# Rate of Intrachain Diffusion of Unfolded Cytochrome c

# Stephen J. Hagen, James Hofrichter, and William A. Eaton\*

Laboratory of Chemical Physics, Building 5, National Institutes of Health, Bethesda, Maryland 20892-0520 Received: July 30, 1996; In Final Form: December 3, 1996<sup>®</sup>

How fast can a protein fold? The rate of initial collapse from the unfolded state to a compact structure provides one upper limit to the folding rate. Although hydrophobic collapse of heteropolymers is not well understood, its rate may be controlled by the rate at which contacts form between distant parts of an unfolded polypeptide chain. The rate of this intrachain diffusion has not been measured directly. However, information about that time scale is contained in the experimental results of Jones et al. (Proc. Natl. Acad. Sci. U.S.A. **1993**, 90, 11860), who triggered the folding of reduced cytochrome c by nanosecond photolysis of the carbon monoxide complex. Jones et al. found that the methionine residues at positions 65 and 80 bind to the heme at position 18 at a rate of  $(40 \mu s)^{-1}$ , while the histidine residues at positions 33 and 26 bind at a rate of  $(400 \mu s)^{-1}$  $\mu$ s)<sup>-1</sup>. To identify the separate contributions of chain dynamics and chemical bond formation (*i.e.* the geminate binding rates) to the observed rates, we have used nanosecond-resolved absorption spectroscopy to determine the bimolecular and geminate rates for free methionine and histidine binding to the hemepeptide of cytochrome c under the solvent conditions of Jones et al. The rate of his33 (and his26) binding to the heme of the intact polypeptide appears to be limited by a slow geminate rate and by the equilibrium probability that the required loop will form spontaneously. In contrast, the binding of met65 and met80 is rate limited only by the diffusion of the polypeptide chain to form an encounter complex. The  $(40 \,\mu\text{s})^{-1}$  rate observed by Jones et al. therefore allows us to calculate the met80-his18 intrachain diffusion rate  $k_{\rm D+} \approx (35-40~\mu \rm s)^{-1}$ . From this result we estimate that the smallest intrachain loops in a polypeptide will form in no less than  $\sim 1 \mu s$ . This may set a limit of  $\sim 10^6$  s<sup>-1</sup> on the rate of collapse of the polypeptide chain under folding conditions. We also use the theory of Szabo et al. (J. Chem. Phys. 1980, 72, 4350) to calculate the relative diffusion constant D of the heme and methionine residues, obtaining a value  $D \approx 4 \times 10^{-7}$  cm<sup>2</sup>/s.

### Introduction

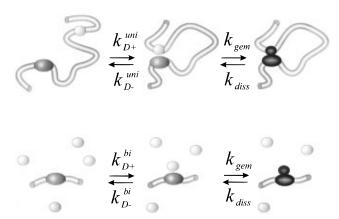
Recent advances in the study of protein folding on the submillisecond time scale $^{1-3}$  and the observation of very rapid folding in many systems<sup>4,5</sup> raise an important question: how rapidly can a protein fold? Energy landscape theory<sup>6</sup> suggests that conditions exist under which some proteins may fold without encountering a thermodynamic barrier, so that folding is limited only by the speed of collapse to a compact structure. The mechanism of heteropolymer collapse is poorly understood, however, and its rate cannot be predicted by existing theory. If collapse requires regions of the polypeptide chain to come into contact by diffusion, then the rate of intrachain diffusion provides an upper limit to the rate of collapse, and therefore to the rate of folding. Estimates of the intrachain diffusion rate have appeared from time to time on the basis of simple arguments.<sup>7-9</sup> Intrachain contacts are believed to form within microseconds, while subsequent stages of folding may proceed at far slower rates. We have measured the rate at which contact occurs between sites separated by 50-60 residues on the chain of unfolded cytochrome c. By extrapolating this rate down to the limit of the shortest intrachain loops, we are able to estimate the maximum speed of collapse and folding of a polypeptide.

To resolve the rapid rebinding, we utilize the optical trigger for folding employed by Jones  $et\ al.^1$  Those authors greatly enhanced the time resolution of folding studies by using a nanosecond laser pulse to photodissociate a carbon monoxide (CO) ligand from the heme group covalently attached to unfolded cytochrome c. With the CO dissociated, the heme is free to bind the methionine or other residues on the chain, and folding of the polypeptide can occur.

The experiment of Jones et al. was performed at 40 °C, pH 6.5, in both 5.6 and 4.6 M guanidine hydrochloride (GuHCl). In the presence of CO the protein is completely unfolded in both solvents, as judged by circular dichroism and fluorescence spectroscopies.<sup>1</sup> Upon CO dissociation, the protein remains unfolded in 5.6 M GuHCl, while in 4.6 M GuHCl more than 95% of molecules are in the native conformation at equilibrium. Measurements of the fluorescence quenching of tryptophan by the heme<sup>10</sup> show no evidence of collapse to a compact structure prior to the rebinding of CO at  $\sim$ 1 ms which aborts the folding process. The native conformation appears with a time constant of about 1 s at 4.6 M GuHCl when continuous illumination is used to prevent CO rebinding or when folding is initiated by chemical dilution from 5.6 to 4.6 M in a stopped flow experiment.<sup>1,11</sup> The data so far suggest that at 4.6–5.6 M GuHCl cytochrome c behaves very much like a two-state system, with an unfolded state that is near random coil and a compact, native conformation.<sup>1,3</sup> The slow folding, then, is presumably due to a large thermodynamic barrier separating these states, which results from the competition between energetically-stabilizing contacts and configurational entropy. This simple two-state folding is described as "scenario I" in the statistical energy landscape theory of Wolynes and co-workers.<sup>6</sup>

Jones *et al.*<sup>1</sup> did not observe the folding of cytochrome c but did observe important early events in folding. Using time-resolved absorption spectroscopy with nanosecond lasers, they observed the formation of both native (met80) and non-native (met65, his33, and his26) intrachain complexes with the heme iron at his18 at times beginning hundreds of nanoseconds after photolysis of the CO complex. Their modeling of the time-resolved spectra indicated that the heme complex with met65 or met80 (which cannot be distinguished by the optical spectra)

<sup>&</sup>lt;sup>®</sup> Abstract published in *Advance ACS Abstracts*, January 15, 1997.



**Figure 1.** Two-step model for ligand binding by a heme on a polypeptide. The heme and ligand diffuse together at a rate  $k_{\rm D+}$  to form the "encounter complex", which then either separates (rate  $k_{\rm D-}$ ) or reacts ("geminate" rate  $k_{\rm gem}$ ) to give the covalent complex. (The covalent complex dissociates spontaneously at a rate  $k_{\rm diss} \ll k_{\rm gem}$ .) The model describes both the unimolecular reaction (upper), in which heme and ligand are bound to the polypeptide, and the bimolecular reaction (lower), in which the ligand is free in solution.

forms in  $\approx$ 40  $\mu$ s and dissociates in  $\approx$ 4  $\mu$ s. The complex with his33 or his26 (also spectrally indistinguishable) forms in  $\approx$ 400  $\mu$ s and dissociates in  $\approx$ 170  $\mu$ s. Thus the methionine—heme complexes form 10 times faster than the histidine—heme complexes, even though the histidine residues are positioned much closer to the heme.

Jones et al. encountered two difficulties in analyzing their experimental results. First, their assignment of the observed  $(40 \mu s)^{-1}$  and  $(400 \mu s)^{-1}$  binding rates to methionine and histidine residues, respectively, rather than the reverse assignment, was based on a comparison of the time-resolved absorption spectra obtained through kinetic modeling with reference spectra obtained from the dimethyl sulfide and imidazole complexes of unfolded cytochrome c. In order to help confirm these assignments, we have conducted experiments on photolysis of the CO complex of dicarboxymethylated horse cytochrome c in GuHCl. In the dicarboxymethylated protein, residues met65 and met80 are chemically modified to prevent their binding to the heme iron. We find that the rapid kinetic phase at  $\sim$ 4  $\mu$ s (identified by Jones *et al.* as the time scale,  $\tau \approx (k_{\rm on} + k_{\rm off})^{-1}$ , for met65 and met80 binding) is absent, while the slower  $\sim$ 100 μs kinetic phase (attributed to his33 or his26 binding) remains. This result supports Jones et al.'s description of the sequence of heme-ligand binding reactions following breakage of the heme-CO bond.

A second difficulty encountered by Jones *et al.* was that the measured complex formation rates reflect not only the diffusion rate of the chain but also the heme—ligand bond formation. Because the intrinsic rates for histidine and methionine binding to the heme of cytochrome *c* were not known, the rates of chain diffusion could not be obtained from the measured covalent complex formation rates.

We therefore measured the intrinsic rates of histidine and methionine binding to microperoxidase, a heme-containing peptide that serves as an analog for the heme and nearby residues of cytochrome c. We used nanosecond-resolved absorption spectroscopy to measure the bimolecular binding rate of free histidine and methionine to the reduced heme peptide under solution conditions identical to those of Jones  $et\ al$ . From these bimolecular rates and the data of Jones  $et\ al$ ., we have calculated the rate of relative diffusion of the heme and ligand residue on the unfolded cytochrome c chain. This calculation is based on an analogy (Figure 1) between unimolecular and bimolecular binding of ligands to a heme on a polypeptide chain. It assumes

that the intrinsic or "geminate" rate  $k_{\rm gem}$  for formation of the iron—ligand bond from an "encounter complex" (consisting of the heme—ligand pair) is the same for the unimolecular and bimolecular reactions. Wang and Davidson<sup>12</sup> used a similar approach to show that the ring-closure reaction of  $\lambda$  DNA is reaction-limited.

One might expect that the bimolecular rates could be measured by mixing the heme peptide with free histidine or methionine in a stopped-flow kinetic spectrometer. Unfortunately, the unliganded peptide forms a slowly dissociating aggregate<sup>13</sup> even at 5.6 M GuHCl, making this experiment impossible. The rebinding kinetics that follow photolysis of the heme-ligand complex therefore offer a better means for obtaining these rates. Since the yield for photolysis of the methionine and histidine complexes is poor with a nanosecond laser, 14 we employed a strategy first used by Gibson 15 for ligandbinding studies in hemoglobin. We formed the complex of the heme peptide with a tightly binding but photodissociable ligand, such as CO or *n*-butyl isocyanide, in the presence of excess histidine or methionine. Upon photolysis of the complex, the free amino acid binds transiently to the unliganded heme. Timeresolved spectroscopy was used to measure the bimolecular rate for this reaction.

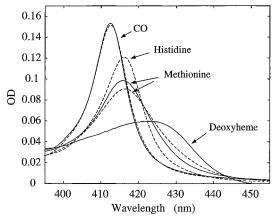
A brief account of some of this work has already been presented.<sup>14</sup>

#### **Materials and Methods**

To test Jones et al.'s interpretation of the intrachain ligandbinding rates observed in unfolded cytochrome c, we have repeated their experiment with dicarboxymethylated cytochrome c. In this case the methionine residues at positions 65 and 80 are alkylated and cannot bind to the heme iron, with the consequence that only his 26 and his 33 (and the free CO) should be observed binding to the heme. In order to make a detailed comparison between the kinetics of the unmodified and carboxymethylated proteins, we chose to work at a pH significantly above the pK ( $\approx$ 6.5–7) of his33 and his26; although the greater deprotonation of the histidine residues was expected to accelerate histidine-heme binding, it was also expected to make the histidine-heme reaction rate less sensitive to any variations in the pH of the samples. We therefore performed the experiment at T = 23 °C, pH 8.0, while Jones et al. worked at T = 40 °C, pH 6.5. The slightly lower temperature partially compensated for the acceleration of rebinding. The results presented in the next section confirm that the reaction rates under our conditions are quite similar to those observed by Jones et al.

Horse heart cytochrome c was obtained from Sigma and dicarboxymethylated by the method of Schejter. The modification of met80 and met65 was confirmed by the absence of the characteristic ferricytochrome c absorption band at 695 nm. Like Jones  $et\ al.$ , we studied the formation of transient intrachain complexes in the unfolded protein by collecting time-resolved absorption difference spectra after photolysis of the unfolded cytochrome c/CO complex. Samples of cytochrome c and dicarboxymethylated cytochrome c were prepared in 5.2 M GuHCl with 1 atm of CO at pH 8.0 (phosphate 0.1 M). The protein concentration was 100  $\mu$ M, and sodium dithionite (400  $\mu$ M) was added to maintain the heme iron in the ferrous state. The samples were loaded anaerobically into quartz cuvettes with 350  $\mu$ m path length and then sealed and placed in the transient spectrometer (see below).

For measurements of the rate of free histidine and methionine binding, we used the 11-residue heme-containing peptide that is obtained by enzymatic digestion of horse cytochrome c. This peptide, available commercially as microperoxidase (Sigma



**Figure 2.** Spectra of heme complexes in 5.6 M GuHCl, pH 6.5,  $T=40~^{\circ}\text{C}$ : deoxyheme spectrum is from unfolded cytochrome  $c^1$ ; dashedline spectra are equilibrium spectra of the heme peptide (100 μM) measured in the presence of 1 atm CO, 78 mM *N*-acetyl-L-methionine, or 234 mM *N*-α-acetyl-L-histidine, as labeled; solid-line spectra labeled CO and methionine are calculated spectra for the pure CO and methionine complexes, obtained from the transient spectra observed following photodissociation of a heme peptide/*N*-acetyl-L-methionine/CO sample (see Appendix A). The equilibrium methionine spectrum is broader than the methionine-complex spectrum, most likely because in 78 mM methionine the methionine-liganded and deoxyheme are both present.

