the first estimate of this diffusional "speed limit" for protein folding:^{2,3} the formation of a 62-residue intrachain loop that joins the heme iron (at His18) of horse cytochrome *c* with its native ligand, Met80, can be triggered and detected by optical spectroscopy.⁴ Under unfolding conditions, diffusion of the randomly coiled polypeptide brings Met80 into contact with the heme iron at a rate $\sim (2.5\text{--}2.9) \times 10^4 \text{ s}^{-1}$. Based on this result, a simple scaling model then suggested that the shortest loops commonly found in protein structures ($n = 6\text{--}10$ residues) should form at a rate no greater than $\sim (0.3\text{--}1) \times 10^6 \text{ s}^{-1}$; thus intrachain diffusion should not allow a protein to fold in less than about 1 μs .

This prediction raises the question of how the rate of contact formation between distinct sites on an unfolded chain depends on the length of the intervening segment of chain, its amino acid composition, and the solvent conditions, including temperature and denaturant concentration. Two recent studies used the transfer and quenching of triplet-state excitation to measure the rate of end-to-end

Abbreviations used: GdnHCl, guanidine hydrochloride; SVD, singular value decomposition; k_B , Boltzmann's constant.

E-mail address of the corresponding author:
sjhagen@ufl.edu

contact formation in short synthetic peptides of varying length.^{5,6} In both experiments, glycine-rich chains consisting of about ten peptide bonds formed end-to-end loops at rates $>1 \times 10^7 \text{ s}^{-1}$. This rate significantly exceeds the prediction from the cytochrome *c* result, and probably indicates at the very least that chain composition will strongly influence the upper limit for protein folding speed.

Previous studies have not, however, investigated the role of denaturant and temperature on the rate of loop formation. Since the cytochrome *c* study was conducted under unfolding conditions (4.6–5.6 M GdnHCl, pH 6.5, 40 °C), one might anticipate that lower temperatures or denaturant concentrations may enhance the influence of intrachain interactions, with consequences for the diffusional rate of the polypeptide chain. A current view of protein folding holds that solvent-averaged, non-native intrachain interactions, e.g. hydrogen bonds, hydrophobic interactions, etc., in unfolded states create “roughness” in the free energy surface that may impede the conformational diffusion of a molecule toward its native configuration. This may also give rise to non-Arrhenius temperature-dependence in the folding kinetics⁷ and in conformational diffusion.⁸ Although computer simulations clearly indicate that such interactions should have strong effects on the dynamics of folding, relatively few experiments have provided much insight into the magnitude of roughness effects in real proteins.⁹ One may therefore ask whether the removal of denaturant and the restoration of folding conditions in fact slows the intrachain diffusion of cytochrome *c*, and whether it leads to any qualitative changes in the temperature dependence of this diffusion.

This work examines the effect of denaturant and temperature on the rate of intrachain loop formation in cytochrome *c*. In the native state, the heme iron of horse cytochrome *c* has two axial ligands, the imidazole group of His18 and the sulfur ion of Met80.¹⁰ Met80 dissociates from the heme in the denatured protein, but the heme remains covalently attached to the polypeptide chain at His18 and through thioether links at Cys14 and Cys17. In its Fe(II) state, the heme iron of unfolded cytochrome *c* can reversibly bind exogenous ligands such as CO in place of Met80. It can also bind intrachain ligands such as Met80, Met65, His26, or His33. If CO is bound to the iron, photodissociation of CO then allows transient binding of the intrachain ligands to the heme, prior to the eventual rebinding of CO. But while binding of His26 and His33 to the heme is rate-limited by the relatively slow formation of the covalent link, heme-methionine binding is nearly diffusion-limited;^{2,3} thus, the rate of Met80 or Met65 loop formation primarily reflects the rate of intrachain diffusion of the polypeptide, rather than the rate of heme-ligand chemistry. Using a nanosecond laser to photodissociate the CO ligand from the heme iron, we can then collect transient optical absorption spectra in the heme Soret region and detect

the formation of the heme-methionine bond.⁴ From the observed rate, and from a knowledge of the intrinsic heme-methionine binding rates, we then calculate the rate of the intramolecular motions that bring Met80 into proximity with the heme iron.

With this technique, we can measure the rate of heme-Met80 contact formation in the unfolded protein, even under solution conditions that correspond to a thermodynamically stable native state. In the absence of CO, Fe(II) cytochrome *c* unfolds at $\sim 5 \text{ M GdnHCl}$, at 40 °C and neutral pH; at $\sim 4 \text{ M GdnHCl}$ at 40 °C, fewer than 1% of cytochrome *c* molecules would remain unfolded at equilibrium in the absence of CO. However, the CO ligand destabilizes the folded state, shifting the unfolding transition to $\sim 3.5 \text{ M GdnHCl}$.^{4,11,12} Photodissociation of CO from unfolded cytochrome *c* at GdnHCl concentrations near 4–5 M therefore initiates intrachain loop formation phenomena on microsecond time-scales, and would lead also to folding on millisecond time-scales,^{4,11,13,14} if CO recombination did not occur first.

Results

Kinetic phases

Figure 1 shows typical transient optical absorption spectra $A(\lambda, t)$ measured after a CO ligand is photodissociated from the heme iron of cytochrome *c* unfolded in guanidine hydrochloride (GdnHCl). Immediately after photodissociation ($t \sim 10 \text{ ns}$), the sample exhibits the high-spin ferrous deoxyheme Soret spectrum, with a broad low peak near 425 nm. At long times ($t \sim 1 \text{ ms}$), CO recombination with the heme produces the low-spin carbonmonoxyheme spectrum with its narrow peak near 413 nm. The transient spectra at intermediate times appear to exhibit two nearly isosbestic points in sequence, first near 424 nm and then near 419 nm; this suggests that about three spectroscopically distinct species account for most of the observed spectral changes. More detailed analysis by singular value decomposition (SVD; see Materials and Methods) in fact shows that $A(\lambda, t)$ contains four independent spectra: each set of absorption difference spectra, $\Delta A(\lambda, t) = A(\lambda, t) - A(\lambda, 0^-)$, contains three independent spectral components, each with its own time dependence.

We first analyzed the Soret spectral changes by reconstructing each set of difference spectra as a series of four exponential relaxations: $\Delta A(\lambda, t) = \sum_{i=1-4} B_i(\lambda) \exp(-t/\tau_i)$ (see Materials and Methods). Figure 2 shows the time constants τ_1 , τ_2 , and τ_3 that result from such a fit ($\tau_4 \geq 10 \times \tau_3$ is not shown); the associated $B_i(\lambda)$ appear in previously published studies of CO-cytochrome *c* photodissociation.¹⁵ The first relaxation, at $\tau_1 \sim 10^{-6} - 10^{-5}$ second, corresponds to the appearance of a weak maximum near 420 nm in

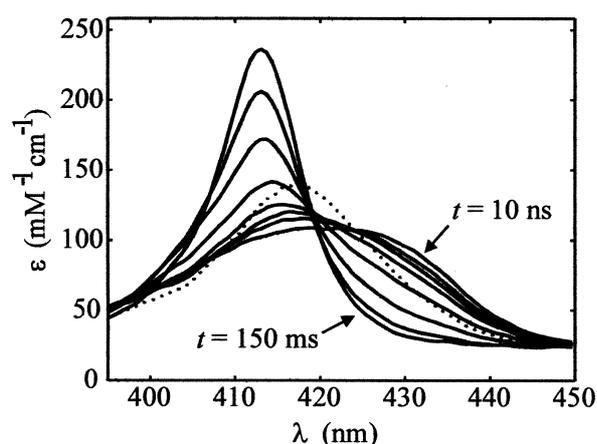


Figure 1. Transient absorption spectra of horse Fe(II) cytochrome *c* unfolded in 4.4 M GdnHCl (pH 6.0) at 40.3 °C, at various time delays t after photodissociation of the CO heme ligand. Continuous lines: spectra at times $t = 10$ ns, 3.1 μ s, 9.4 μ s, 32 μ s, 106 μ s, 350 μ s, 1.0 ms, and 150 ms. Broken curve: spectrum of the heme-methionine complex under these conditions, as estimated in the fit to the kinetic model of Scheme I (see Materials and Methods).

