

## Ligand Exchange during Unfolding of Cytochrome *c*\*

(Received for publication, February 3, 1999, and in revised form, April 9, 1999)

Syun-Ru Yeh‡ and Denis L. Rousseau

From the Department of Physiology and Biophysics, Albert Einstein College of Medicine, Bronx, New York 10461

**The productive folding pathway of cytochrome *c* passes through an obligatory HW intermediate in which the heme is coordinated by a solvent water molecule and a native ligand, His-18, prior to the formation of the folded HM state with both the native His-18 and Met-80 heme coordination. Two off pathway intermediates, a five-coordinated state (5C) and a bis-histidine state (HH), were also identified during the folding reaction. In the present work, the thermodynamics and the kinetics of the unfolding reaction of cytochrome *c* were investigated with resonance Raman scattering, tryptophan fluorescence spectroscopy, and circular dichroism. The objective of these experiments was to determine if the protein opens up and diverges into the differing heme ligation states through a many pathway mechanism or if it passes through intermediate states analogous to those observed during the folding reaction. Equilibrium unfolding results indicate that, in contrast to 5C, the stability of HH with respect to HW decreases as the concentration of GdnHCl increases. The difference in their response to the denaturant indicates that the polypeptide structure of 5C is relatively loose as compared with HH in which the polypeptide is misfolded. Time-resolved resonance Raman measurements show that strikingly similar ligand exchange reactions occur during unfolding as were observed during folding. Combined with fluorescence data, a kinetic model is proposed in which local structural rearrangements controlled by heme ligand exchange reactions appear prior to the global relaxation of the polypeptide chain.**

To perform its biological function, a nascent polypeptide produced at the ribosome needs to fold to its native structure accurately and efficiently. The “folding codon” encoded in its amino acid sequence and the mechanism by which the polypeptide folds to its native structure remain central problems in biology (1–7). Recent advances have shed light on the nucleation condensation model in which the growth of the polypeptide architecture of a small single domain protein starts in a particular region of the molecule, the folding nucleus, and subsequently propagates throughout the rest of the molecule until the native conformation is reached (3, 8, 9). The reversible folding often appears as two-state cooperative processes in which only the totally unstructured unfolded state (U)<sup>1</sup> and the

well ordered folded state (F) are populated. The folding transition state (‡) relies on the formation of the folding nuclei (Equation 1) and may be determined by site-directed mutagenesis along with kinetic analysis (8).



The folding of larger proteins, on the other hand, is much more complex. The polypeptide commonly passes through various intermediates (*I<sub>i</sub>*) prior to reaching the unique native structure (Equation 2).



In addition, the productive folding pathway sometimes needs to compete with nonproductive off pathway reactions.



The off pathway intermediates (*D<sub>i</sub>*) are dead-end products that may form deep kinetic traps during folding and have been reported to be linked to human pathogenesis (10, 11). To unify the kinetic studies from a large variety of proteins, a “new view” of protein folding has been postulated in which unique folding channels are replaced by a statistical distribution of many pathways (4, 12–14). Although there is little direct experimental evidence for the statistical nature of the folding reactions, a diversity of unfolding pathways has been postulated for chymotrypsin inhibitor 2 by molecular dynamics simulations (13).

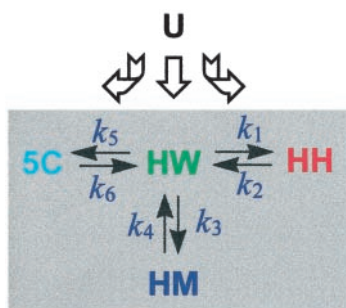
In contrast to the folding reactions, the unfolding transition in single domain proteins usually appears to be a cooperative two-state process regardless of the size of the protein. It is generally considered as a “melting” of the protein crystal, because the closed packed interior of a protein resembles the ordered structure of molecular crystals. In this report, we use cytochrome *c* (cyt *c*) as a model to unravel the statistical nature of the unfolding reaction. Cyt *c* is a convenient model for folding and unfolding studies, because the native structure is well known and the folding kinetics can be followed by various spectroscopic techniques (15–20). In addition, the protein contains a heme prosthetic group that can be used as a probe for the local structure of the protein during the folding reaction (21). We have demonstrated in previous reports that the folding reaction of cyt *c* follows a two-step mechanism (15, 21–23). In the nascent phase of folding, cyt *c* collapses from a random coil-like structure to a semi-compact form that involves little or none of the specific tertiary interactions present in the native structure. This collapsing process is kinetically controlled and thus happens very fast, in less than 100 μs (21, 24, 25). The size of the conformational space is dramatically reduced during this phase. It facilitates the searching of the native structure in the subsequent phase, which is a thermodynamically controlled heme ligand exchange phase that is regulated by random struc-