M-6756), consists of residues 11–21 together with the heme iron complex that is covalently attached to his 18 of the intact protein. We studied the bimolecular binding of N- $\alpha$ -acetyl-Lhistidine (Sigma A-8133) and N-acetyl-L-methionine (Aldrich 85,534-0) to the heme of microperoxidase under the solution conditions of Jones et al.1 The lyophilized peptide was dissolved in 5.6 M guanidine hydrochloride (GuHCl), phosphate 0.1 M, pH 6.5, with free histidine (20–200 mM) or methionine (10-100 mM) present. The peptide concentration was approximately 100  $\mu$ M in all experiments. The samples were equilibrated with 1 atm of CO, and sodium dithionite was added to a concentration of  $\approx$ 500  $\mu$ M in order to reduce the heme iron to the ferrous state. Samples were prepared anaerobically and loaded into 350 µm quartz cuvettes, sealed, and placed in the transient spectrometer for reaction-rate measurements. Figure 2 shows the equilibrium spectra of the samples.

We used flash photolysis and nanosecond-resolved absorption spectrometry to study the binding reactions. The spectrometer has been described in detail elsewhere. Briefly, it consists of two Nd:YAG lasers generating 10 ns pulses at 1066 nm. The 533 nm second harmonic of the first laser photodissociates the heme—ligand complex. The 355 nm third harmonic of the second laser excites a fluorescent dye (Stilbene 420) to generate a broad band probe source which allows transient absorption spectra of the heme Soret region of the sample to be recorded on a silicon vidicon detector.

Under each set of temperature and solution conditions, time-resolved absorption difference (*i.e.* photolyzed minus unphotolyzed) spectra were collected at times logarithmically spaced from  $\sim 10$  ns to  $\sim 100$  ms after photolysis. These spectra, organized as a matrix  $\mathbf{a}(\lambda,t)$  of absorption values in wavelength  $\lambda$  and time t, were analyzed by singular value decomposition (SVD). SVD represents the matrix  $\mathbf{a}(\lambda,t)$  as the product of three matrices

$$\mathbf{a}(\lambda,t) = \mathbf{V}(t) \mathbf{S} \mathbf{U}^{\mathrm{T}}(\lambda)$$

where the columns of  $\mathbf{U}(\lambda)$  describe orthonormal basis spectra from which the observed spectra in  $\mathbf{a}(\lambda,t)$  can be constructed and the columns of  $\mathbf{V}(t)$  describe time-dependent, normalized

amplitudes associated with the basis spectra. Thus, the columns of V(t) describe the time dependence of spectral changes that occur after photolysis. The diagonal matrix S contains the singular values, or weights associated with each basis spectrum. The singular values are ordered by decreasing size, so that the first n singular values (and associated columns of  $U(\lambda)$  and V(t)) provide the best n-component least-squares fit to the full data set. Amplitude offsets resulting from variations in the detector dark current were removed from  $a(\lambda,t)$  by the method of Jones  $et\ al.^{19}$ 

### Results

**Dicarboxymethylated Cytochrome c.** By modeling the spectral changes that follow photodissociation of the heme-CO complex of unfolded cytochrome c at pH 6.5, T = 40 °C, Jones et al. concluded that met65 and met80 bind to the heme at a rate  $k_{\rm on} \approx (40 \,\mu{\rm s})^{-1}$  and dissociate at a rate  $k_{\rm off} \approx (4 \,\mu{\rm s})^{-1}$ . Therefore methionine binding is incomplete ( $k_{\rm on}/k_{\rm off} \approx 0.1$ ) but accounts for an exponential process in the time-resolved spectra with a time constant  $\tau = (k_{\rm on} + k_{\rm off})^{-1} \approx 3-4 \,\mu{\rm s}$ . The same analysis found that his33 (and perhaps also his26) binds to the heme at a rate  $k_{\rm on} \approx (400 \ \mu \rm s)^{-1}$  and dissociates at  $k_{\rm off} \approx (170 \ \rm m)^{-1}$  $\mu$ s)<sup>-1</sup>, generating a process with a time constant  $\tau = (k_{on} + 1)^{-1}$  $k_{\rm off}$ )<sup>-1</sup>  $\approx 120 \,\mu s$ . However, the analysis of Jones *et al.* was not definitive. The Soret spectra of the methionine and histidine complexes of the reduced heme are difficult to distinguish under unfolding conditions (Figure 2), and the transient heme spectra observed in the experiment of Jones et al. did not perfectly match those of model compounds. Jones et al. therefore could not establish with certainty that the  $3-4 \mu s$  process is due to methionine (as opposed to histidine) binding to the heme and that the 120 µs process is due to histidine (vs methionine) binding. If the analysis of Jones et al. is correct, the rapid (approximately a few microseconds) process will be present in the photolysis spectra of unfolded cytochrome c/CO but not in the spectra of the dicarboxymethylated cytochrome c/CO, while the slower ( $\sim$ 100  $\mu$ s) process will be present in both data sets.

Figure 3 gives a simple demonstration of the early spectral changes observed in our photolysis study of cytochrome c/CO unfolded in 5.2 M GuHCl at pH 8, T = 23 °C. For both the carboxymethylated and unmodified protein samples, the figure compares the Soret absorption difference spectrum at time t after photolysis with the initial (t = 10 ns) spectrum. The sum-ofsquares difference between the spectrum at t and the initial spectrum (normalized to the same amplitude as the spectrum at t) is calculated as a function of t. The figure therefore accentuates changes in the shape of the absorption difference spectrum during ligand rebinding. Both samples are seen to exhibit major spectral changes with time constants of  $\sim 100 \,\mu s$ and  $\sim$ 2 ms. These can be attributed to the formation of the heme-histidine complex, followed by the displacement of the histidine ligand by CO. However, Figure 3 also shows that the unmodified cytochrome c sample exhibits an additional, early (approximately a few microseconds) spectral change just prior to the larger change at  $\sim 100 \,\mu s$ . The dicarboxymethylated cytochrome c sample does not exhibit this early process. Because the only difference between the two samples is the carboxymethylation of the methionine residues in the second sample, we conclude that transient binding of the methionine residues to the heme iron is responsible for the small amplitude process occuring within a few microseconds in the unmodified sample. This appears to confirm the rate assignments of Jones et al. for the intrachain ligand binding events that follow photodissociation of the CO ligand from unfolded cytochrome c.

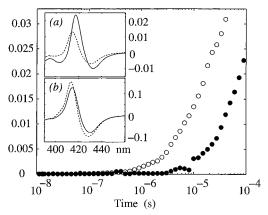
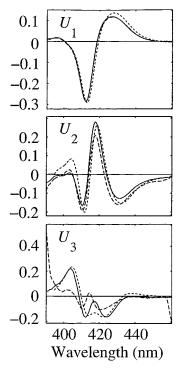
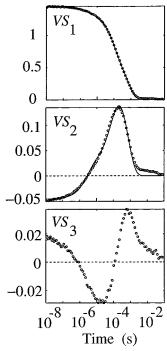


Figure 3. Evolution of the time-resolved photolysis difference spectra of unmodified (open circles) cytochrome c and dicarboxymethylated (solid circles) cytochrome c unfolded in 1 atm CO, 5.2 M GuHCl, pH 8.0, at T = 23 °C. For each experimental time t (and for each sample), we scaled the initial (t = 10 ns) difference spectrum so that it had as nearly as possible the same amplitude as the photolysis difference spectrum observed at t. We then subtracted this scaled spectrum from the observed difference spectrum to obtain a new difference spectrum  $\Delta(\lambda,t)$ .  $\Delta(\lambda,t)$  measures the difference in the shape (as opposed to the amplitude) of the spectrum at t relative to that at t = 10 ns. The figure shows the sum-of-squares magnitude of  $\Delta(\lambda,t)$  (i.e.  $\sum_{\lambda}(\Delta(\lambda,t))^2$ ) as a function of t, indicating the time scale of changes in the shape of the sample spectra. The unmodified sample appears to exhibit an additional early ( $\tau < \sim 10 \ \mu s$ ) spectral change. This is presumably due to methionine binding to the deoxyheme, as suggested by Jones et al.1 The insets show the double-difference spectrum (i.e. the simple difference between the absorption spectrum at t and at 10 ns) for unmodified cytochrome c (solid curve) and carboxymethylated cytochrome c (dashed curve) at  $t = 3.7 \mu s$  (a) and 110  $\mu s$  (b).

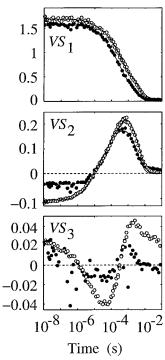


**Figure 4.** Significant spectral components  $U_i(\lambda)$  (i = 1-3) obtained from SVD of the absorption difference spectra measured after photodissociation of the CO complex of unfolded cytochrome c: (solid curve) unmodified cytochrome c and (long dash) dicarboxymethylated cytochrome c in 5.2 M GuHCl at T = 23 °C, pH 8.0; (short dash) unmodified cytochrome c in 5.6 M GuHCl at T = 40 °C, pH 6.5 (data of Jones et al.1).

Figures 4-6 present a more detailed analysis of the data. Figure 4 shows the significant spectral components  $U_i(\lambda)$  (i =1-3) obtained from SVD of our photolysis data for cytochrome



**Figure 5.** Amplitudes  $VS_1(t) - VS_3(t)$  associated with the spectral components of Figure 4, for pH 6.5, T = 40 °C data of Jones et al. The spectral changes are dominated by three phases, as shown by a fit of  $VS_1(t)$  and  $VS_2(t)$  to a sum of three exponentials (solid curves).



**Figure 6.** Amplitudes  $VS_1(t) - VS_3(t)$  associated with the spectral components of Figure 4 for the carboxymethylated (solid circles) and unmodified (open circles) cytochrome c at pH 8.0, T = 23 °C. The fit (solid curve) to  $VS_1$  and  $VS_2$  is to a sum of four exponentials for the unmodified protein and three exponentials for the carboxymethylated protein. The fit requires that two of the time constants be the same for both the modified and unmodified protein.

c/CO and dicarboxymethylated cytochrome c/CO, and for the data of Jones et al. (i.e. cytochrome c/CO at 40 °C in 5.6 M GuHCl, pH 6.5). The first two components  $U_1(\lambda)$  and  $U_2(\lambda)$ are very similar for all three data sets, since in all three samples similar reactions occur upon photodissociation of the CO complex—i.e. formation of a deoxyheme which transiently binds histidine (or methionine) before rebinding CO-and because the methionine and histidine complexes of the reduced heme exhibit similar spectra (Figure 2). The effect of carboxymethylation is primarily visible as a slight blue shift and reduction in amplitude of the  $\approx$ 418 nm peak in  $U_2(\lambda)$ . This reflects the absence of the heme—methionine complex, whose Soret absorption peak (near 418 nm) in unfolded cytochrome c is slightly red shifted relative to the heme—histidine complex. A difference is also visible in the third component  $U_3(\lambda)$ , which is similar in the samples of unmodified cytochrome c but quite different for the carboxymethylated sample. These differences confirm that the unmodified protein undergoes a process, i.e. transient binding of met65 and met80, that does not occur in the carboxymethylated protein.

The kinetics of these spectra are analyzed in Figures 5 and 6. Figure 5 shows the amplitudes  $VS_1-VS_3$  associated with  $U_1-U_3$  in the pH 6.5, T=40 °C cytochrome c data of Jones et~al. The spectral changes are dominated by three phases, as shown by fitting both  $VS_1(t)$  and  $VS_2(t)$  to a sum of three exponentials. The fit gives time constants  $\tau_1 \approx 2.8~\mu s$ ,  $\tau_2 \approx 93~\mu s$ , and  $\tau_3 \approx 710~\mu s$ . In the model of Jones et~al., these times were interpreted as the time constants for heme-methionine binding, heme-histidine binding, and heme-CO recombination, respectively.

Figure 6 shows  $VS_1-VS_3$  for our samples of carboxymethylated and unmodified cytochrome c at 23 °C, pH 8.0. As with the spectral components  $U_i(\lambda)$ , the kinetics for the unmodified protein closely resemble the data of Jones et al., confirming that the sequence of reactions and time constants occuring under our experimental conditions is similar to that at pH 6.5, T = 40°C. However, the kinetics are clearly different for the carboxymethylated protein, as evident from the early time (t < t $\sim 1 \,\mu s$ ) behavior of  $VS_2$ . The kinetics of the unmodified protein can be well-characterized by a sum of exponentials, as with the Jones data, Figure 5. The carboxymethylated protein—lacking a methionine-binding phase—can be fit with one fewer process. Figure 6 shows  $VS_1$  and  $VS_2$  of the unmodified protein fit to a sum of four exponentials with time constants  $\tau_1 - \tau_4$  and  $VS_1$ and  $VS_2$  of the carboxymethylated protein fit to a sum of three exponentials with time constants  $\tau_5 - \tau_7$ . For both cases, we interpret one of the time constants as resulting from heme-CO rebinding and two time constants as resulting from hemehistidine binding. Heme-histidine binding is very plausibly a biexponential process, since his26 and his33 may encounter very different steric constraints in binding to the heme. Elöve et al.5 observed two kinetic phases in the dissociation of histidine residues from the heme of unfolded oxidized cytochrome c upon solvent acidification. The rates for these phases differed by a factor ~4, and site-directed mutagenesis studies indicated that they result from distinct rates for his26 vs his33 binding to the heme.