$A(\lambda, t)$. Previous work strongly indicates that this phase corresponds to the transient binding of intrachain methionine residues (Met80 and Met65) to the heme iron, since (1) the transient spectrum developing at τ_1 closely resembles the Soret spectrum of native cytochrome *c*, in which Met80 forms the heme axial ligand on the distal side;⁴ (2) the τ_1 relaxation was not observed in cytochrome *c* samples in which Met65 and Met80 were carboxymethylated;³ and (3) time-resolved magnetic circular dichroism studies, along with equilibrium studies of model systems, appear most consistent with a methionine ligand at the heme during this phase.^{15,16} Based on this evidence, we interpret τ_1 as heme-methionine intrachain loop formation in the unfolded protein.

A second, weaker relaxation occurs in the spectra on a time-scale $\tau_2 \sim 10^{-4} - 10^{-3}$ second; previous authors attributed this phase to the transient binding of intrachain His26 or His33 to the heme.^{4,15} Mutation studies further suggest that His33, rather than His26, is primarily responsible.¹⁷ The heme-histidine loops form more slowly than the heme-methionine loops, despite their shorter length, because of intrinsically slower heme-histidine reaction chemistry.³ The third and fourth exponential phases (τ_3 and τ_4) coincide with the recovery of the equilibrium carbonmonoxyheme spectrum. Chen *et al.*¹⁵ found that both τ_3 and τ_4

vary with CO partial pressure, indicating that both phases arise from the recombination of CO. For example, τ_3 may correspond to bimolecular recombination of CO with the 5-coordinate heme, while τ_4 may correspond to the displacement of Met80 by CO in molecules that progressed to the native state after photodissociation.

Extraction of loop formation rates

The τ_1 data of Figure 2 show that the methionine-binding process accelerates roughly fivefold over a temperature range of 15–45 °C, and also accelerates with decreasing denaturant concentration. However, the optical absorption spectrum at $t \sim \tau_1$ (i.e. at $t \sim 10^{-5} - 10^{-4}$ second) varies little with experimental conditions, indicating that the overall degree of methionine binding (after photodissociation of CO) does not vary greatly with temperature or denaturant. Thus, the data initially suggest that temperature and denaturant affect both the rate of formation (k_{M80}) and the rate of dissociation (k_{-M}) of the heme-Met80 link, but only weakly affect the association constant, k_{M80}/k_{-M} .

We used the kinetic model of Scheme I (Figure 3) to analyze the spectroscopic data more quantitatively. Photodissociation of the iron-CO bond at $t = 0$ creates the 5-coordinate deoxyheme state. The heme iron can then bind transiently to Met80 and Met65, as diffusion of the polypeptide chain brings these into proximity. Both methionine residues dissociate from the heme iron at a rate k_{-M} , but Scheme I provides for different rates, k_{M80} and k_{M65} , for the binding of Met80 and Met65 to the heme. We cannot expect to distinguish between the Met80- and Met65-coordinated hemes in optical absorption spectra; the spectra can reveal only a single overall rate ($1/\tau_1$) and equilibrium constant for the binding of intrachain methionine to the heme under a given set of experimental conditions. In order to separately obtain k_{M80} and k_{M65} , our model therefore fixes the ratio $k_{M80}/k_{M65} = 0.66$. This choice of the ratio assumes slightly faster binding of Met65 than Met80, and arises from a simple Gaussian chain model of loop formation (see Discussion). With this ratio fixed, the analysis of τ_1 and the observed spectra according to Scheme I yields separate values for k_{M80} , k_{M65} , and k_{-M} (see Materials and Methods). In what follows we discuss k_{M80} , although similar conclusions will apply to k_{M65} .

Like the kinetic model described by Jones *et al.*,⁴ Scheme I also allows an intrachain heme-histidine complex to form at a rate k_H and dissociate at a rate k_{-H} , and bimolecular CO recombination occurs at a rate k_{CO} . Finally, Scheme I includes only the rapid intrachain dynamics following photolysis; it does not include a route to the folded state on longer time-scales: although the native fold is thermodynamically stable under some of our experimental conditions, we expect little if any folding on loop-formation time-scales ($\sim 10^{-6} - 10^{-5}$ second)^{13,14†}

† We note that Goldbeck *et al.*¹⁶ have argued that a heterogeneous description more accurately reproduces certain features of the heme-His transient MCD spectrum. For modeling heme-methionine contacts, we consider Scheme I adequate.

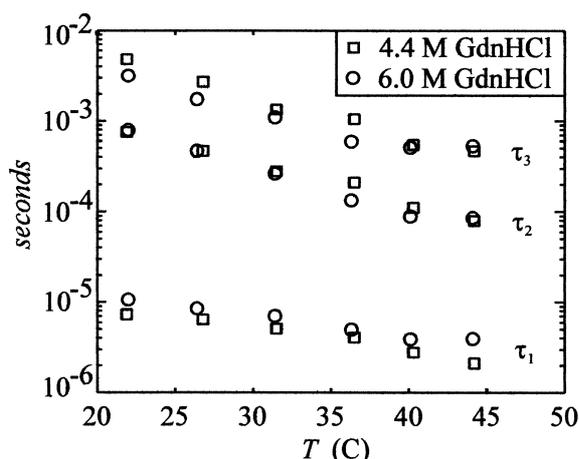


Figure 2. Relaxation times τ_1 , τ_2 , and τ_3 obtained from multi-exponential fit to the time-course of transient optical absorption changes of unfolded cytochrome *c* after photodissociation of heme-CO bond. Intrachain heme-methionine (Met65 and Met80) loops form on the time-scale τ_1 , while τ_2 and τ_3 characterize formation of intrachain heme-histidine loops and bimolecular CO recombination, respectively.

The kinetic equations for Scheme I generate about three sequential exponential relaxations, in good agreement with the spectroscopic data between $t \approx 10$ ns and $t \sim 1$ ms (i.e. $t \ll \tau_4$). Fitting each set of transient spectra, representing a fixed temperature and denaturant concentration, to Scheme I (see Materials and Methods), we find that k_{M80} and k_{-M} both increase as temperature (T) increases or as denaturant concentration decreases. Figure 4 shows k_{M80} and the ratio k_{M80}/k_{-M} and supports the initial conclusion that the loop formation probability varies little with experimental conditions. Although k_{M80}/k_{-M} decreases with increasing T in 4.4 M GdnHCl (indicating a lower probability of heme-methionine contact at higher T), the data at higher denaturant concentrations reveal very little temperature or denaturant-dependence in k_{M80}/k_{-M} .

Rate of chain diffusion

Formation of the heme-methionine loop requires diffusion of the polypeptide chain, followed by the formation of an iron-sulfur covalent bond. Therefore the rate of intrachain diffusion must exceed, to some extent, the rate of loop formation, k_{M80} , obtained above. Here we account for the speed of the heme-methionine chemistry and calculate the chain diffusion rate from the measured loop formation rate. We adopt the approach of Wang & Davidson,^{18,19} who studied an analogous loop-formation process, the cyclization of λ DNA. The formation of the intramolecular link in the polypeptide is described as a two-step process:

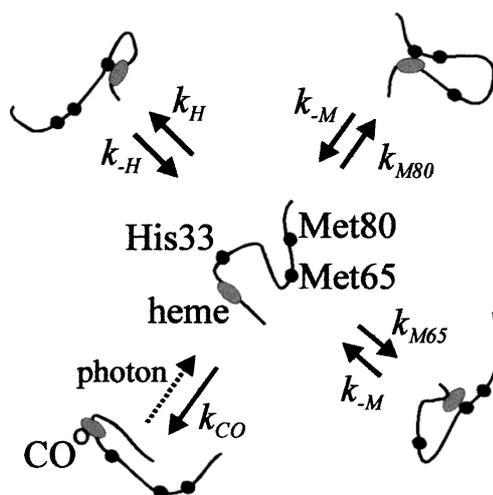


Figure 3. Scheme I, summarizing the kinetic model for the formation of intrachain loops in unfolded cytochrome *c* after photolysis of the heme-CO bond. Photolysis leaves the heme iron (located at His18) in the five-coordinate state, which allows the transient binding of Met80, Met65, or a histidine residue (presumably His33). Bimolecular CO recombination restores the initial state at longer times.



Rate k_1 characterizes the diffusive motion of the randomly coiled polypeptide that leads to an intermediate encounter complex, a configuration of the chain in which Met80 makes contact with the heme iron but has not yet bound. The encounter complex can proceed to form the covalently linked complex, at a rate k_3 , or else the loop may diffuse apart, at a rate k_2 . The covalent link breaks spontaneously at a rate k_4 . The steady-state approximation then gives the overall rates of Met80 binding and dissociation:

$$k_{M80} = k_1 k_3 / (k_2 + k_3) \quad (1)$$

$$k_{-M} = k_4 k_2 / (k_2 + k_3)$$

We can then find the rates k_1 and k_2 , describing diffusion of the polypeptide chain, from the measured rates k_{M80} and k_{-M} :

$$k_1 = k_4 k_{M80} / (k_4 - k_{-M}) \quad (2)$$

$$k_2 = k_3 k_{-M} / (k_4 - k_{-M})$$

Since k_3 and k_4 describe only the rate of heme-methionine chemistry, independent of the process that brings the two reactants into proximity, we can estimate their values by measuring the rate at which free methionine binds from solution to a heme-containing fragment of cytochrome *c*.^{2,3} This analysis (see Materials and Methods) gives k_3 in excess of $5 \times 10^{10} \text{ s}^{-1}$ at all T studied, and

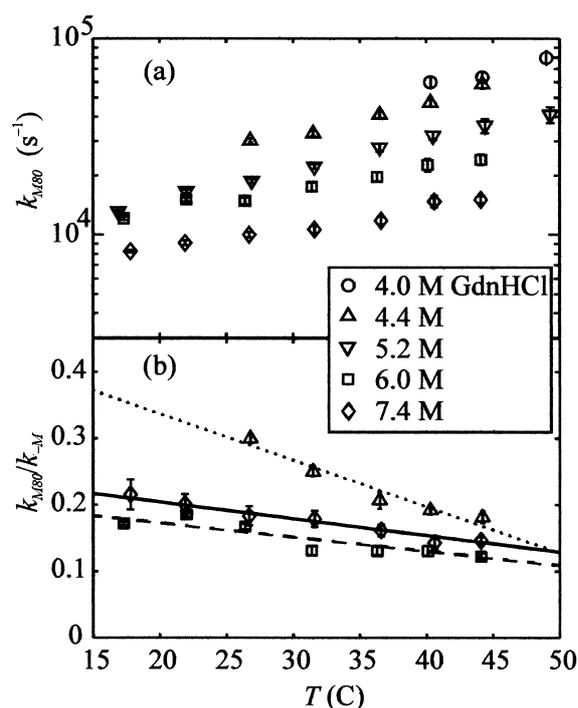


Figure 4. (a) Rate k_{M80} of heme-Met80 binding versus temperature and denaturant, obtained from fit of transient spectra to the kinetic model of Scheme 1. Error bars indicate 1σ uncertainty, as calculated in a Monte Carlo histogram analysis of the data and fit. (b) Ratio k_{M80}/k_{-M} versus temperature, indicating the weak temperature dependence of the heme-Met80 loop formation probability at high denaturant concentrations. The lines represent linear fits to the data for 4.4 M, 6.0 M, and 7.4 M GdnHCl.

$k_3/k_4 \approx 1 \times 10^5$. We insert k_3 , k_4 , k_{M80} , and k_{-M} into equation (2) and obtain k_1 and k_2 , the rates of the diffusional motion of the polypeptide chain that brings Met80 into and out of close proximity with the heme in the unfolded molecule. Figures 4 and 5 show generally similar magnitudes for k_1 and k_{M80} , essentially because the covalent binding step (k_3) occurs so rapidly that the rate of chain diffusion k_1 controls the loop formation rate k_{M80} .^{2,3}

Discussion

Equilibrium constant for loop formation

We first compare our measured equilibrium constant for loop formation, k_{M80}/k_{-M} , with that expected if the unfolded cytochrome *c* adopts a random coil configuration. If the polypeptide resembles a Gaussian random coil, the encounter complex between the heme and Met80 forms with a probability:^{18–20}

$$k_1/k_2 = (3/2\pi)^{3/2} v_s/L^3$$

Here, $L = \langle r^2 \rangle^{1/2}$ is the rms straight-line distance

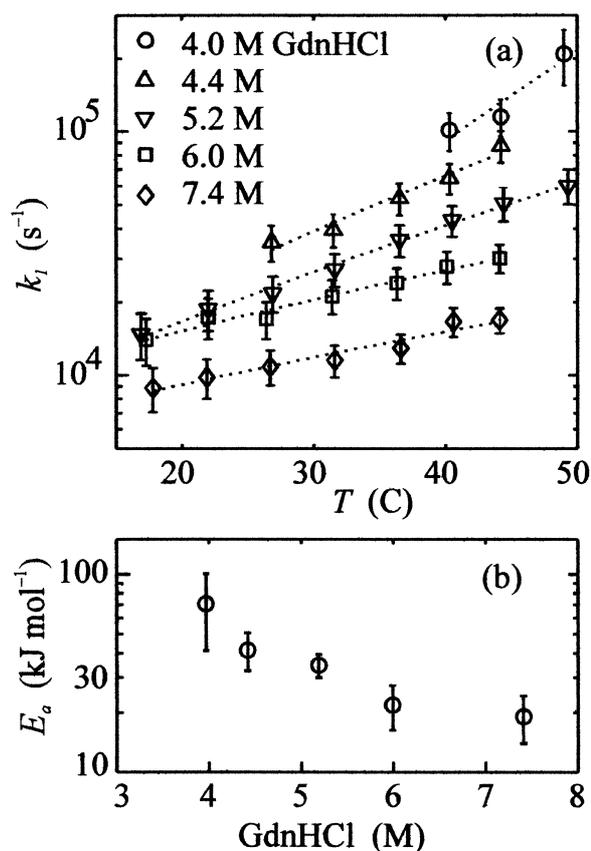


Figure 5. (a) Rate k_1 of the intrachain diffusion that brings heme at His18 into contact with Met80, calculated from the data of Figure 4 and equation (2) (see Materials and Methods). Broken curves indicate fit to simple Arrhenius relation, $k_1 \propto \exp(-E_a/k_B T)$. (b) Activation energies E_a versus denaturant, from Arrhenius curves in (a).

between His18 and Met80, and the encounter complex is defined as any configuration of the chain that brings the heme iron and its ligand into the same small volume v_s . Equation (1) gives the loop formation probability k_{M80}/k_{-M} :

$$k_{M80}/k_{-M} = k_1 k_3 / k_2 k_4 = (k_3/k_4) (3/2\pi)^{3/2} v_s / L^3 \quad (3)$$