Because the heme-histidine kinetics should be the same for both samples,<sup>20</sup> we impose the constraints  $\tau_5 = \tau_2$  and  $\tau_6 = \tau_3$ . This fit gives good agreement with the data (Figure 6) and generates the time constants  $\tau_1 \approx 5.2 \,\mu\text{s}$ ,  $\tau_2 \approx 41 \,\mu\text{s}$ ,  $\tau_3 \approx 334$  $\mu$ s, and  $\tau_4 \approx 4.6$  ms for the unmodified cytochrome c and  $\tau_5 =$  $\tau_2 \approx 41 \ \mu s$ ,  $\tau_6 = \tau_3 \approx 334 \ \mu s$ , and  $\tau_7 \approx 3.7 \ ms$  for the carboxymethylated sample. The 41 and 334 µs processes are presumably the formation of the heme-histidine complex (for which  $\tau \approx 120 \ \mu s$  in the analysis of Jones at pH 6.5, T=40 $^{\circ}$ C), while the  $\sim$ 3–5 ms phase is due to heme–CO recombination. Evidently carboxymethylation of the cytochrome c eliminates the fourth,  $\sim$ 5  $\mu$ s phase. Since the only effect of carboxymethylation should be to prevent the binding of met80 and met65 to the heme, we conclude that the binding of methionine to the heme is indeed a microsecond process, as proposed by Jones et al.

Heme Peptide. Measurement of the ligand-binding rate of the heme peptide presents two problems. The first is that the unliganded peptide in solution forms aggregates which dissociate very slowly.<sup>13</sup> This aggregation is easily detected by circular dichroism measurements at equilibrium, since the heme—heme intermolecular interaction within the aggregate produces an exciton splitting of the heme Soret band. Under our solution conditions, aggregation is not observed once the heme binds a ligand such as histidine, methionine, or carbon monoxide.

A second problem, however, is that with a 10 ns laser pulse the probability of photolyzing the heme—histidine and heme—methionine complexes is found to be small, roughly  $\sim 6\%$  for heme—methionine and  $\leq 1\%$  for heme—histidine. Although the photolysis yield for heme—methionine is large enough to allow measurement of the bimolecular rebinding rate, <sup>14</sup> the yield for heme—histidine is too small. The small yield may indicate that photoexcitation of the heme—histidine electronic system results in a rapid radiationless transition to the ground state, rather than photodissociation. It may also indicate that the proximal histidine (his18) rather than the extrinsic ligand is dissociated by the light pulse and rebinds during the 10 ns photolysis pulse.

Because direct photolysis of histidine and methionine was inefficient, we undertook a ligand-replacement experiment. Here a small amount of an additional, tighter-binding but photodissociable ligand L is added to a heme peptide solution containing excess histidine or methionine. L replaces the free amino acid as the dominant heme ligand at equilibrium but is readily photodissociated by a 10 ns laser pulse. The laser pulse displaces L, and the free amino acid, present in high concentrations ( $\sim 10-200$  mM), binds to the heme. This generates an easily observable spectral change. At longer times the ligand dissociates from the heme and is replaced by L, again producing a spectral changes, both the dissociation and binding rates for methionine and histidine may be calculated.

We found that n-butyl isocyanide (BuNC) and carbon monoxide (CO) are both suitable choices for the photolabile ligand L in this experiment. Equilibrium circular dichroism studies of heme peptide samples prepared with CO or BuNC (in 5.6 M GuHCl at pH 6.5) confirmed that the peptide was not aggregated. We will not present here the photolysis data from all experiments.<sup>21</sup> Instead we discuss only the data obtained for the histidine-CO and methionine-CO samples. Figure 7 shows the SVD of the photolysis spectra for hemehistidine-CO at T = 40 °C. Two major phases are evident in the spectral changes, and the rates associated with these phases are related to the bimolecular heme-binding rates for histidine and CO and the dissociation rate for histidine (see below). Figure 8 shows the SVD of the photolysis spectra for hememethionine-CO at T = 40 °C. Again the rates of the two major processes are related to the ligand binding and dissociation rates.

Figure 9 shows the Soret absorption spectrum of a heme–peptide/methionine/CO sample at equilibrium and after photolysis by a 10 ns laser pulse. Figure 9 also shows representative photoproduct spectra for a methionine—heme—CO sample. Immediately after photodissociation of the heme—CO complex, the sample exhibits a broad, five-coordinate heme spectrum centered near 425 nm. As expected, this is a deoxyheme spectrum like that of deoxymyoglobin, in which the proximal histidine is the single axial ligand of the heme iron. This photoproduct evolves with time to form the intermediate ( $t \sim 10 \, \mu s$ ) spectrum shown. Comparison to the equilibrium spectra in Figure 2 indicates that the sample at 10  $\mu s$  represents an equilibrium between the methionine—heme complex and the deoxyheme, *i.e.* that *N*-acetylmethionine has bound to the

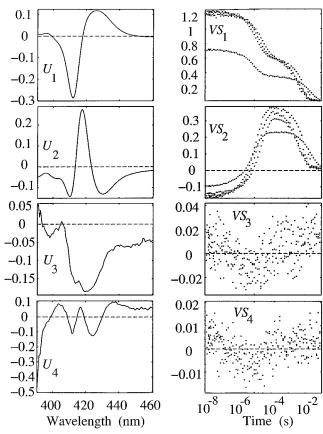


Figure 7. SVD of the time-resolved difference spectra of heme peptide/ CO/N- $\alpha$ -acetyl-L-histidine in 5.6 M GuHCl, pH 6.5, at T = 40 °C. The spectra for four samples (at histidine concentrations 20, 50, 100, 200 mM) are analyzed together so that a single set of component spectra  $U_i(\lambda)$  is obtained. The first two amplitudes (i.e. first two columns of the matrix VS)  $VS_1(t)$  and  $VS_2(t)$  are almost sufficient to describe all the observed spectral changes following photolysis; the third and higher amplitudes are  $\leq 2-3\%$  as large and are dominated by noise.

deoxyheme. At longer times  $t \sim 1-100$  ms, the methionine ligand is replaced by CO and the equilibrium spectrum is recovered. A similar sequence of spectral changes is observed with the histidine-heme-CO samples.

Calculation of the Bimolecular Reaction Rates. The rates of methionine and histidine bimolecular binding and dissociation are obtained from the SVD amplitudes. We use the following scheme to describe the bimolecular heme ligation reactions that follow photolysis

heme + CO 
$$\xrightarrow{k_1}$$
 heme-CO (1)  
heme + L  $\xrightarrow{k_2}$  heme-L

Here L represents N-acetylmethionine or N-acetylhistidine,  $k_1$ and  $k_2$  are bimolecular binding rates, and  $k_3$  is a unimolecular dissociation rate. If  $L_0 \gg H_0$  and  $C_0 \gg H_0$ , where  $L_0 \approx 10$ 200 mM,  $C_0 \sim 1$  mM, and  $H_0 \approx 100 \ \mu M$  are the total concentrations of free amino acid, CO, and heme respectively, the kinetic equations associated with eq 1 can be greatly simplified. The time course of each species concentration after photolysis at t = 0 is then well approximated by a sum of two exponential functions, with rates  $1/\tau_1$  and  $1/\tau_2$  that are simply related<sup>22</sup> to the rates  $k_1$ ,  $k_2$ , and  $k_3$ . If in addition  $k_2L_0 + k_3 \gg$  $k_1C_0$  (i.e. the heme—amino acid system attains equilibrium much faster than does the heme-CO system),  $\tau_1$  and  $\tau_2$  further simplify to

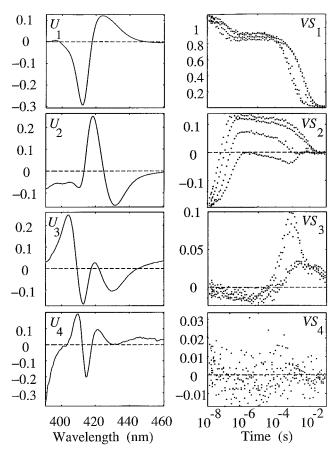


Figure 8. SVD of the time-resolved difference spectra of heme peptide/ CO/N-acetyl-L-methionine in 5.6 M GuHCl, pH 6.5, at T = 40 °C. As in Figure 7, the spectra for four samples (at methionine concentrations 10, 25, 50, 100 mM) are analyzed together so that a single set of component spectra  $U_i(\lambda)$  is obtained.

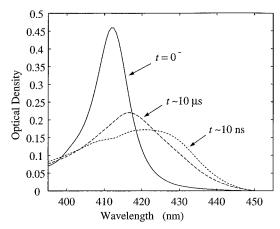
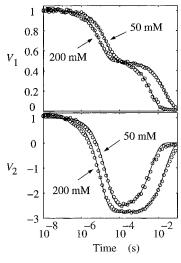


Figure 9. Absorption spectra of a heme peptide/CO/N-acetyl-Lmethionine sample before and after photolysis by a 10 ns laser pulse. The initial spectrum (solid curve) is essentially that of the CO-heme complex, while the immediate photoproduct (short dashes) exhibits the high-spin spectrum of deoxyheme (cf. Figure 2). Free methionine binds to the deoxyheme, and by  $\sim 10 \,\mu s$  (long dashes) the methionine—heme complex is at equilibrium with the deoxyheme. At longer times, CO displaces methionine as the heme ligand. The sample contained 100 μM heme peptide (microperoxidase) in 5.6 M GuHCl, pH 6.5, 1 atm CO, 10 mM N-acetyl-L-methionine, at T = 40 °C.

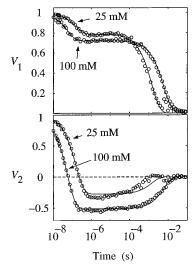
$$\tau_1 \approx 1/(k_3 + k_2 L_0)$$

$$\tau_2 \approx (1 + k_2 L_0 / k_3) / (k_1 C_0)$$
(2)

The first and second SVD amplitudes in the experimental data,  $v_1(t)$  and  $v_2(t)$  (Figures 7 and 8), are dominated by two phases,

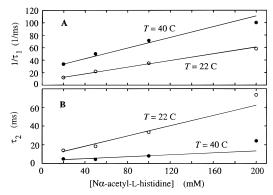


**Figure 10.** SVD amplitudes  $V_1(t)$  and  $V_2(t)$  for time-resolved difference spectra of heme peptide/CO/N-α-acetyl-L-histidine samples after photolysis. Data are shown for 50 and 200 mM histidine. For each ligand concentration, VS1 and VS2 were simultaneously fit to biexponential functions (solid curves) to obtain the time constants  $\tau_1$  and  $\tau_2$  that are characteristic of the ligand-binding kinetics (see text). In the figure, the amplitudes  $V_1$  and  $V_2$  are normalized to their  $t \approx 0$  values for clarity.

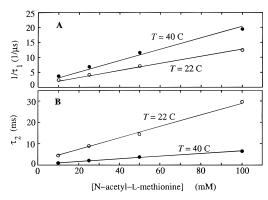


**Figure 11.** SVD amplitudes  $V_1(t)$  and  $V_2(t)$  for time-resolved difference spectra of heme peptide/CO/N-acetyl-L-methionine samples after photolysis. Data are shown for 25 and 100 mM methionine. Solid curves show the fit to the biexponential functions, as in Figure 10. In the figure, the amplitudes are normalized to their t = 0 values for clarity.

i.e. by a fast and a slow relaxation.<sup>23</sup> We fit these SVD amplitudes to a sum of exponentials, i.e. at each ligand concentration and temperature, column i (i = 1, 2) of the matrix **VS** is fit to a biexponential,  $vs_i = A_i \exp(-t/\tau_1) + B_i \exp(-t/\tau_2)$  $\tau_2$ ). Figures 10 and 11 show representative fits to  $vs_1$  and  $vs_2$ for histidine and methionine experiments. Figures 12 and 13 show that the times  $\tau_1$  and  $\tau_2$  associated with the two phases vary inversely  $(\tau_1)$  or linearly  $(\tau_2)$  with the amino acid concentration, as expected from eq 2. For each ligand (methionine and histidine), we obtain four values of  $\tau_1$  and four values of  $\tau_2$  by studying rebinding at four different ligand concentrations. A three-parameter fit to the  $1/\tau_1$  and  $\tau_2$  data at each temperature, through eq 2, therefore gives  $k_1C_0$ ,  $k_2$ , and  $k_3$ . Figures 12 and 13 show the fit to eq 2 for the methionine and histidine experiments. Table 1 gives the results of both fits. The fits determine the methionine binding (and dissociation) rates to within  $\sim$ 5%; the histidine rates are determined to within  $\sim$ 10% (at T = 22 °C) or  $\sim$ 25% (at T = 40 °C). (These



**Figure 12.** (A)  $1/\tau_1$  and (B)  $\tau_2$  vs histidine concentration, from the biexponential fits to the SVD data (Figure 10) at T = 22 °C (open circles) and 40 °C (solid circles). At each temperature, a three-parameter fit to eq 2 generates the solid lines through the  $1/\tau_1$  and  $\tau_2$  data and gives the histidine bimolecular binding and dissociation rates (and the CO-heme bimolecular binding rate).