We can estimate $(k_3/k_4)v_s$ from the known equilibrium constant for the bimolecular binding of heme and methionine. That equilibrium constant is the bimolecular analog of equation (3):

$$K_{\text{bimolecular}} = (k_1/k_2)_{\text{bimolecular}} (k_3/k_4) = v_s (k_3/k_4)$$

Therefore $k_{M80}/k_{-M} = (3/2\pi)^{3/2} K_{\text{bimolecular}} / L^3$. For free methionine binding to an Fe(II)-heme peptide, Tezcan *et al.*²¹ report $K_{\text{bimolecular}} \approx (2.4 \text{ mM})^{-1}$, i.e. $v_s (k_3/k_4) \approx 6.9 \times 10^{-19} \text{ cm}^3$, at $\approx 20^\circ \text{C}$. We can estimate L for an idealized polypeptide chain of typical amino acid composition unfolded in GdnHCl. Early studies of proteins unfolded in high concentrations of GdnHCl established that the chain dimensions approximate those expected for ran-

dom coils: the mean-squared separation between endpoints of a chain of n amino acid residues is approximately $L^2 \approx n (0.95 - 1.15 \text{ nm})^2$, but varies with chain composition, solution conditions, etc.²² For the $n = 62$ residue segment between Met80 and the heme at His18, we then expect $k_{M80}/k_{-M} \approx 0.3 - 0.55$. By comparison, Figure 4 shows $k_{M80}/k_{-M} \approx 0.2 - 0.35$ at 22 °C, in GdnHCl concentrations of 4.4-7.4 M. Thus the data agree roughly with a random coil configuration of the unfolded protein: they fall below the predicted value, but the prediction employs the most idealized description of the unfolded chain. The unperturbed Gaussian random coil model ignores the excluded volume of the chain as well as any differences between the statistical mechanics of internal versus end-to-end loops. The generally weak temperature dependence of k_{M80}/k_{-M} (Figure 4), which falls by 1-2 % per degree C in 4.4 M GdnHCl, suggests (through equation (3)) that the dimension L of the unfolded chains increases by less than $\sim 1\%$ per degree C between 25 and 45 °C. At higher GdnHCl concentrations, k_{M80}/k_{-M} varies even less with temperature and denaturant; consistent with X-ray studies²³ showing that guanidine concentration does not significantly affect the dimensions of unfolded cytochrome *c*. Therefore, temperature and denaturant effects on the chain diffusion rate k_1 should reflect changes in the dynamics of the molecule, rather than systematic changes in its average dimensions.

The rate of chain diffusion

Figure 6 compares our results for k_1 with those obtained in recent studies on short peptides. Several authors have analyzed the theoretical problem of diffusion-controlled loop formation in biopolymers.²⁴⁻²⁹ The rate k_1 is determined by the flexibility and mobility of the unfolded chain, and by the average dimensions of the unfolded molecules. Szabo *et al.*²⁶ modeled the dynamics of the end-to-end distance r of a Gaussian random coil as the three-dimensional diffusion of a particle on a harmonic energy surface $U \propto r^2$: the particle moves on this surface with a diffusion constant D , and its approach towards the origin (or more precisely the approach of r towards a value a , where $a \ll L$ and $v_s = 4\pi a^3/3$) corresponds to the formation of a contact between the two ends of the chain. The characteristic loop formation time, $1/k_1$, is the mean first passage time for the particle to reach $r = a$:

$$k_1 = 3Da(6/\pi)^{1/2}/L^3 \quad (4)$$

Equation (4) introduces a phenomenological diffusion constant D to characterize the internal motions of the chain. This D need not equal the diffusion constant of a single amino acid molecule in solution, or of the center of mass of the unfolded molecule. Since $L^2 \propto n$ for an unperturbed random coil, equation (4) predicts $k_1 \propto n^{-3/2}$. Cyclization studies

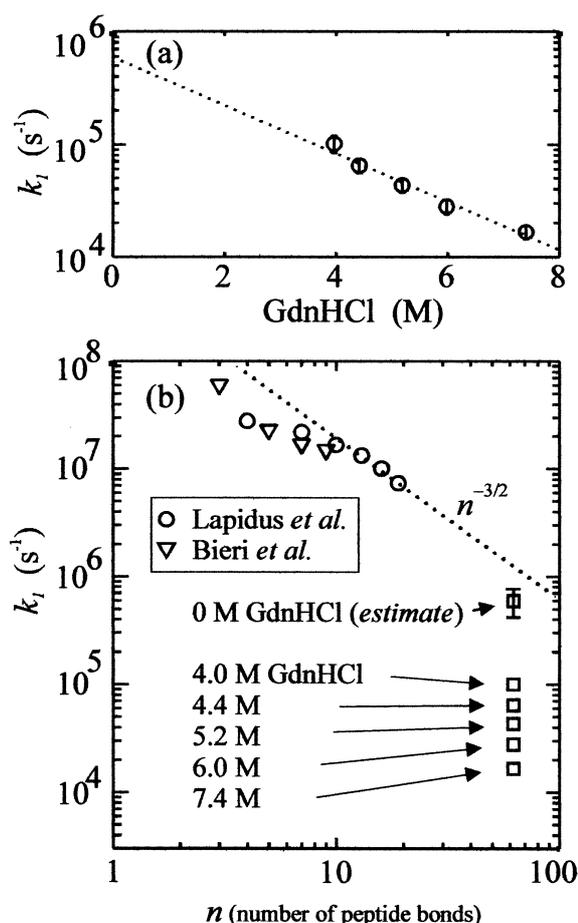


Figure 6. (a) Dependence of diffusion rate k_1 on GdnHCl concentration, at 40 °C (pH 6.0). The dotted line extrapolates (on a semilog scale) to zero denaturant, indicating $k_1 \approx k_1^0 \exp(-m[\text{GdnHCl}])$, where $m = 0.49(\pm 0.05) \text{ M}^{-1}$ and $k_1^0 = (6.0 \pm 1.9) \times 10^5 \text{ s}^{-1}$. (b) Comparison of the intrachain contact formation rates k_1 measured on polypeptides of varying length. Data of Lapidus *et al.*⁶ and Bieri *et al.*⁵ represent synthetic peptides consisting of repeating units of [Gln-Gly-Ala] and [Ser-Gly], respectively. Data at $n = 62$ indicate present results for the heme-Met80 loop of unfolded cytochrome *c* at 40 °C (pH 6.0). Also shown is k_1^0 , the extrapolation of the present data to zero denaturant.

on long synthetic polymers, such as polystyrene with $n \sim 100-2500$, generally confirm this behavior.³⁰ However, equation (4) presumably fails for the smallest values of n , as the intrinsic stiffness of the polypeptide chain will prevent loop formation;^{28,29} a recent study on short peptides appears to confirm that k_1 levels off for $n \leq \sim 10$.⁶

The dotted line in Figure 6(b) extrapolates from the $n \sim 10-20$ data of reference 6, according to $k_1 \propto n^{-3/2}$. This extrapolation predicts $k_1 \approx 1.2 \times 10^6 \text{ s}^{-1}$ for the $n = 62$ loop of cytochrome *c*: our measured values of k_1 clearly fall short of this extrapolation. However, Figure 6(a) shows

that k_1 increases at low GdnHCl concentrations (see below). If we make the (admittedly long) extrapolation from 4 M to 0 M GdnHCl, Figure 6(a) suggests a limiting value k_1 ($n = 62$) $\approx (6.0 \pm 1.9) \times 10^5 \text{ s}^{-1}$ at 40 °C, in the absence of denaturant. This value lies within a factor of 2 of the dotted line of Figure 6(b).

The amino acid composition of the polypeptide chain probably accounts for the remaining difference. While the 62 residue loop connecting His18 and Met80 of cytochrome *c* consists of 15% glycine, the short n data of Figure 6(b) was acquired on peptides of 30-50% glycine: Bieri *et al.*⁵ measured the rate of end-to-end contact formation by studying triplet-triplet energy transfer from a thioxanthone donor D to a naphthylalanine acceptor A in synthetic peptides consisting of D-[Gly-Ser] $_n$ -A; Lapidus *et al.*⁶ measured contact formation by observing the contact quenching of the tryptophan triplet state by cysteine in peptides of Cys-(Ala-Gly-Gln) $_N$ -Trp. Equation (4) indicates that k_1 varies inversely as L^3 , where L for a Gaussian chain depends on the chain flexibility through $L^2 = C n l^2$.³¹ Here $l \approx 0.38 \text{ nm}$ is the C α -C β distance and C is Flory's characteristic ratio. Thus $k_1 \propto C^{-3/2}$. C falls dramatically upon the introduction of glycine into the chain; the presence of ~33% glycine residues decreases C to ~0.5 of its value for a chain containing ~15% glycine.³¹ This implies (through equation (4)) that the cytochrome *c* loop should form $\sim(0.5)^{3/2} = 0.35$ as fast as a (Ala-Gly-Gln) $_n$ chain of equal length, or $\sim 0.35 \times (1.2 \times 10^6 \text{ s}^{-1}) = 4.2 \times 10^5 \text{ s}^{-1}$. This estimate agrees well with our extrapolation to zero denaturant (Figure 6(a)); thus we find a satisfactory agreement between the loop formation rates measured in short and long chains.

Temperature and solvent-dependence of chain diffusion

Since the folded state of Fe(II) cytochrome *c* becomes thermodynamically stable at 5 M GdnHCl^{4,11} at 40 °C, our data at lower temperatures and denaturant concentrations include conditions where the polypeptide could fold after photodissociation of the CO, if not for the recombination of CO. We can therefore ask whether the loop formation kinetics differ above and below the equilibrium folding transition: do the intrachain interactions that stabilize the native state affect the diffusion of the unfolded chain? We first estimate the conformational diffusion constant D of equation (4). From equation (4) and the discussion following equation (3), we have $D = (K_{\text{bimolecular}}/4\pi a) (k_1 k_{-M}/k_{M80})$, or $D \approx 4.2 \times 10^{-7} \text{ cm}^2 \text{ s}^{-1}$ in

7.4 M GdnHCl (at 40 °C). This rises fivefold in 4 M GdnHCl, to $D \approx 2.3 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$. Here we take $a = 0.15 \text{ nm}$ and $K_{\text{bimolecular}} = (2.4 \text{ mM})^{-1}$. (Fluorescence resonance energy transfer studies yielded comparable values for the intrachain diffusion constant in a denatured protein.³²)

Thus one may interpret the increase of k_1 at low denaturant concentration as arising from a growing diffusion constant D . One might have anticipated that lower denaturant concentration would enhance the effect of random, non-native intrachain interactions, slowing the diffusion of the chain. The contrary experimental result probably cannot be attributed to a denaturant-dependent solvent viscosity. The viscosity of aqueous guanidine solutions increases by just a factor of 2 between 0 M and 7.4 M GdnHCl at 25 °C.³³ Binding of denaturant to the unfolded chain, however, may slow its internal diffusion at high GdnHCl concentrations: denaturant molecules attached to the polypeptide chain may increase its effective hydrodynamic radius enough to slow its internal diffusion, but need not necessarily increase either the average end-to-end distance L or the molecule's center-of-mass diffusion constant.

Alternatively, the faster loop formation at low denaturant concentration may reflect a stronger thermodynamic driving force. Closing the intrachain loop presumably conceals certain hydrophobic regions from exposure to the solvent; this entropic effect accelerates the folding of many proteins at low denaturant concentrations. Folding rates commonly increase by 200-400 times as [GdnHCl] decreases from 3 M to 0 M.^{11,34,35} For Fe(II) cytochrome *c*, the overall folding rate k_f increases as $\ln(k_f) = \ln(k_f^0) - m \times [\text{GdnHCl}]$, where $m = 1.9 \text{ M}^{-1}$.¹¹ By comparison, Figure 6(a) suggests $\ln(k_1) = \ln(k_1^0) - m \times [\text{GdnHCl}]$, where $m = 0.49(\pm 0.05) \text{ M}^{-1}$ and $k_1^0 = (6.0 \pm 1.9) \times 10^5 \text{ s}^{-1}$. The 20-fold weaker dependence of k_1 (or D) on [GdnHCl], compared to k_f , would then reasonably indicate that formation of the heme-methionine loop buries a far smaller hydrophobic area than does folding of the protein.

From the temperature-dependence of k_1 (or, equivalently, D), we can estimate the magnitude of the enthalpic barriers to diffusion of the unfolded chain in the vicinity of the folding transition. An Arrhenius fit to the k_1 data of Figure 5(a), near 20-45 °C, indicates an activation energy $E_a = -\partial \ln(k_1)/\partial(1/k_B T) \approx 19(\pm 5) \text{ kJ/mol}$ at 7.4 M GdnHCl \dagger , and $E_a \approx 42(\pm 9) \text{ kJ/mol}$ at 4.4 M GdnHCl (Figure 5(b)). Thus, removing denaturant appears to increase the activation energy for conformational diffusion by ~23 kJ/mol, or ~9 $k_B T$. Since the rate of diffusion should reflect both the strength of intrachain interactions and the viscosity of the solvent, we suggest that this activation energy contains a contribution E_a^0 from the solvent viscosity and a contribution E_a^{chain} from the intrachain interactions of the polypeptide, giving a total activation energy $E_a = E_a^0 + E_a^{\text{chain}}$. The viscosity of water near room temperature can be characterized

\dagger The data do not allow us to confirm either Arrhenius or non-Arrhenius kinetics. We note that in a non-Arrhenius model, $k_1 \propto D = D_0 \exp(-E^2/2(k_B T)^2)$,⁸ the data indicate $E \approx 9-11 \text{ kJ/mol}$ under folding conditions of 4.4 M GdnHCl.

by $E_a^0 \approx 17$ kJ/mol, so that $E_a^{\text{chain}} \approx E_a - E_a^0 = (2 \pm 5)$ kJ/mol, consistent with zero, at high [GdnHCl]. This increases to $E_a^{\text{chain}} \approx 25 (\pm 9)$ kJ/mol at 4.4 M, where the native fold becomes stable. (This simple analysis does not consider any effect of denaturant on E_a^0 or differentiate the activation energies above and below the folding temperature T_f at each denaturant concentration.) Thus the energetic barriers may grow under folding conditions, although diffusion still occurs faster at low denaturant concentrations, i.e. these barriers do not control the rate of intrachain diffusion.

Conclusions

By examining the kinetics of contact formation in cytochrome *c*, we have measured the rate of intrachain diffusion of an unfolded protein under solvent and temperature conditions that span both sides of the unfolding transition. We find that the 62 residue loop in cytochrome *c* forms at a rate fully consistent with extrapolations from short peptide data. Although this rate, and its value in the limit of very short chains, may define one fundamental limit (~ 0.1 - $1 \mu\text{s}$) to the speed of protein folding, recent studies suggest that kinetic barriers to the initial collapse of a polypeptide may impose a more significant limitation (~ 50 - $100 \mu\text{s}$) on folding speed.³⁶⁻³⁸ Therefore, diffusional contact formation may not represent the most important physical constraint on folding speed.