**Figure 13.** (A)  $1/\tau_1$  and (B)  $\tau_2$  vs methionine concentration, from the biexponential fits to the SVD data (Figure 11) at T = 22 °C (open circles) and 40 °C (solid circles). As in Figure 12, the solid lines are fits to eq 2, giving the methionine-heme bimolecular binding and dissociation rates.

error margins give a 50% increase in the sum of squares in the fits to  $1/\tau_1$  and  $\tau_2$ .) In all cases our initial assumption  $k_2L_0$  +  $k_3 \gg k_1 C_0$  is satisfied by the observed rates  $(k_1 C_0 \le 2 \times 10^3/\text{s})$ .

Alternative Determination of the Equilibrium Constants. Above we used the kinetics of spectral changes to evaluate the ligand-binding rates. Although the changes in the kinetics that result from changes in ligand concentration are consistent with

the model of eq 1 and Figure 1, the purely kinetic analysis does not directly test that the observed spectral changes are as expected from the scheme of eq 1. While Figure 9 does demonstrate that the transient spectra appearing in the hemepeptide/methionine/CO sample are as expected for eq 1, we confirmed our interpretation of the data by modeling the timedependent heme-peptide/methionine/CO difference spectra in detail. This procedure, which is explained in Appendix A, generates absolute spectra for the transient heme complexes that form after photodissociation of the heme-CO complex and confirms the equilibrium binding constant obtained above from fits to the kinetics. The absolute spectra that we calculate closely resemble (Figure 2) the heme-CO and heme-methionine spectra, as expected. The equilibrium constants obtained by this procedure,  $K \approx 97 \text{ M}^{-1}$  at T = 22 °C and  $K \approx 74 \text{ M}^{-1}$ at 40 °C, are close to the values  $K \approx 140$  and 150 M<sup>-1</sup>, respectively, obtained from the kinetics.

#### Discussion

Jones et al.1 studied the rates of intrachain heme-ligand reactions that occur in the unfolded state of cytochrome c. In

TABLE 1: Measured and Calculated Rates in 5.6 M GuHCl, pH 6.5a

rate	reaction	$T = 22  ^{\circ}\text{C}$	<i>T</i> = 40 °C
$k_{ m on}^{ m bi}(k_2) \ k_{ m of}^{ m bi}(k_3) \ k_{ m on}^{ m bi}(k_2) \ k_{ m of}^{ m bi}(k_3) \ k_{ m gem}^{ m M} \ k_{ m gem}^{ m H} \ k_{ m on}^{ m uni} \ k_{ m of}^{ m uni}$	heme + his binding heme-his dissoc heme + met binding heme-met dissoc heme + met geminate binding heme + his geminate binding heme + met80/met65 binding heme-met80/met65 dissoc	$(2.7 \pm 0.2) \times 10^{5} \text{M} \cdot \text{s}$ $(7.2 \pm 0.8) \times 10^{3} \text{/s}$ $(1.19 \pm 0.05) \times 10^{8} \text{/M} \cdot \text{s}$ $(1.04 \pm 0.06) \times 10^{6} \text{/s}$ $\geq 4.3 \times 10^{10} \text{/s}$ $\approx 6.3 \times 10^{7} \text{/s}$	$(4.2 \pm 1.0) \times 10^{5}/\text{M} \cdot \text{s}$ $(2.6 \pm 0.6) \times 10^{4}/\text{s}$ $(1.93 \pm 0.09) \times 10^{8}/\text{M} \cdot \text{s}$ $(1.28 \pm 0.07) \times 10^{6}/\text{s}$ $\geq 1.0 \times 10^{11}/\text{s}$ $\approx 9.7 \times 10^{7}/\text{s}$ $2.5 \times 10^{4}/\text{s} \text{ (Jones } et al.^{1})$ $2.5 \times 10^{5}/\text{s} \text{ (Jones } et al.^{1})$
$k_{ m on}^{ m  uni} \ k_{ m off}^{ m  uni}$	heme + his33/his26 binding heme-his33/his26 dissoc		$2.5 \times 10^3$ /s (Jones <i>et al.</i> <sup>1</sup> ) $5.9 \times 10^3$ /s (Jones <i>et al.</i> <sup>1</sup> )

<sup>&</sup>lt;sup>a</sup> Indicated errors represent a 50% increase in the sum of squares in the joint fit to  $1/\tau_1$  and  $\tau_2$  vs ligand concentration.

order to identify the separate roles that intrachain diffusion and heme-ligand covalent bond formation play in these unimolecular rates, we have studied the bimolecular heme-ligand reaction rates in an analogous system—the binding of N- $\alpha$ acetyl-L-histidine and N-acetyl-L-methionine to the heme of microperoxidase. The bimolecular rate of binding to the deoxyheme is ~400 times faster for methionine than for histidine, suggesting that bimolecular and unimolecular methionine-heme binding may be diffusion-limited, while histidine-heme binding may be reaction-rate limited. In what follows we quantify this result and calculate the rate of intrachain diffusion of the unfolded cytochrome c molecule. From this we estimate the rate at which the smallest internal loops should form in an unfolded polypeptide. This may provide an upper limit for the rate of protein folding.

Model for Ligand Binding. In order to relate bimolecular ligand binding to intrachain diffusion and binding, we use the two-step model of Figure 1. This model can describe hemeligand binding and any other complex formation. For binding to occur, the heme and ligand (either in solution or on the polypeptide chain) must first diffuse together, at a rate  $k_{D+}$ , to form an "encounter complex." The encounter complex is identical to the geminate pair formed by photodissociation of the covalent complex: it either dissociates (at rate  $k_{D-}$ ) or else forms the covalent complex (at rate  $k_{\rm gem}$ ). The heme-ligand bond also breaks spontaneously at a rate  $k_{\rm diss} \ll k_{\rm gem}$ . In the steady-state approximation, this model gives an overall binding rate

$$k_{\rm on} = (k_{\rm gem} k_{\rm D+})/(k_{\rm gem} + k_{\rm D-})$$

or

$$1/k_{\rm on} = 1/k_{\rm D+} + (k_{\rm D-}/k_{\rm D+})(1/k_{\rm gem}) = 1/k_{\rm D+} + 1/(Kk_{\rm gem})$$
 (3)

for either bimolecular or intrachain binding. Here  $K \equiv k_{\rm D+}/$  $k_{\rm D-}$  is the equilibrium constant for formation of the encounter complex. Although the model of Figure 1 is a simple representation of ligand binding, it is not an oversimplification. For both the bimolecular and unimolecular cases, the rate  $k_{\rm on}$ in eq 3 been shown to be equivalent to the rate calculated from a detailed theory of partially diffusion-controlled reactions based on solution of the Smoluchowski equation.<sup>9,24</sup> We assume that  $k_{\text{gem}}$  is the same for both bimolecular and intrachain reaction with the heme, while K and the rates  $k_{D+}$  and  $k_{D-}$  must depend on whether the reacting ligand is attached to the polypeptide chain. With this model, we can use eq 3 to calculate the chain diffusion rate  $k_{D+}^{uni}$  from the experimentally measured bimolecular and unimolecular rates  $k_{\rm on}$ . From the bimolecular form of eq 3

$$1/k_{\text{gem}} = K^{\text{bi}} (1/k_{\text{on}}^{\text{bi}} - 1/k_{\text{D+}}^{\text{bi}})$$

while from the unimolecular form

$$1/k_{\rm D+}^{\rm uni} = 1/k_{\rm on}^{\rm uni} - 1/K^{\rm uni}k_{\rm gem}$$

Therefore

$$1/k_{D+}^{\text{uni}} = 1/k_{\text{on}}^{\text{uni}} - (K^{\text{bi}}/K^{\text{uni}})(1/k_{\text{on}}^{\text{bi}} - 1/k_{D+}^{\text{bi}})$$
 (4)

In order to calculate  $k_{\rm D+}^{\rm uni}$  from eq 4, we can obtain  $k_{\rm on}^{\rm uni}$  and  $k_{\rm on}^{\rm bi}$  from experiment and  $k_{\rm D+}^{\rm bi}$  and  $K^{\rm bi}/K^{\rm uni}$  from theory. For a bimolecular reaction between spherically symmetric neutral molecules,  $k_{D+}^{bi}$  is given by the Smoluchowski rate,  $k_{D+}^{bi}$  =  $4\pi D_0 a$ , where  $D_0$  is the relative diffusion constant and a is the reaction radius. However, a ligand can access the reactive site of a heme on a peptide chain only from the distal side, since his 18 occupies the proximal side of the heme. We expect that this limited solid angle of access will reduce both K and  $k_{D+}$ roughly by a factor of 2 without affecting  $k_{\rm D}$ . The fact that the iron atom of the heme is partially embedded in the porphyrin ring further reduces its accessibility to the ligand. Following Miers et al., 25 we account for these effects by using  $k_{D+}^{bi} =$  $2\pi D_0 a$  in eq 4 and then allowing the reaction radius a to be a parameter that is determined empirically. Miers et al., who modeled binding of carbon monoxide to protoheme, found that  $a \approx 1.5 \text{ Å}$  gives the best agreement with experiment.<sup>25</sup>

Equilibrium Constant for the Encounter Complex. Kbi in eq 4 is easily calculated. The equilibrium constant K for forming a heme-ligand reaction complex is given by

$$K = \int_{v} d\mathbf{r} \exp(-U(\mathbf{r})/k_B T) / \int_{v} d\mathbf{r} \exp(-U(\mathbf{r})/k_B T)$$
 (5)

where **r** is the heme-ligand separation,  $U(\mathbf{r})$  is the potential of mean force between the heme and ligand, v is the reaction volume (i.e. the volume of the encounter complex), and V is the volume accessible to the free ligand. For a bimolecular reaction, V is the inverse of the ligand concentration. If  $U \approx$ 0, we have K = v/V. To obtain the bimolecular equilibrium constant  $K^{bi}$ , we divide K by the ligand concentration, giving  $K^{\rm bi} = v$ . One may then reasonably estimate  $v \approx (4/3)\pi a^3$ , although since only one side of the heme is accessible to the ligand we use  $K^{\text{bi}} = v \approx (2/3)\pi a^3$ .

Calculation of Kuni requires some assumptions about equilibrium properties of the unfolded polypeptide. When eq 5 is applied to the unimolecular case, V represents the total volume available to the ligand, and U reflects the conformations adopted by the segment of polypeptide connecting heme and ligand. We assume that this segment resembles a Gaussian random coil,<sup>26</sup> for which  $U(\mathbf{r})/k_BT = 3r^2/(2\langle r^2 \rangle)$ . For small reaction volumes ( $v \ll \langle r^2 \rangle^{3/2}$ ),  $K^{\text{uni}} = v / \int d\mathbf{r} \exp(-U(\mathbf{r})/k_B T) =$  $v/(2\pi\langle r^2\rangle/3)^{3/2}$ . The ratio appearing in eq 4 is then  $K^{\rm bi}/K^{\rm uni}\approx$  $(2\pi \langle r^2 \rangle/3)^{3/2}$ . The reaction volume v disappears from this relation, so that uncertainty in v does not directly affect our calculation of  $k_{\rm D+}^{\rm uni}$ .

To calculate  $K^{\text{bi}}/K^{\text{uni}}$  we therefore need to estimate  $\langle r^2 \rangle$ . For a Gaussian segment of polymer, the mean-squared separation  $\langle r^2 \rangle$  of the end points depends on the number n of residues as

$$\langle r^2 \rangle = C_n n^{2\nu} l^2$$

Here l=3.8 Å is the distance between consecutive  $C_{\alpha}$  atoms on the chain,  $C_n$  is Flory's characteristic ratio,  $^{26}$  and the exponent  $\nu$  is determined by the strength of chain—solvent interactions relative to the effects of intrachain interactions and the excluded volume of the chain. In a "good solvent" favorable chain—solvent interactions and intrachain interactions together expand the polymer chain, giving  $\nu \approx ^3/_5$ . In a poor solvent the chain is compacted, with  $\nu = ^1/_3$ . In a " $\theta$  solvent", chain—solvent interactions precisely compensate for intrachain interactions and  $\nu = ^1/_2$  as in the unperturbed chain.