The rate of intrachain diffusion in the unfolded molecule accelerates significantly as denaturant concentrations decrease, an effect that does not seem to result from changes in the dimensions of the unfolded chain. However, the Arrhenius temperature dependence of the diffusion rate gradually increases at lower denaturant. If we interpret this enthalpic contribution to the diffusion as an effect of non-covalent intrachain interactions, i.e. "landscape roughness", we conclude that the roughness grows as denaturant decreases, although it does not appear to control the conformational diffusion of the unfolded molecule under our experimental conditions. Like the kinetics of folding, the kinetics of intrachain loop formation may primarily reflect a single dominant barrier and give little indication of the presence of a more complicated energy landscape:³⁹ recent experimental work has shown that landscape effects do not necessarily generate non-exponential kinetics in folding, even at temperatures as low as $\sim 70^\circ\text{C}$ below the folding transition temperature.⁴⁰ The observed acceleration of loop formation at low

denaturant, in the presence of rising enthalpic barriers, may indicate that solvent entropy favors the formation of the intrachain contact, even if chain entropy does not; alternatively, denaturant interactions with the unfolded chain may significantly impede the dynamics of chain motion. Studies at very low denaturant concentrations might begin to show a larger role for landscape effects, but would require a different experimental system. Although our results do not necessarily indicate the absence of roughness in unfolded states, they do suggest that non-Arrhenius kinetics and energetic frustration may play a more important role in compact configurations of a polypeptide. Studies of intrachain diffusion in molten globule proteins could offer the best perspective on energy landscape effects in the dynamics of polypeptide chains.

Materials and Methods

Sample preparation and instrumentation

Horse cytochrome *c* (Sigma-Aldrich) was used without further purification. The protein was dissolved to a concentration of 70-100 μM in solutions containing variable concentrations of guanidine hydrochloride (GdnHCl, 99 + %, from Sigma), and 100 mM citric acid buffer (pH 6.0). The GdnHCl concentrations were checked refractometrically. The samples were then passed through a 0.2 μm syringe filter and stirred under a flow of deoxygenated CO for 0.5 hour in a glove-box. A trace amount of sodium dithionite (Fluka) was then added to reduce the heme iron to Fe(II). Each sample was then loaded under 1 atm CO pressure into a quartz spectrophotometry cell (NSG Precision Cells) with 0.5 mm path length, sealed anaerobically, and equilibrated overnight. The equilibrium absorbance of the samples was close to 1.0 at the carbonmonoxyheme Soret peak.

The cuvette was then secured in a thermally regulated aluminum mount in an anaerobic box on an optical table. The sample temperature was controlled by the fluid flow from a circulating bath, and monitored by fine thermocouples (125 μm gauge, from Omega Engineering) attached to the mount and the cuvette surface. Cuvette temperatures were uniform and stable to ± 0.2 deg. C or better during each experiment. Pulses ($\lambda = 532$ nm) from an Nd:YAG laser (Spectra-Physics), at ~ 4 mJ energy, 5-7 ns duration, and 0.33-10 Hz repetition rate, photodissociated the CO from the heme iron. At a variable time delay t after photolysis, a microsecond Xe flashlamp (EG&G) was triggered, producing a broadband probe light that was split and focused onto two points on the cuvette, in order to probe both photolyzed and unphotolyzed regions of the sample. Imaging optics collected the transmitted light and directed it through an imaging spectrometer (Acton) and onto a gated, intensified CCD camera (Roper). The nanosecond gating of the ICCD allowed the collection of absorption spectra with time resolution of 10 ns or better, despite the $\sim 1 \mu\text{s}$ duration of the Xe flash. These spectra indicated that the sample was fully photolyzed by each Nd:YAG pulse[†]. The spectrometer and associated electronics operated under the control of a desktop computer, through a Lab-View (National Instruments) software interface.

[†] At $T \leq 25$ - 30°C , the shape of the Soret spectrum of the immediate photoproduct suggests that a small fraction (up to $\sim 5\%$) of the hemes are in a low-spin configuration, indicating the presence of a ligand in the CO position. Similar spectra were observed at equilibrium in the Fe(II) cytochrome *c* octapeptide, and were attributed to partial ligation by H_2O .²¹

Kinetic modeling of spectra

We collected time-resolved absorption difference spectra $\Delta A(\lambda, t) = A(\lambda, t) - A(\lambda, 0^-)$, where $A(\lambda, t)$ is the optical absorbance at wavelength λ (=390-450 nm) and pump-probe time delay t . Each data set consisted of ~ 90 difference spectra collected at logarithmically spaced delay times t between 10 ns and 0.3-1 second, at fixed denaturant concentration and temperature. We used singular value decomposition (SVD) to reformulate each data set $\Delta A(\lambda, t)$ as a product of three matrices, $\Delta A(\lambda, t) = \mathbf{u}(\lambda) \mathbf{s} \mathbf{v}(t)^T$.⁴¹ The columns of $\mathbf{u}(\lambda)$ are orthonormal spectra, while the columns of $\mathbf{v}(t)$ are orthonormal functions that describe the time evolution of the corresponding columns of $\mathbf{u}(\lambda)$. A single column of $\mathbf{u}(\lambda)$, together with its associated column of $\mathbf{v}(t)$, represents one component of the SVD of $\Delta A(\lambda, t)$. The singular values s_i , or diagonal elements of \mathbf{s} , are ordered by decreasing size: we generally found s_2 and s_3 to be 2-7% as large as s_1 , while s_4 (and successive components) were less than 1% as large as s_1 and had signal-to-noise ratio less than unity. Thus each dataset contained, to a very good approximation, three significant spectral components (Figures 7 and 8).

The exponential decay constants τ_1 , τ_2 , τ_3 , and τ_4 of Figure 2 were obtained by fitting the first three SVD components, $s_i v_i(t)$ ($i = 1-3$), simultaneously to a sum of exponential decays:

$$s_i v_i(t) = c_{i1} \exp(-t/\tau_1) + c_{i2} \exp(-t/\tau_2) + c_{i3} \exp(-t/\tau_3) + c_{i4} \exp(-t/\tau_4)$$

Each decay constant τ_i corresponds to a spectral change $B_i(\lambda)$, i.e. $\mathbf{B}(\lambda) = \mathbf{u}(\lambda) \mathbf{c}$ and $\Delta A(\lambda, t) = \sum_{i=1-4} B_i(\lambda) \exp(-t/\tau_i)$.

We fit each data set to the kinetic model of Scheme I, in which Met80 binds to the heme at rate k_{M80} and Met65 binds at a rate $k_{M65} = k_{M80}/r$. We find that the fits do not clearly favor a particular value of r , since similar kinetics occur for any value of r in the reasonable range of $r \sim 0.1-1$: therefore we fix $r = (47/62)^{3/2} = 0.66$, which equation (4) predicts as the ratio of binding rates for Met80 and Met65 to the heme at His18. The kinetic rate matrix of the deoxyheme/Met65/Met80 subsystem then has eigenvalues $\lambda_1 = k_{-M} + (1 + r^{-1}) k_{M80}$, and $\lambda_2 = k_{-M}$. For the complete Scheme I, we expect roughly

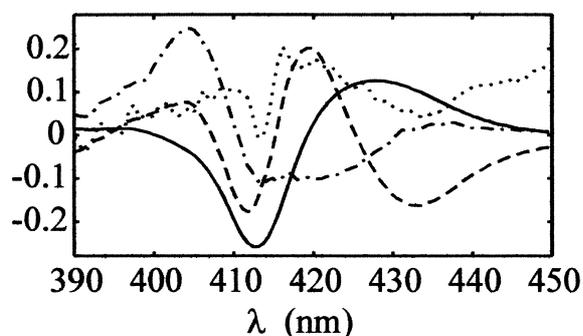


Figure 7. Leading spectral components $u_i(\lambda)$ obtained from singular value decomposition (SVD) of transient absorption difference spectra, same dataset as in Figure 1. u_1 , continuous curve; u_2 , broken curve; u_3 , dot-dashed curve; u_4 , dotted curve.