Damaschun et al. 27 recently used dynamic light scattering to measure the Stokes radius  $R_S$  of a variety of proteins unfolded in 6 M GuHCl at T=20 °C. They calculated the radius of gyration  $R_{\rm G}$  from  $R_{\rm S}$  through the relation  $R_{\rm G}=1.6R_{\rm S}$  and  $\langle r^2\rangle$ from  $R_{\rm G}$  through the relation  $\langle r^2 \rangle = 6R_{\rm G}^2$  (which is accurate as n becomes large). Those authors found the characteristic ratio  $C_n$  to be essentially independent of both amino acid sequence and chain length n, i.e.  $C_n \approx C_\infty \approx 8.1 \pm 0.5$ , for  $100 \le n \le$ 700. They also found that  $\nu \approx 0.50 \pm 0.02$  for proteins unfolded under these conditions, implying that 6 M GuHCl acts as a  $\theta$ solvent for many proteins.<sup>28</sup> Therefore, under their experimental conditions, excluded volume does not affect the size of the unfolded chain. Since our work and that of Jones et al. were performed at slightly lower denaturant concentration (5.6 M) but slightly higher T (40 °C), we assume that near- $\theta$  conditions exist also in our experiment, so that the mean-square end-toend length of a chain of n amino acid residues under our solvent conditions is approximately

$$L^2 \equiv \langle r^2 \rangle \approx C_n n l^2$$

with  $C_n \approx 8$  for n large and l = 3.8 Å. If this relation also describes the mean-square distance between internal (*i.e.* not just the N- and C-termini) residues on the chain, we expect  $L \approx 74$  Å for met65—his18 and  $L \approx 85$  Å for met80—his18. We assume that  $C_{62}$  is close to  $C_{100} \approx C_{\infty} \approx 8$ . (For homopolypeptides as well as for many realistic model chains with  $C_{\infty} \approx 5-10$ ,  $C_n$  has already attained a value within  $\sim 10\%$  of  $C_{\infty}$  once<sup>26</sup>  $n \geq \sim 50$ .) With this value of  $\langle r^2 \rangle$  we can calculate the ratio  $K^{\text{bi}}/K^{\text{uni}}$  of eq 4.

Calculation of the Intrachain Diffusion Rates. We estimate that the mean-squared distance between met80 and the heme (at his18) is  $L^2 \approx (85 \text{ Å})^2$ , giving  $K^{\text{bi}}/K^{\text{uni}} \approx 1.9 \times 10^{-18} \text{ cm}^3 = 1.1 \times 10^3 \text{ M}^{-1}$ . Since  $1/k_{\text{on}}^{\text{uni}} \approx 40 \ \mu \text{s}$  and  $k_{\text{on}}^{\text{bi}} \approx 1.9 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$  at  $T = 40 \ ^{\circ}\text{C}$ , eq 4 gives  $1/k_{\text{D}+}^{\text{uni}} \approx 34-40 \ \mu \text{s}$ . The range of values for  $k_{\text{D}+}^{\text{uni}}$  reflects the range of possible values for  $1/k_{\text{D}+}^{\text{bi}}$  (i.e.  $0 \leq 1/k_{\text{D}+}^{\text{bi}} \leq 1/k_{\text{on}}^{\text{bi}}$ ). We therefore conclude that diffusion brings together the ends of this segment of polypeptide on a time scale of  $34-40 \ \mu \text{s}$ . A similar calculation for met65, also using  $1/k_{\text{on}}^{\text{uni}} \approx 40 \ \mu \text{s}$ , gives  $1/k_{\text{D}+}^{\text{uni}} \approx 36-40 \ \mu \text{s}$  at  $T = 40 \ ^{\circ}\text{C}$ .

We therefore calculate from the methionine data a time scale of  $35-40~\mu s$  for relative diffusion of two points separated by 50-60 residues in an unfolded polypeptide. This calculation is based on the model of Figure 1, the measured unimolecular and bimolecular binding rates for methionine, and the assumption that the statistics of the chain are nearly Gaussian. It does not require a microscopic model for  $k_{\rm D+}^{\rm bi}$ , intrachain diffusion, or the encounter complex. The rate  $k_{\rm D+} \approx (2.5-3) \times 10^4~{\rm s}^{-1}$  should therefore be a good indication of the characteristic time

scale required for formation of intrachain contacts in any similar unfolded polypeptide as it attempts to fold in a denaturing solvent. Below we use this rate to estimate the time scale for formation of the smallest polypeptide loops.

It is more difficult to obtain information about chain diffusion from the his33 binding rate. If we assume that the chain segment connecting the heme with his33 acts as a Gaussian random coil with  $C_n \approx 8$ , we expect  $L \approx 42$  Å, which gives  $K^{\rm bi}/K^{\rm uni} \approx 2.2 \times 10^{-19}$  cm³ for his33—heme binding. With  $k_{\rm on}{}^{\rm bi} \approx 4.3 \times 10^5$  M<sup>-1</sup> s<sup>-1</sup> (Table 1) and  $1/k_{\rm on}{}^{\rm uni} \approx 400$   $\mu$ s, we have  $(K^{\rm bi}/K^{\rm uni})(1/k_{\rm on}{}^{\rm bi} - 1/k_{\rm D+}{}^{\rm bi}) \approx 310$   $\mu$ s. Equation 4 then gives  $1/k_{\rm D+}{}^{\rm uni} \approx 400$  –  $310 \approx 90$   $\mu$ s. This estimate is insensitive to  $k_{\rm D+}{}^{\rm bi}$ , since  $k_{\rm D+}{}^{\rm bi} \approx 2\pi D_0 a \approx 5.7 \times 10^{-13}$  cm³/s =  $3.4 \times 10^8$  M<sup>-1</sup> s<sup>-1</sup> is ~1000 times larger than  $k_{\rm on}{}^{\rm bi}$ . Here we have used  $D_0 \approx 0.6 \times 10^{-5}$  cm²/s (as estimated²9 for 5.6 M GuHCl at 40 °C) and  $a \approx 1.5 \times 10^{-8}$  cm.²5 The calculated diffusion time,  $1/k_{\rm D+}{}^{\rm uni} \approx 90$   $\mu$ s, is unexpectedly long. From our result for met80 and met65, we expect that the his33 residue should diffuse to the heme within a time less than 40  $\mu$ s.

There are three possible explanations for this result. First, the estimate for  $k_{\rm D+}^{\rm uni}$  is sensitive to the values of  $K^{\rm bi}/K^{\rm uni}$  and  $k_{\rm on}^{\rm bi}$ , while the experimental uncertainty in  $k_{\rm on}^{\rm bi}$  ( $\Delta k_{\rm on}^{\rm bi}\sim 25\%$  at T=40 °C; see Table 1) is large. A diffusion time of  $\sim 40$   $\mu \rm s$  would be calculated from eq 4 if  $k_{\rm on}^{\rm bi}$  were  $\sim 16\%$  smaller than our measured value. This variation lies within our uncertainty.

Second, it is likely that the equilibrium constant  $K^{\rm uni}$  is given inaccurately by the Gaussian-chain value, since the 15-residue length of polypeptide connecting his33 with the heme of cytochrome c is probably too short to exhibit the statistics of a Gaussian chain. Damaschun et al.'s finding<sup>27</sup>  $C \approx 8$  indicates that  $\sim 8$  residues in the unfolded polypeptide comprise one segment of the equivalent random coil: the 15-residue length of chain between the heme and his33 therefore represents barely two "effective" segments and therefore should not exhibit Gaussian coil statistics. If  $K^{\rm uni}$  were  $\sim 16\%$  smaller than the Gaussian-chain value, a diffusion time of  $\sim 40~\mu s$  would be calculated from eq 4.

Third, the histidine geminate rate  $k_{gem}$  measured in the bimolecular experiment may not equal the  $k_{gem}$  that is relevant to the folding of cytochrome c. Our experiments and those of Jones et al. were performed at pH 6.5, near the pK of free histidine. Because protonation of the histidine imidazole hinders its binding to the heme iron, a small difference between the pKof free N- $\alpha$ -acetyl-L-histidine and that of the histidine residue on the polypeptide would result in a different degree of protonation in the bimolecular and intrachain binding experiments and thus alter the measured geminate rate  $k_{\text{gem}}$ . If  $k_{\text{gem}}$ were  $\sim 15-20\%$  smaller for a histidine residue in the native protein than for the free amino acid, owing to a slightly higher pK, then the chain diffusion time calculated from eq 4 would drop to a value  $\sim 10 \ \mu s$ . Because of these uncertainties, the histidine-heme binding rate provides less information about intrachain diffusion rates than does the methionine-heme binding rate.

Calculation of the Geminate Reaction Rates. *Histidine*. Although the histidine binding rates were not as well determined as the methionine rates, we can calculate the geminate reaction rates for binding of these ligands. For histidine, eq 3 gives  $1/k_{\rm gem} = K^{\rm bi}(1/k_{\rm on}^{\rm bi} - 1/k_{\rm D+}^{\rm bi}) \approx K^{\rm bi}/k_{\rm on}^{\rm bi}$ . The bimolecular equilibrium constant is essentially the reaction volume, or  $K^{\rm bi}$  =  $(2/3)\pi a^3 \approx 7.1 \times 10^{-24} \, {\rm cm}^3 = 4.3 \times 10^{-3} \, {\rm M}^{-1}$ . Thus  $k_{\rm gem} \approx 1.0 \times 10^8 \, {\rm s}^{-1}$  for histidine—heme binding at 40 °C. This result is insensitive to the uncertainty in  $k_{\rm D+}^{\rm bi}$  because  $k_{\rm D+}^{\rm bi} \gg k_{\rm on}^{\rm bi}$ .

From the histidine rate  $k_{\rm gem}$  and the observed dissociation rate  $k_{\rm off}^{\rm bi} \approx 2.6 \times 10^4 \, {\rm s}^{-1}$  (at T=40 °C), we can roughly estimate the his33—heme unimolecular dissociation rate,  $k_{\rm off}^{\rm uni}$ . In the model of Figure 1, the total spontaneous dissociation rate is

$$k_{\rm off} = (k_{\rm diss}k_{\rm D-})/(k_{\rm D-} + k_{\rm gem})$$
 (6)

in the steady-state approximation. From  $k_{\rm D-}=k_{\rm D+}/K$  we have  $k_{\rm D-}^{\rm bi}=3D_0/a^2\approx 8\times 10^{10}~{\rm s}^{-1}$ . Since  $k_{\rm D-}^{\rm bi}\gg k_{\rm gem}$ , we find  $k_{\rm off}^{\rm bi}\approx k_{\rm diss}$ , or  $k_{\rm diss}\approx 2.5\times 10^4~{\rm s}^{-1}$ . If  $k_{\rm D-}^{\rm uni}\gg k_{\rm gem}$ , we estimate the unimolecular dissociation rate as  $k_{\rm off}^{\rm uni}\approx k_{\rm diss}\approx (40~\mu{\rm s})^{-1}$ . Experimentally, Jones  $et~al.^1$  observed  $k_{\rm off}^{\rm uni}\approx (170~\mu{\rm s})^{-1}$  for his33. Because our experiment measures  $k_{\rm off}^{\rm bi}$  only indirectly, and because Jones et~al.'s determination of  $k_{\rm off}^{\rm uni}$  contains some uncertainty, we consider the agreement to be satisfactory.

We find the overall equilibrium constant for formation of the bimolecular covalent complex to be  $k_{\rm on}^{\rm bi}/k_{\rm off}^{\rm bi} \approx 17~{\rm M}^{-1}$  at 40 °C and 38  $M^{-1}$  at 23 °C. These values are smaller than the reported equilibrium constant for imidazole binding to the ferrous heme of dicarboxymethylated cytochrome c: Schejter et al.<sup>30</sup> found  $K_{\rm imid} \approx 870~{\rm M}^{-1}$  at pH 7.4 (25 °C). Their higher value for  $K_{\text{imid}}$  may in part reflect deprotonation of the imidazole at the higher pH. For example, if the pK of the histidine residue is approximately 7.0, Schejter's higher pH should result in a  $\sim$ 2.5–3 times smaller degree of protonation than exists in ours. In the simplest picture this could account for a  $\sim$ 2.5-3 times greater heme-imidazole affinity than would be observed in our experiment. Furthermore, there may be noncovalent interactions between the imidazole and residues lining the heme pocket that increase the affinity of cytochrome c over that of the heme peptide.

**Methionine.** The bimolecular binding of methionine is much faster than that of histidine. For methionine at 40 °C, eq 3 gives a lower limit  $k_{\rm gem} \geq 1 \times 10^{11} \, {\rm s}^{-1}$ , where the uncertainty arises from uncertainty in  $D_0$ . The lower limit applies if  $k_{\rm D+}{}^{\rm bi} \gg k_{\rm on}{}^{\rm bi}$ . The estimate  $k_{\rm D+}{}^{\rm bi} = 2\pi D_0 a \approx 3.4 \times 10^8 \, {\rm M}^{-1} \, {\rm s}^{-1}$  gives  $k_{\rm gem} \approx 1 \times 10^{11} \, {\rm s}^{-1}$  at T = 40 °C. In the case of cytochrome c peroxidase, where the heme has methionine and histidine as axial ligands, it is believed that the methionine photodissociates and then rebinds with a half time of 35 ps.  $^{31}$ 

Our data on the direct photolysis of the methionine-heme complex (without CO present) allow us to estimate an upper limit for  $k_{\text{gem}}$ . At a methionine concentration of 100 mM (T =40 °C), the maximum amplitude of the observed spectral change after photolysis indicates that about 6-7% of the heme ligands are photodissociated and rebind bimolecularly. 14 Given that the photolysis laser generates a Gaussian pulse of width ~10 ns (full width at half-maximum) with intensity sufficient to photolyze the heme-ligand complex at a peak rate no greater than about  $^{32}$  2  $\times$  10 $^{8}$ /s, we can calculate the maximum degree of photolysis that would result from any value of  $D_0$  or  $k_{gem}$ . We integrated numerically the kinetic equations associated with Figure 1 for different values of  $k_{gem}$  (or, equivalently, values of  $D_0$ , since eq 3 fixes the relation between  $D_0$  and  $k_{\text{gem}}$ ), and we determined the maximum value of  $k_{gem}$  that is consistent with the observed bimolecular yield. This analysis shows that  $k_{\text{gem}}$ probably does not exceed  $\sim 1 \times 10^{12} \text{ s}^{-1}$  at  $T = 40 \,^{\circ}\text{C}$ . If  $k_{\text{gem}}$ were larger than this value, the diffusion constant  $D_0$  would have to be greater than  $\sim 10^{-5}$  cm<sup>2</sup>/s in order to allow sufficient ligand escape to the solvent. If  $2 \times 10^8$ /s were a severe overestimate of the photolysis rate, owing perhaps to a misalignment of the photolysis laser, then  $k_{gem}$  smaller than 1  $\times$  10<sup>12</sup>/s would be required to explain the data. Therefore the experimental data are most consistent with a geminate-binding rate  $k_{\rm gem}$  for methionine to heme (in 5.6 M GuHCl at T=40 °C) lying in the range  $6\times 10^{10}/{\rm s} \le k_{\rm gem} \le \sim 1\times 10^{12}/{\rm s}$ . The calculated upper limit is close to the rate for barrierless ligand binding to a deoxyheme,  $k_{\rm gem}^0 \approx 5\times 10^{11}/{\rm s}$ , as obtained from molecular dynamics simulations.<sup>33</sup>

We can use eq 6 and the observed dissociation rates  $k_{\rm off}$  to estimate the rate  $k_{\rm D}$ -uni for breakup of the heme-methionine encounter complex. Inserting  $k_{\rm D}$ -bi =  $3D_0/a^2$  and  $k_{\rm off}$ bi  $\approx 1.3 \times 10^6 \ {\rm s}^{-1}$  into eq 6 gives  $k_{\rm diss} = (1 + k_{\rm gem}/k_{\rm D}$ -bi) $k_{\rm off}$ bi  $\approx 2.9 \times 10^6 \ {\rm s}^{-1}$ . This result, together with  $k_{\rm off}$ uni  $\approx (4 \ \mu {\rm s})^{-1}$  and  $k_{\rm gem} \approx 1 \times 10^{11} \ {\rm s}^{-1}$ , in turn gives  $k_{\rm D}$ -uni  $\approx 9.4 \times 10^9 \ {\rm s}^{-1}$ . In the next section we use this rate to estimate the intrachain diffusion constant D.

**Dynamics of the Polypeptide Chain.** Residues met65 and met80 make contact with the heme at a rate  $k_{\rm D+}^{\rm uni} \approx (35-40~\mu s)^{-1}$ . It is difficult to compare this experimental rate with a theoretical value, since dynamical properties of realistic polymers are not well understood. The rate  $k_{\rm D+}^{\rm uni}$  that controls loop formation and intrachain reactions in peptides is described only by approximate theories. We describe here the approach of Szabo, Schulten, and Schulten<sup>9</sup> (SSS), who used a Smoluchowski-equation method to estimate  $k_{\rm D+}$  for a polymer whose ends move with a relative diffusion constant D. This theory provides an estimate for  $k_{\rm D+}$  for a Gaussian chain and, because  $K = k_{\rm D+}/k_{\rm D-}$  is known for this case, also gives  $k_{\rm D-}$ . From our measurement of  $k_{\rm D+}$ , we can use the SSS theory to estimate the end-to-end diffusion constant D, which has previously been measured by fluorescence energy transfer.  $^{34-36}$ 

For a sufficiently long chain, the equilibrium distribution W(r)of end-to-end distances r is well approximated by a Gaussian distribution; the theoretical problem is to determine the average time required for the two ends to approach to within a distance a of each other. That is, given a Gaussian distribution in r at time t = 0, we wish to know the mean first passage time for r  $\leq a$  to occur by diffusion of the chain. Rather than consider all the degrees of freedom of a diffusing polymer, SSS modeled the problem as diffusion of a particle (located at position r) on a harmonic potential energy surface. The potential surface is designed so that the equilibrium probability distribution in r for the particle is identical to W(r) for the polymer. If the particle is absorbed (i.e. undergoes a reaction) at rate  $k_{gem}$  when it reaches r = a, then SSS found that the rate of reaction (for  $a \ll L$ ) can be approximated by eq 3, with  $k_{\rm D+} = 3Da(6/\pi)^{1/2}$  $L^3$ . The equilibrium constant K is that expected for a freely jointed Gaussian chain,  $K \approx (a/L)^3 (6/\pi)^{1/2}$ , so that  $k_{D-} = k_{D+}/K$  $=3D/a^2$ . The assumption of a Gaussian distribution, although poor for chains of only a few segments, becomes quite adequate (as far as this theory is concerned) once the chain has more than about four segments.

Because the ligand can bind only to the distal side of the heme of cytochrome c, we expect that  $k_{\rm D+}$  and K are both roughly half as great as those estimated by the SSS theory. We therefore take  $k_{\rm D+}^{\rm uni} \approx 3Da(3/2\pi)^{1/2}/L^3$  and  $K^{\rm uni} \approx (a/L)^3(3/2\pi)^{1/2}$ . Using our finding that  $k_{\rm D+} \approx (2.5-2.8) \times 10^4/{\rm s}$  for met80, we can estimate D. From the SSS theory, with  $L \approx 85$  Å,  $D \approx k_{\rm D+}^{\rm uni} (2\pi/3)^{1/2}L^3/(3a) \approx 5.4 \times 10^{-7} {\rm cm}^2/{\rm s}$ . A similar calculation based on met65 instead of met80 gives  $k_{\rm D+}^{\rm uni} \approx 2.6 \times 10^4/{\rm s}$  and  $D \approx 3.5 \times 10^{-7} {\rm cm}^2/{\rm s}$ . Thus we obtain a fairly direct estimate,  $D \sim (4-6) \times 10^{-7} {\rm cm}^2/{\rm s}$ , for the rate of diffusion between end points of the unfolded chain in 5.6 M GuHCl at T = 40 °C. We note that this value is at least 1 order of magnitude smaller than the monomer value  $D_0 \sim 10^{-5} {\rm cm}^2/{\rm s}$  describing diffusion of a free amino acid molecule in H<sub>2</sub>O.

The SSS theory predicts that the encounter complex will dissociate at a rate  $k_{\rm D-}{}^{\rm uni}=k_{\rm D+}{}^{\rm uni}/K^{\rm uni}\approx 3D/a^2$ , which is similar

to the bimolecular result  $k_{\rm D-}^{\rm bi}=3D_0/a^2$ . Our estimate  $k_{\rm D-}^{\rm uni}\approx 9.4\times 10^9~{\rm s^{-1}}$ , obtained from the heme–methionine dissociation rates (see above), then implies an intrachain diffusion constant  $D=a^2k_{\rm D-}^{\rm uni}/3\approx 7\times 10^{-7}~{\rm cm^2/s}$ . The good agreement between our two estimates of D demonstrates that our analysis of the unimolecular and bimolecular reactions is self-consistent.

Only one other experimental technique has provided an estimate of D. Haas and co-workers $^{34-36}$  have used the nonradiative transfer of fluorescence energy to study the equilibrium conformations and end-to-end diffusion of polypeptides and other polymers. In these experiments, a donor and an acceptor group are attached to different positions on an unfolded chain: after excitation of the donor fluorescence, the rate of energy transfer from donor to acceptor varies strongly with the distance r separating the two groups (i.e. as  $1/r^6$ , in Förster theory). The fluorescence decay of the excited donor molecules is therefore sensitive both to the equilibrium distances between segments of the chain and to the relative rate of diffusion of these segments. Haas and co-workers<sup>34–36</sup> model the experimental fluorescence decay to obtain parameters describing the equilibrium distances and the diffusion. They assume that the probability distribution W(r) for the distance r between donor and acceptor obeys a Smoluchowski-like diffusion equation, with a Gaussian equilibrium distribution and a (usually) r-independent diffusion constant D.

This technique is effective for measuring equilibrium distributions W(r) for large molecules. However, one difficulty with the method is that the diffusion constant D obtained from the data analysis is correlated with W(r) in such a way that the uncertainty in D may be large. A short fluorescence lifetime for the donor may result either from rapid diffusion of the chain or from a short equilibrium separation of the donor-acceptor pair. This led to large uncertainty in Gottfried et al.'s estimate<sup>35</sup> of the rate of relative diffusion between residues 1 and 46 of bovine pancreatic trypsin inhibitor unfolded at 35 °C,  $D \sim (0-$ 16)  $\times$  10<sup>-7</sup> cm<sup>2</sup>/s. Nevertheless, the method has produced estimates of D that are consistent with our value  $D \sim (4-6) \times 10^{-6}$  $10^{-7}$  cm<sup>2</sup>/s, obtained by a very different technique. For the relative diffusion of residues 104 and 125 of reduced ribonuclease A unfolded in 6.7 M GuHCl at T = 22 °C, fluorescence energy transfer indicated<sup>36</sup>  $D \sim (1.5-8) \times 10^{-7}$  cm<sup>2</sup>/s.

Other Models for the Rate of Chain Diffusion. In their diffusion-collision model of protein folding, Karplus and Weaver<sup>7</sup> use a more simplified description of polypeptide chain dynamics. They consider an unfolded protein with two separate regions that can form transient structures or microdomains such as helices; these microdomains are modeled as diffusing spheres tethered by a frictionless string. By solving a diffusion equation for the relative separation of two particles, constrained by a maximum and a minimum possible separation (*i.e.* by the length of the string and the size of the spheres), Karplus and Weaver estimate the time scale for the spheres to diffuse into contact. The rate is<sup>7</sup>

$$\tau^{-1} \sim 3DR_{\min}/(R_{\max})^3 \tag{7}$$

Equation 7, which is the Smoluchowski rate for the reaction of two tethered particles, may be obtained by multiplying the Smoluchowski bimolecular rate  $4\pi DR_{\rm min}$  by the effective concentration of reactants, *i.e.* the inverse of the volume (4/3) $\pi R_{\rm max}^3$ . In order to apply this model to the relative diffusion of met80 and the heme of cytochrome c, we use a smaller bimolecular rate  $2\pi DR_{\rm min}$ , which predicts a reaction rate  $\tau^{-1} \sim (3/2)DR_{\rm min}/(R_{\rm max})^3$ . Then with  $R_{\rm max} \approx nl = 240$  Å,  $R_{\rm min} \approx a \approx 1.5$  Å, and  $\tau \approx 35~\mu s$ , we find  $D \sim 2 \times 10^{-5}~{\rm cm}^2~{\rm s}^{-1}$ , which is approximately three times larger than our estimate for the

diffusion constant of the free amino acid molecule ( $D_0 \approx 0.6 \times 10^{-5} \ {\rm cm^2 \ s^{-1}}$ ) and ~40 times larger than the D obtained in the SSS model. Whether or not this is a reasonable value for the rate of diffusion of a chain, it is important to note that eq 7 arises from a very different picture of chain diffusion than was used by SSS. Equation 7 is based on diffusion of a particle on a "flat" potential surface, in which the equilibrium probability distribution W(r) for the separation r does not tend toward a Gaussian; thus the model does not take into account the entropy of the string connecting the spheres, *i.e.* its tendency to form a random coil. The SSS model includes this contribution by placing the diffusing particle on a harmonic potential energy surface. Both models assume that a single constant D describes the relative motion of the chain end points at all separations.

Thirumalai and co-workers<sup>8</sup> have also offered an estimate for the rate of loop formation. Rather than solve a diffusion equation, these authors use a dimensional argument to estimate the time required for two residues on the polypeptide chain to diffuse together. For two residues separated by n positions on the chain, they estimate a loop formation time

$$\tau(n) \sim L^2(n)/(D_0 P(n))$$
 (8)

Here  $D_0$  is the monomer diffusion constant and P(n) is the equilibrium probability for formation of an n-residue loop within the interior of a long peptide. P(n) decreases as n increases, since large loops are less common than small loops:  $^{8,37}$ 

$$P(n) = P_0 n^{-2.2} / (\exp(1.8(9 - n)) + 1)$$
 (9)

Equation 8 is based on equilibrium statistics and dimensional analysis; it does not arise from an actual model for the dynamics of a diffusing chain. Using eq 8, Guo and Thirumalai estimate that an end-to-end loop will require a time  $\tau \sim 20~\mu s$  to form in a chain of 10 residues.<sup>8,38</sup>

**Upper Limit to the Folding Rate.** The mechanism by which a fully unfolded polypeptide collapses to a compact state is unknown. Even for homopolymers, the dynamics of hydrophobic collapse are not well understood. de Gennes<sup>39</sup> and Grosberg *et al.*<sup>40</sup> originally described homopolymer collapse as a hydrodynamically-controlled crumpling, in which hydrophobic forces cause the chain to pull inwardly on itself. This is clearly not a process that is governed by the rate of diffusive loop formation. Experimental studies of the time scale of collapse have proven very difficult<sup>41</sup> and have not yet confirmed de Gennes' predictions for the collapse rate. However, Ostrovsky and Bar-Yam<sup>42</sup> have proposed a model in which hydrophobic collapse does occur through intrachain diffusion. In this model the two ends of the polymer form aggregates that diffuse together slowly.