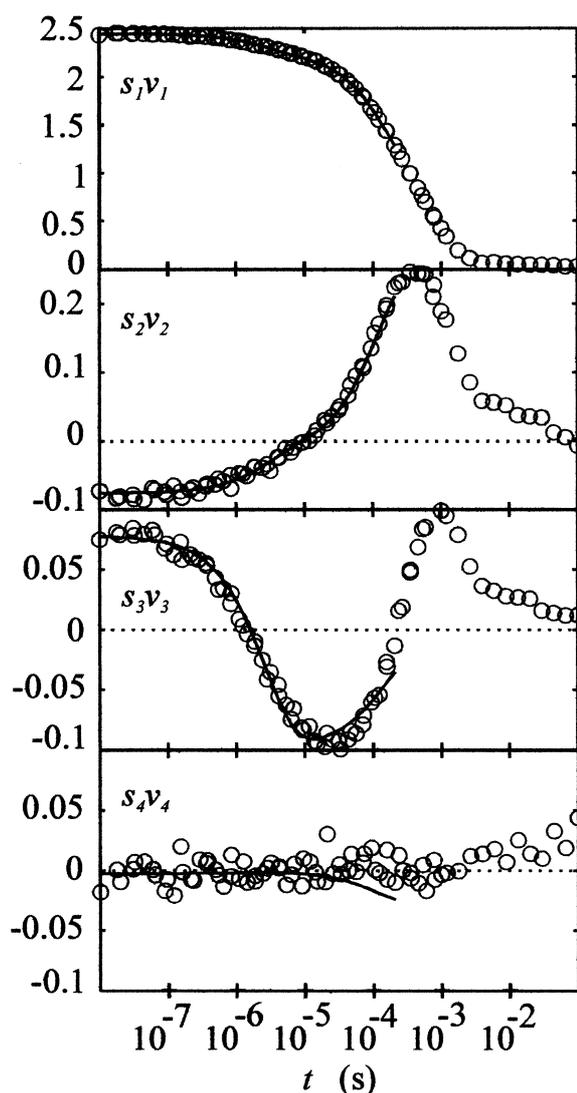


Figure 8. Amplitudes $s_i v_i(t)$ associated with the component spectra $u_i(\lambda)$ of Figure 7, from SVD of transient spectra, same dataset as Figure 1. Only the first three components exhibit a signal-to-noise ratio exceeding unity. Circles are experimental data; the continuous curve is fit to the kinetic model of Scheme I for $t = 10$ ns-200 μ s. Sum of squares for this fit is 0.045.

$$(\tau_1)^{-1} \approx k_{M80} + k_{M65} + k_{-M'} \quad (\tau_2)^{-1} \approx k_H + k_{-H'} \quad \text{and} \\ (\tau_3)^{-1} \approx k_{CO}$$

Generating a set of time-resolved spectra according to Scheme I requires the five independent rates (k_{M80} , $k_{-M'}$, k_H , $k_{-H'}$, k_{CO}) and estimates of the deoxyheme, heme-methionine, and Fe-His spectra. The spectrum of the initial deoxyheme photoproduct is a fixed input to the model, obtained by adding the early-time difference spectrum to the equilibrium carbonmonoxyheme spectrum. The fit generates the spectrum for the Fe-His complex; we parametrized the spectrum as a skewed Lorentzian in five parameters, and the fit estimated these parameters. Results agreed with published Fe(II) heme-histidine spectra.^{3,4,42} The methionine-liganded heme

spectrum is also a fixed input, obtained as follows: The fastest exponential decay τ_1 in the data corresponds to the replacement of the deoxyheme by the methionine complex. Thus $B_1(\lambda)$ is proportional to the deoxyheme minus methionine-heme difference spectrum. We obtained the methionine-heme spectrum by adding $B_1(\lambda)$ to the known deoxyheme spectrum until the result generated the best fit to reference Fe(II) heme-methionine spectra; we obtained reference spectra from our measurements of native cytochrome *c* and from published sources.^{3,4,42}

We found very good agreement between the observed transient spectra and the fit to Scheme I. The slow τ_4 process, which is presumably related to refolding/unfolding kinetics omitted from Scheme I, erodes the quality of fits that include millisecond time-scales (although it does not affect our analysis of the very fast heme-Met80 loop kinetics). Fits were therefore restricted to include times $t = 10$ ns to $t \geq 200$ μ s; each fit involved one set of experimental conditions (temperature and denaturant) and typically 5432 absorbance values, i.e. 56 times and 97 wavelengths in the range of $\lambda = 395$ –445 nm. The sum of squares, defined as $\sum_{t,\lambda} (A_{\text{measured}}(t,\lambda) - A_{\text{fit}}(t,\lambda))^2$, ranged from 0.023 to 0.17, with an average value of 0.045. We found no clear correlation between the sum of squares and the solvent conditions. Figures 7 and 8 show the SVD of a dataset for which the quality of fit was average.

Estimation of the heme-methionine binding rates

We estimated the chemical binding and dissociation rates k_3 and k_4 from the rate of bimolecular binding of free *N*-acetyl-L-methionine (Sigma) to Fe(II) microperoxidase-11 (Sigma), a heme-containing peptide obtained by the enzymatic cleavage of horse cytochrome *c*. We measured the bimolecular binding rate in 5.9 M GdnHCl for $T = 22$ –45 °C, and found $k^{\text{bi}} = (1.5 \pm 0.04) \times 10^8$ M⁻¹ s⁻¹ at 22 °C, which rises to $(2.1 \pm 0.2) \times 10^8$ M⁻¹ s⁻¹ at 40 °C. We use the bimolecular analogy of equation (1) to estimate k_3 from k^{bi} , which gives:

$$k_3 = k_2^{\text{bi}} k^{\text{bi}} / (k_1^{\text{bi}} - k^{\text{bi}})$$

Here k_1^{bi} and k_2^{bi} represent bimolecular (not intrachain) diffusion, and we use $k_1^{\text{bi}} = 2\pi D_0 a$ and $k_2^{\text{bi}} = 3 D_0 / a^2$, where D_0 is the bimolecular diffusion constant and a is the reaction radius. This differs from the usual $k_1^{\text{bi}} = 4\pi D_0 a$ because methionine can bind only to the distal side of the heme iron. Using these relations and the values $a = 0.15$ nm for a heme-ligand reaction⁴³ and $D_0 \approx 6.8 \times 10^{-6}$ cm² s⁻¹ for free methionine at $T = 22$ °C, we find $k_3 \approx 6.1 \times 10^{10}$ s⁻¹; for $T = 40$ °C this becomes $k_3 \approx 9.0 \times 10^{10}$ s⁻¹.

We can estimate k_4 from the association constant $K_{\text{bimolecular}}$ of free methionine to the Fe(II) heme. Our rate studies at 5.9 M GdnHCl indicate $K_{\text{bimolecular}} \approx 3(\pm 1)$ mM⁻¹ at 22 °C, while equilibrium studies find $K_{\text{bimolecular}} = 2.4$ mM⁻¹.²¹ Since the Tezcan *et al.* value, measured in water, matches our value in 5.9 M GdnHCl, we conclude that $K_{\text{bimolecular}}$ for methionine does not vary greatly with GdnHCl. This was already shown for histidine ligands.⁴⁴ We find only a weak temperature-dependence in $K_{\text{bimolecular}}$. Using the association constant of Tezcan *et al.*,²¹ we then have $k_4 = k_1^{\text{bi}} / k_2^{\text{bi}} k_3 / K_{\text{bimolecular}} \approx 6.2 \times 10^5$ s⁻¹ at 22 °C, and $k_4 \approx 9.2 \times 10^5$ s⁻¹ at 40 °C, with $k_3/k_4 \approx 1 \times 10^5$.

Acknowledgments

The authors thank Arthur S. Edison for helpful discussions. This work is supported by the National Science Foundation, Division of Molecular and Cellular Biosciences (MCB 0077907), and by the University of Florida, College of Liberal Arts and Sciences.