If such a picture describes heteropolymer collapse, the rate of intrachain diffusion allows us to set an upper limit on the rate of polypeptide collapse and folding. The theory of Szabo et al.<sup>9</sup> predicts that a loop of n residues forms at a rate proportional to  $n^{-3/2}$ . The smallest loops in proteins have  $\sim 6-10$  residues<sup>43</sup> and should therefore form 10-30 times faster than loops of 50-60 residues. We therefore expect that the smallest loops in an unfolded protein will form in  $\sim 1-3$   $\mu$ s. Thus an unfolded polypeptide collapsing to a compact state should form native-like loops at a rate no faster than  $\sim 10^6$  s<sup>-1</sup>. Since the presence of additional folding steps or thermodynamic barriers would only slow the total folding process, this estimate  $\sim 10^6$  s<sup>-1</sup> provides an approximate upper limit for the rate of folding of a polypeptide from a chemically denatured state.

We have studied intrachain diffusion under strongly denaturing solution conditions. It is difficult to predict whether the intrachain diffusion rate would increase or decrease under folding conditions. Under folding conditions, interactions between amino acid residues presumably introduce roughness to the energy landscape. This may slow diffusion,<sup>44</sup> while at the same time the interactions create an infundibuliform landscape, which favors the collapsed states of the molecule and may therefore accelerate collapse.<sup>6</sup> In order to determine which effect dominates, it would be necessary to model the system and (for example) perform simulations.

#### **Conclusions**

Laser photolysis can provide important information about the earliest events in the collapse and folding of a polypeptide. Jones et al. measured the rate at which met80 (or met65) and his33 (or his26) bind to the heme iron at his18 as the cytochrome c molecule diffuses under unfolding conditions. They found a 10-fold difference in the rates for his33 and met80 binding. Our studies of intrachain ligand binding in dicarboxymethylated cytochrome c appear to confirm Jones et al.'s finding that the methionine residues bind to the heme more quickly than do the histidine residues, despite their greater distance from the heme on the unfolded chain. Further, our studies of ligand binding to the heme peptide (microperoxidase) allow a more complete interpretation of the rates observed by Jones et al. We find that the relatively slow binding of histidine to the heme is primarily due to the chemistry of this ligand—heme interaction, rather than to slow chain dynamics. By contrast, the 40 µs time constant for met80 (or met65) binding to the heme largely represents the time required for the intervening 62-residue length of polypeptide to form a loop bringing the reactants together. Therefore we have established that a large scale diffusive motion of the polypeptide chain in aqueous solvent, under chemically denaturing conditions, requires a time of the order  $35-40 \mu s$ .

From this result we can estimate a diffusion constant D for the relative motion of two positions on an unfolded polypeptide. However, this calculation is model dependent, and a description of polymer diffusional dynamics in terms of a single diffusion constant may oversimplify the problem. Further, intrachain interactions may stabilize intermediate conformations on the diffusive pathway. These can be expected to complicate the relationship between the overall rate  $k_{\rm D+}$  and D. For these reasons, the "diffusion-limited" rate  $1/k_{\rm D+} \sim 3 \times 10^4/{\rm s}$  for bringing together two side chains separated by 62 residues along an unfolded polypeptide is more meaningful than the associated diffusion constant  $D \sim 4 \times 10^{-7}$  cm²/s calculated from the SSS theory.

Finally, we note that rapid protein collapse may have biological significance. In order to prevent the aggregation of unchaperoned proteins after their synthesis *in vivo*, the hydrophobic residues of the protein must be concealed rapidly by collapse of the polypeptide. Once a molecule is compact, with no "sticky" hydrophobic regions on its surface, the native structure may develop more slowly. The rate of aggregation of unfolded polypeptides under cytoplasmic conditions therefore sets a lower limit for the collapse rate that can be tolerated biologically. A successful sequence must not only form a functional folded structure but must also collapse at or above this minimum rate.

**Acknowledgment.** We thank Attila Szabo for many helpful discussions of theoretical issues raised in this work. We also benefited from discussions with Abel Schejter and Devarajan Thirumalai.

# Appendix A: Modeling of the Heme—Peptide/ Methionine/CO Photolysis Spectra

Our modeling of the photolysis spectra confirms our interpretation of the reaction kinetics that follow photolysis. It also yields a spectral determination of the ligand-binding equilibrium constant obtained from the kinetic data. This analysis takes as input the assumption that CO is dissociated from the heme by the photolysis laser pulse, that methionine subsequently binds to the heme, and that CO replaces methionine as the ligand at equilibrium. It gives as output the methionine—heme and CO—heme Soret spectra, which can be compared to the known spectra. It also gives the equilibrium constant for methionine—heme binding.

We represent the various liganded and unliganded complexes of microperoxidase by

$$\mathbf{D}(\lambda) = \text{deoxyheme spectrum}$$

 $\mathbf{M}(\lambda) = \text{methionine} - \text{heme complex}$ 

 $\mathbf{C}(\lambda) = \text{carbonmonoxy-heme complex}$ 

The spectra collected at  $t \le 10 \mu s$ , before the CO begins to rebind, can be written as

$$\mathbf{A}(\lambda,t) = A_0(m(t)\mathbf{M} + d(t)\mathbf{D} + c(t)\mathbf{C}) \tag{A1}$$

where m, c, and d are the concentrations of methionine—heme, carbonmonoxy—heme, and deoxyhemes, respectively, and  $A_0$  is a constant (related to properties of the cuvette, etc.). If the time constant for methionine—heme binding is  $\tau_1$ , the total concentration of hemes is  $N^{\text{Total}}$ , and the photolyzed concentration is  $N^{\text{P}}$ , eq A1 can be rewritten

$$\mathbf{A}(\lambda,t) = A_0 N^{p} (m_0 (1 - \exp(-t/\tau_1)) \mathbf{M} + (1 - m_0 (1 - \exp(-t/\tau_1))) \mathbf{D} + (N^{\text{Total}}/N^{p} - 1) \mathbf{C})$$
(A2)

where  $m_0$  is the equilibrium fraction of methionine—heme at the current methionine concentration,  $m_0 = K_{\text{eq}}[\text{methionine}]/(1 + K_{\text{eq}}[\text{methionine}])$ . Meanwhile the unphotolyzed spectrum is

$$\mathbf{A}(\lambda, 0^{-}) = A_0 N^{\text{Total}} \mathbf{C}$$

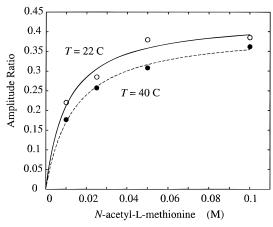
so the observed difference spectrum should be

$$\Delta \mathbf{A}(\lambda, t) = A_0 N^{\text{p}}(m_0 (1 - \exp(-t/\tau_1)) \mathbf{M} + (1 - m_0 (1 - \exp(-t/\tau_1))) \mathbf{D} - \mathbf{C})$$

This can be rewritten as the sum of a time-independent difference spectrum and an exponentially decaying difference spectrum

$$\Delta \mathbf{A} = A_0 N^p (m_0 \mathbf{M} + (1 - m_0) \mathbf{D} - \mathbf{C}) + A_0 N^p \exp(-t/\tau_1) (\mathbf{D} - \mathbf{C})$$
(A3)

As this equation suggests, the experimentally observed difference spectra up until  $t \sim 10~\mu s$  can be described by two components; one is time independent, while the other decays as a single exponential. We therefore identify the time-independent component in the data with the first term on the right-hand side of eq A3, and the decaying component with the second term. We fit the amplitude of the decaying term to a single exponential to obtain the rate  $1/\tau_1$ , and we extrapolate the fit back to t=0 to obtain the initial amplitude of the time-dependent term. Thus we obtain experimentally the two difference spectra



**Figure 14.** Least-squares ratio of  $\Delta A_1 - \Delta A_2$  to  $\Delta A_2$  for heme peptide/methionine/CO photolysis spectra at T = 22 °C (open circles) and T = 40 °C (solid circles). The ratio is plotted as a function of methionine concentration, and the curves represent the fit (see text) that gives the equilibrium binding constant for heme—methionine,  $K_{\rm eq} = 97$  M<sup>-1</sup> (22 °C) and 74 M<sup>-1</sup> (40 °C).

$$\Delta \mathbf{A}_1 = A_0 N^p (m_0 \mathbf{M} + (1 - m_0) \mathbf{D} - \mathbf{C})$$
$$\Delta \mathbf{A}_2 = A_0 N^p (\mathbf{D} - \mathbf{C})$$

with

$$\Delta \mathbf{A}_1 - \Delta \mathbf{A}_2 = A_0 N^p m_0 (\mathbf{M} - \mathbf{D})$$

At this point the deoxy–CO and methionine—deoxy difference spectra are obtained, since they are proportional to  $\Delta A_2$  and  $\Delta A_1 - \Delta A_2$  respectively. The constant  $A_0N^p$  is uninteresting, but the factor  $m_0$  is related to the methionine-binding equilibrium constant  $K_{\rm eq}$  and can be determined by comparing the amplitude of  $\Delta A_1 - \Delta A_2$  observed at different methionine concentrations. For each of the four samples at methionine concentrations from 10 to 100 mM, we plot the (least-squares) ratio of  $\Delta A_1 - \Delta A_2$  to  $\Delta A_2$ , Figure 14. This ratio should be proportional to  $m_0$ , and a fit of the ratio to  $K_{\rm eq}$ [methionine]/(1 +  $K_{\rm eq}$ [methionine]) is a two-parameter fit that gives  $K_{\rm eq} = 97$  M<sup>-1</sup> at 22 °C and  $K_{\rm eq} = 74$  M<sup>-1</sup> at 40 °C. Thus, from modeling the difference spectra—independent of the kinetics—we obtain an equilibrium constant that is close to that obtained from the kinetics.

Knowing  $K_{\rm eq}$ , we can subtract (or add) the absolute deoxyheme spectrum from (or to)  $\Delta {\bf A}_2$  and  $\Delta {\bf A}_1 - \Delta {\bf A}_2$  to obtain the absolute carbonmonoxy—heme and methionine—heme spectra. These are shown in Figure 2, together with the equilibrium spectrum of a CO/heme/methionine sample and the equilibrium spectrum of a methionine/heme sample. Clearly agreement is excellent, indicating that the observed spectral changes are as expected from the kinetic model of eq 1.

### References and Notes

- (1) Jones, C. M.; Henry, E. R.; Hu, Y.; Chan, C. K.; Luck, S. D.; Bhuyan, A.; Roder, H.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11860.
- (2) (a) Ballew, R. M.; Sabelko, J.; Gruebele, M. Proc. Natl. Acad. Sci. U.S.A. 1996, 93, 5759. (b) Huang G. S.; Oas, T. G. Proc. Natl. Acad. Sci. U.S.A. 1995, 92, 6878. (c) Nölting, B.; Golbik, R.; Fersht, A. R. Proc. Natl. Acad. Sci. U.S.A. 1995, 92, 10668. (d) Phillips, C. M.; Mizutani, Y.; Hochstrasser, R. M. Proc. Natl. Acad. Sci. U.S.A. 1995, 92, 7292. (e) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. Biochemistry 1996, 35, 691. (f) Eaton, W. A.; Thompson, P. A.; Chan, Chi-Kin; Hagen, S. J.; Hofrichter, J. Structure 1996, 4, 1133. (g) Wolynes, P. G.; Luthey-Schulten, Z.; Onuchic, J. N. Chem. Biol. 1996, 3, 425.