References

- Eaton, W. A., Muñoz, V., Hagen, S. J., Jas, G. S., Lapidus, L. J., Henry, E. R. & Hofrichter, J. (2000). Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327–359.
- Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996). Diffusion-limited contact formation in unfolded cytochrome *c*: estimating the maximum rate of protein folding. *Proc. Natl Acad. Sci. USA*, **93**, 11615–11617.
- Hagen, S. J., Hofrichter, J. & Eaton, W. A. (1997). Rate of intrachain diffusion of unfolded cytochrome *c*. *J. Phys. Chem. ser. B*, **101**, 2352–2365.
- Jones, C. M., Henry, E. R., Hu, Y., Chan, C.-K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J. & Eaton, W. A. (1993). Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl Acad. Sci. USA*, **90**, 11860–11864.
- Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. & Kiefhaber, T. (1999). The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl Acad. Sci. USA*, **96**, 9597–9601.
- Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000). Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl Acad. Sci. USA*, **97**, 7220–7225.
- Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Struct. Funct. Genet.* **21**, 167–195.
- Zwanzig, R. (1988). Diffusion in a rough potential. *Proc. Natl Acad. Sci. USA*, **85**, 2029–2030.
- Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. & Socci, N. D. (1995). Toward an outline of the topography of a realistic protein-folding funnel. *Proc. Natl Acad. Sci. USA*, **92**, 3626–3630.
- Bushnell, G. W., Louie, G. V. & Brayer, G. D. (1990). High resolution three dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.* **214**, 585–595.
- Pascher, T., Chesick, J. P., Winkler, J. R. & Gray, H. B. (1996). Protein folding triggered by electron transfer. *Science*, **271**, 1558–1560.
- Telford, J. R., Wittung-Stafshede, P., Gray, H. B. & Winkler, J. R. (1998). Protein folding triggered by electron transfer. *Acc. Chem. Res.* **31**, 755–763.
- Chan, C.-K., Hofrichter, J. & Eaton, W. A. (1996). Optical triggers of protein folding. *Science*, **274**, 628–629.
- Akiyama, S., Takahashi, S., Ishimori, K. & Morishima, I. (2000). Stepwise formation of α -helices during cytochrome *c* folding. *Nature Struct. Biol.* **7**, 514–520.
- Chen, E., Wood, M. J., Fink, A. J. & Kliger, D. S. (1998). Time-resolved circular dichroism studies of protein folding intermediates of cytochrome *c*. *Biochemistry*, **37**, 5589–5598.

16. Goldbeck, R. A., Thomas, Y. G., Chen, E., Esquerra, R. M. & Kliger, D. S. (1999). Multiple pathways on a protein folding energy landscape: kinetic evidence. *Proc. Natl Acad. Sci. USA*, **96**, 2782-2787.
17. Colón, W., Wakem, L. P., Sherman, F. & Roder, H. (1997). Identification of the primary non-native histidine ligand in unfolded cytochrome *c*. *Biochemistry*, **36**, 12535-12541.
18. Wang, J. C. & Davidson, N. (1966a). Thermodynamic and kinetic studies on the interconversion between the linear and circular forms of phage lambda DNA. *J. Mol. Biol.* **15**, 111-123.
19. Wang, J. C. & Davidson, N. (1966b). On the probability of ring closure of lambda DNA. *J. Mol. Biol.* **19**, 469-482.
20. Jacobsen, H. & Stockmayer, W. H. (1950). Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* **18**, 1600-1606.
21. Tezcan, F. A., Winkler, J. R. & Gray, H. B. (1998). Effects of ligation and folding on reduction potentials of heme proteins. *J. Am. Chem. Soc.* **120**, 13383-13388.
22. Tanford, C. (1968). Protein denaturation. *Advan. Protein Chem.* **23**, 121-282.
23. Segel, D. J., Fink, A. L., Hodgson, K. O. & Doniach, S. (1998). Protein denaturation: a small-angle X-ray scattering study of the ensemble of unfolded states of cytochrome *c*. *Biochemistry*, **37**, 12443-12451.
24. Wilemski, G. & Fixman, M. (1974). Diffusion-controlled intrachain reactions of polymers. II Results for a pair of terminal reactive groups. *J. Chem. Phys.* **60**, 878-890.
25. Doi, M. (1975). Diffusion controlled reaction of polymers. *Chem. Phys.* **9**, 455-466.
26. Szabo, A., Schulten, K. & Schulten, Z. (1980). First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **72**, 4350-4357.
27. Chan, H. S. & Dill, K. A. (1989). Intrachain loops in polymers: effects of excluded volume. *J. Chem. Phys.* **90**, 492-509.
28. Camacho, C. J. & Thirumalai, D. (1995). Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic trypsin inhibitor. *Proc. Natl Acad. Sci. USA*, **92**, 1277-1281.
29. Thirumalai, D. (1999). Time scales for the formation of the most probable tertiary contacts in proteins with applications to cytochrome *c*. *J. Phys. Chem. ser. B*, **103**, 608-610.
30. Winnik, M. A. (1985). End-to-end cyclization of polymer chains. *Acc. Chem. Res.* **18**, 73-79.
31. Flory, P. J. (1969). *Statistical Mechanics of Chain Molecules*, John Wiley, New York.
32. Buckler, D. R., Haas, E. & Scheraga, H. A. (1995). Analysis of the structure of ribonuclease A in native and partially denatured states by time-resolved nonradiative excitation energy transfer between site-specific extrinsic probes. *Biochemistry*, **34**, 15965-15978.
33. Kawahara, K. & Tanford, C. (1966). Viscosity and density of aqueous solutions of urea and guanidine hydrochloride. *J. Biol. Chem.* **241**, 3228-3232.
34. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry*, **30**, 10428-10435.
35. Chan, C.-K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A. & Hofrichter, J. (1997). Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc. Natl Acad. Sci. USA*, **94**, 1779-1784.
36. Shastry, M. C. R. & Roder, H. (1998). Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nature Struct. Biol.* **5**, 385-392.
37. Park, S. H., Shastry, M. C. R. & Roder, H. (1999). Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing. *Nature Struct. Biol.* **6**, 943-947.
38. Hagen, S. J. & Eaton, W. A. (2000). Two-state expansion and collapse of a polypeptide. *J. Mol. Biol.* **301**, 1019-1027.
39. Karplus, M. (2000). Aspects of protein reaction dynamics: deviations from simple behavior. *J. Phys. Chem. ser. B*, **104**, 11-27.
40. Gillespie, B. & Plaxco, K. W. (2000). Nonglassy kinetics in the folding of a single-domain protein. *Proc. Natl Acad. Sci. USA*, **97**, 12014-12019.
41. Henry, E. R. & Hofrichter, J. (1992). Singular value decomposition - application to analysis of experimental-data. *Methods. Enzymol.* **210**, 129-192.
42. Telford, J. R., Tezcan, F. A., Gray, H. B. & Winkler, J. R. (1999). Role of ligand substitution in ferrocyclochrome *c* folding. *Biochemistry*, **38**, 1944-1949.
43. Miers, J. B., Postlewaite, J. C., Zyung, T., Chen, S., Roemig, G. R., Wen, X., Dlott, D. D. & Szabo, A. (1990). Diffusion can explain the nonexponential rebinding of carbon-monoxide to protoheme. *J. Chem. Phys.* **93**, 8771-8776.
44. Godbole, S., Hammack, B. & Bowler, B. E. (2000). Measuring denatured state energetics: deviations from random coil behavior and implications for the folding of iso-1-cytochrome *c*. *J. Mol. Biol.* **296**, 217-228.

Edited by C. R. Matthews

(Received 2 October 2000; received in revised form 30 November 2000; accepted 30 November 2000)