- (3) (a) Pascher, T.; Chesick, J. P.; Winkler, J. R.; Gray, H. B. *Science* **1996**, *271*, 1558. (b) Mines, G. A.; Pascher, T.; Lee, S. C.; Winkler, J. R.; Gray, H. B. *Chem. Biol.* **1996**, *3*, 491.
- (4) (a) Alexander, P.; Orban, J.; Bryan, P. Biochemistry 1992, 31, 7243. (b) Baldwin, R. L. Proc. Natl. Acad. Sci. U.S.A. 1996, 93, 2627. (c) Khorasanizadeh, S.; Peters, I. D.; Roder, H. Nat. Struct. Biol. 1996, 3, 193. (d) Kragelund, B. B.; Robinson, C. V.; Knudsen, J.; Dobson, C. M.; Poulsen, F. M. Biochemistry 1995, 34, 7217. (e) Kuszewski, J.; Clore, G. M.; Gronenborn, A. M. Protein Sci. 1994, 3, 1945. (f) Ptitsyn, O. B. Adv. Protein Chem. 1995, 47, 83. (g) Schindler, T.; Herrler, M.; Marahiel, M. A.; Schmid, F. X. Nat. Struct. Biol. 1995, 2, 663. (h) Sosnick, T. R.; Mayne, L.; Hiller, R.; Englander, S. W. Nat. Struct. Biol. 1994, I, 149. (i) Waldburger, C. D.; Jonsson, T.; Sauer, R. T. Proc. Natl. Acad. Sci. U.S.A. 1996, 93, 2629.
- (5) Elöve, G. A.; Bhuyan, A. K.; Roder, H. *Biochemistry* **1994**, *33*, 6925
- (6) (a) Bryngelson, J. D.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.*1987, 84, 7524. (b) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Genet.* 1995, 21, 167.
- (7) (a) Karplus, M.; Weaver, D. L. *Nature* **1976**, 260, 404. (b) Karplus, M.; Weaver, D. L. *Biopolymers* **1979**, 18, 1421. (c) Karplus, M.; Weaver, D. L. *Protein Sci.* **1994**, 3, 650. (d) Bashford, D.; Karplus, M.; Weaver, D. L. In *Protein folding. Deciphering the second half of the genetic code*; Gierasch, L., King, J., Eds.; AAAS: Washington, DC, 1990; pp 283–290.
- (8) (a) Camacho, C. J.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1277. (b) Thirumalai, D.; Guo, Z. *Biopolymers* **1995**, *36*, 83.
- (9) Szabo, A.; Schulten, K.; Schulten, Z. J. Chem. Phys. 1980, 72, 4350.
  - (10) Chan, C. K.; Hofrichter, J.; Eaton, W. A. Science 1996, 274, 628.
  - (11) Luck, S. D.; Roder, H. Personal communication.
- (12) (a) Wang, J. C.; Davidson, N. *J. Mol. Biol.* **1966**, *15*, 111. (b) Wang, J. C.; Davidson, N. *J. Mol. Biol.* **1966**, *19*, 469.
- (13) (a) Ehrenberg, A.; Theorell, H. Acta Chem. Scand. 1955, 9, 1193.
   (b) Urry, D. W. J. Am. Chem. Soc. 1967, 89, 4190.
- (14) Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. Proc. Natl. Acad. Sci. U.S.A. 1996, 93, 11615.
  - (15) Gibson, Q. H. Prog. Biophys. Biophys. Chem. 1959, 9, 1.
- (16) (a) Schejter, A.; Aviram, I. J. Biol. Chem. 1970, 245, 1552. (b) Schejter, A. Personal communication.
- (17) Hofrichter, J.; Henry, E. R.; Ansari, A.; Jones, C. M.; Deutsch, R. M.; Sommer, J. H. *Methods Enzymol.* **1994**, 232, 387.
  - (18) Henry, E. R.; Hofrichter, J. Methods Enzymol. 1992, 210, 129.
- (19) Jones, C. M.; Ansari, A.; Henry, E. R.; Christoph, G. W.; Hofrichter, J.; Eaton, W. A. *Biochemistry* **1992**, *31*, 6692.
- (20) The heme—CO recombination rate may differ slightly between samples, owing to the difficulty of preparing two independent samples with identical CO saturation.
- (21) The BuNC complex of the heme peptide, although readily photolyzeable, proved inconvenient to work with. When prepared with excess BuNC, heme peptide samples exhibit multicomponent absorption spectra immediately after photolysis, indicating the presence of an extra, unidentified heme complex in addition to the expected deoxyheme and methionine (or histidine) complexes. This other complex is absent at very low BuNC concentrations. We concluded that BuNC can displace his 18 as the proximal heme ligand, resulting in formation of the dibutyl isocyanide heme complex. To avoid forming this complex, it was necessary to work at very low BuNC concentrations (e.g. 35  $\mu$ M BuNC, 100  $\mu$ M heme peptide), which reduced the photolysis signal and gave a poorer signal-to-noise ratio. For this reason, we eventually used CO instead of BuNC as the photolyzeable ligand in our experiment.
- (22) In the pseudo-first-order limit, the kinetic equations associated with eq 1 have eigenvalues (Moore, J. W.; Pearson, R. G. *Kinetics and Mechanism*, 3rd ed.; John Wiley: New York, 1981)

$$\lambda_1 = -\frac{1}{2} [k' + (k'^2 - 4k_1 C_0 k_3)^{1/2}]$$
  
$$\lambda_2 = -\frac{1}{2} [k' - (k'^2 - 4k_1 C_0 k_3)^{1/2}]$$

where  $k' = (k_1C_0 + k_2L_0 + k_3)$ . If  $k_2L_0 + k_3 \gg k_1C_0$ , then  $\tau_1 \equiv -1/\lambda_1$  and  $\tau_2 \equiv -1/\lambda_2$  simplify to the expressions given in eq 2.

- (23) A weak third process, occuring on a time scale of  $\sim 1 \,\mu s$ , is apparent in the higher SVD components  $VS_2$  and  $VS_3$  of the heme—methionine data, Figure 8. This process most likely indicates partial aggregation of the unliganded peptide molecules prior to rebinding of the CO.
  - (24) Shoup, D.; Szabo, A. Biophys. J. 1982, 40, 33.
- (25) Miers, J. B.; Postlewaite, J. C.; Zyung, T.; Chen, S.; Roemig, G. R.; Wen, X.; Dlott, D. D.; Szabo, A. J. Chem. Phys. **1990**, *93*, 8771.
- (26) Flory, P. J. Statistical mechanics of chain molecules; John Wiley: New York, 1969.
- (27) Damaschun, G.; Gast, K.; Müller-Frohne, M.; Zirwer, D. Biopolymers, in press.
- (28) Our analysis of intrachain diffusion relies on the assumption that the met80—his18 segment of unfolded cytochrome c obeys the statistics of a Gaussian chain under  $\theta$  solvent conditions. We note, however, that Gottfried  $et\ al.^{35}$  found that the mean-square separation  $L^2$  of residues in

the unfolded BPTI molecule (in 6 M GuHCl) does not scale linearly with the number n of intervening residues. In fact,  $L^2$  in that study shows little or no dependence on chain length, contrary to the linear dependence that is predicted by random coil statistics. This result is puzzling, since Damaschun et al.<sup>27</sup> found that the mean-squared radius of gyration  $R_G^2 \propto n$  for many unfolded proteins, as predicted for a Gaussian chain in a  $\theta$  solvent. In addition, Tanford (Tanford, C. Adv. Protein Chem. 1968, 23, 121) presented a summary of  $R_{G}^{2}$  measurements for a variety of proteins unfolded in 6 M GuHCl; these results, based on viscosity data, were generally consistent with random coil statistics. Gottfried et al. suggest that the disagreement may reflect long-range intrachain interactions, possibly unique to the unfolded BPTI molecule, that persist even under strongly unfolding

- (29) Because the peptide is a large molecule, the relevant diffusion constant is that of the free amino acid. In our solvent, this may be roughly estimated from the value  $D_0^{\rm ala} \approx 0.9 \times 10^{-5} \, {\rm cm^2/s}$  measured for alanine in aqueous solution at T = 25 °C (Weast, R., Ed. *Handbook of Chemistry &* Physics, 55th ed.; CRC Press: Cleveland, OH, 1974). From the Einstein-Stokes relation we expect  $D_0^{\text{hist}} \approx D_0^{\text{ala}}(r_{\text{ala}}/r_{\text{hist}})\eta_{\text{H}_2\text{O}}(25 \,^{\circ}\text{C})/\eta_{\text{GuHCl}}(40 \,^{\circ}\text{C})$ , where  $r_{\rm hist} \sim r_{\rm meth} \sim 1.3 r_{\rm ala}$  describes the approximate size of the diffusing molecule and  $\eta$  represents the solvent viscosity. With  $\eta_{\rm H_2O}(40~{}^{\circ}{\rm C}) \approx 0.65$ cP,  $\eta_{\rm H_2O}(25~^{\circ}{\rm C}) \approx 0.9$  cP, (Weast), and  $\eta_{\rm GuHCl}(25~^{\circ}{\rm C})/\eta_{\rm H_2O}(25~^{\circ}{\rm C}) \approx 1.5$ (Fasman, G. D. CRC Handbook of Biochemistry and Molecular Biology, 3rd ed.; CRC Press: Boca Raton, FL, 1976; Vol. 1), we assume that  $\eta_{\text{GuHCl}}$  $(40 \,{}^{\circ}\text{C}) \approx \eta_{\text{GuHCl}}(25 \,{}^{\circ}\text{C})(\eta_{\text{H}_2\text{O}}(40 \,{}^{\circ}\text{C})/\eta_{\text{H}_2\text{O}}(25 \,{}^{\circ}\text{C})) \approx 1 \,\text{cP. Then } D_0^{\text{med}}$  $D_0^{\text{hist}} \approx 0.6 \times 10^{-5} \text{ cm}^2/\text{s}.$ 
  - (30) Schejter, A.; Plotkin, B.; Vig, I. FEBS Lett. 1991, 280, 199.
  - (31) Greenwood, C.; Gibson, Q. J. Biol. Chem. 1989, 264, 19022.
- (32) This upper limit to the heme-ligand photolysis rate was estimated in previous photodissociation studies that used the same instrument.
- (33) Schaad, O.; Zhou, H.-X.; Szabo, A.; Eaton, W. A.; Henry, E. R. Proc. Natl. Acad. Sci. U.S.A. 1993, 90, 9547.
- (34) (a) Haas, E.; Katchalski-Katzir, E.; Steinberg, I. Z. Biopolymers 1978, 17, 11. (b) Beechem, J. M.; Haas, E. Biophys. J. 1989, 55, 245.
  - (35) Gottfried, D. S.; Haas, E. Biochemistry 1992, 31, 12353.
- (36) Buckler, D. R.; Haas, E.; Scheraga, H. A. Biochemistry 1995, 34, 15965.

- (37) des Cloizeaux, J. J. Phys. 1980, 41, 223.
- (38) Equation 8 predicts only the relative speed of formation of loops of different sizes

$$\tau(n_1)/\tau(n_2) \sim (n_1/n_2)^{2.2+2\nu} \approx (n_1/n_2)^{3.2}$$

 $\tau(n_1)/\tau(n_2) \sim (n_1/n_2)^{2.2+2\nu} \approx (n_1/n_2)^{3.2}$  since the absolute normalization for P(n) (i.e.  $P_0$  in eq 9) was not calculated.<sup>8</sup> From the measured rates of intrachain and bimolecular histidine and methionine binding, we can estimate  $P_0$  and test eq 8. For the binding of his33 to the heme at his18, we assume that the rate of intrachain diffusion lies within the limits  $0 \le 1/k_{D+}^{uni} \le 100 \,\mu s$ , which through eq 4 implies

$$0 \le 1/k_{\text{on}}^{\text{uni}} - K^{\text{bi}}/(K^{\text{uni}}k_{\text{on}}^{\text{bi}}) \le 100 \,\mu\text{s}$$

Given this range, the measured rates for histidine binding at 40 °C, and the estimate  $K^{\text{bi}} = (2/3)\pi a^3 \approx 4.3 \times 10^{-3} \text{ M}^{-1}$ , we conclude that the unimolecular equilibrium constant  $K^{\text{uni}}$  lies within the range  $\sim (2.5-3.3) \times$  $10^{-5}$  for his 33—heme binding. If  $P(15) = K^{\text{uni}}$ , we can solve eq 9 to obtain the normalization  $P_0$  for a 104-residue polypeptide (i.e. for cytochrome c). This gives  $P_0 \approx (9.7-12.8) \times 10^{-3}$ . Then the equilibrium probability for formation of a loop between met80 and the heme of cytochrome c is P(62) $\approx (1.1-1.5) \times 10^{-6}$ . With  $L \approx 85$  Å and  $D_0 \sim 10^{-5}$  cm<sup>2</sup>/s, eq 8 then gives

$$\tau(\text{met}80-\text{his}18) \sim L^2/(D_0 P(62)) \approx 50-66 \text{ ms}$$

This prediction compares poorly with our experimental result:  $\tau \approx 35-40$  $\mu$ s. The uncertainty in  $k_{on}^{bi}$  for his 33 has little impact on the sizeable disagreement, while a smaller (and more realistic) value for  $D_0$  would worsen it considerably. We do not understand the origin of the discrepancy in our calculation of the loop formation time from eqs 8 and 9 and that given in

- (39) de Gennes, P. G. J. Phys. Lett. 1985, 46, L639.
- (40) Grosberg, A. Yu.; Nechaev, S. K.; Shakhnovich, E. I. J. Phys. 1988, 49, 2095.
- (41) (a) Yu, J.; Wang, Z.; Chu, B. Macromolecules 1992, 25, 1618. (b) Chu, B.; Ying, Q.; Grosberg, A. Y. Macromolecules 1995, 28, 180. (c) Chu, B.; Ying, Q. Macromolecules 1996, 29, 1824.
- (42) (a) Ostrovsky, B.; Bar-Yam, Y. Europhys. Lett. 1994, 25, 409. (b) Ostrovsky, B.; Bar-Yam, Y. Biophys. J. 1995, 68, 1694.
  - (43) Leszczynski, J. F.; Rose, G. D. Science 1986, 234, 849.
  - (44) Zwanzig, R. Proc. Natl. Acad. Sci. U.S.A. 1988, 85, 2029.