\* This work was supported by the National Institutes of Health Research Grants GM-54806 and GM-54812. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 718-430-4234; Fax: 718-430-4230; E-mail: SYEH@AECOM.YU.EDU.

<sup>1</sup> The abbreviations used are: U, unfolded protein; F, folded protein; cyt *c*, cytochrome *c*; HW, cytochrome *c* with heme coordination by His-18 and a water molecule; HM, cytochrome *c* with heme coordination by His-18 and Met-80; 5C, five-coordinated cytochrome *c* with heme coordination by only a water molecule; HH, cytochrome *c* with heme coordi-

ination by bis-histidine; GdnHCl, guanidine hydrochloride; CD, circular dichroism.



Scheme 1

tural fluctuations biased to the native polypeptide fold. It is at this stage when the specific native contacts are established (15).

The productive folding pathway of *cyt c*, as described in Scheme 1, starts from the totally unstructured polypeptide state, U, and passes through an intermediate, HW, in which His-18 and a water molecule are coordinated to the heme, and then culminates in the native His-18–Met-80 heme-coordinated form (HM). Two off pathway intermediates with different heme coordination have been identified during the folding reaction, a five-coordinated species (5C) and a bis-histidine species (HH) in which the polypeptide chain is misfolded (23). In contrast to the folding reaction that converges to the native HM state, the unfolding reaction of *cyt c* starts from the native protein and diverges to form the conformationally heterogeneous *unfolded* polypeptide structures with differing heme ligation states as formulated in Equation 4.



Recent studies of *cyt c* unfolding by circular dichroism (CD) and tryptophan fluorescence demonstrated nonexponential kinetics indicating the existence of unfolding intermediates (26–28). In addition, hydrogen exchange measurements have suggested that the partially unfolded forms of *cyt c* observed under equilibrium conditions represent kinetic unfolding intermediates (29). To obtain a coherent picture, the equilibrium unfolding reaction of *cyt c* was studied as a function of the GdnHCl concentration at pH 3.0 and 5.9 with CD, tryptophan fluorescence, and resonance Raman scattering. The kinetics of the unfolding reaction at pH 3.8 and 5.8 were monitored by resonance Raman scattering and tryptophan fluorescence. The overall objective of these experiments was to determine if the protein simply opens up and diverges into the differing heme ligation states through a many pathway mechanism or if it passes through intermediate states analogous to those observed during the folding reaction. A direct comparison with the folding reaction under similar ambient conditions is discussed to enlighten the chemical and physical processes that underlie these reactions.

#### MATERIALS AND METHODS

Horse heart ferricytochrome *c* (type IV from Sigma) was treated with ferricyanide to oxidize the residual ferrous form and purified on a Sephadex G-25M column. For the equilibrium studies, the protein was diluted into pH 3.0 or 5.9 acetate buffer (0.1 M) at the desired concentration of guanidine hydrochloride (GdnHCl, ultra pure grade from Pierce) a few hours prior to the experiments. The final protein concentration was 20  $\mu\text{M}$ . The fluorescence measurements were recorded with a SLM Aminco fluorometer (Urbana, IL). The excitation wavelength was 295 nm, and both slits were set to 4 nm. The CD measurements were performed on a JASCO model J720 spectrometer (Easton, MD).

For the kinetic studies, the unfolding of *cyt c* was initiated by a 6-fold dilution of the protein with 4.8 M GdnHCl in 0.1 M acetate buffer solution in a rapid solution mixer with a mixing dead time of  $\sim 100 \mu\text{s}$

(21). The final protein concentration was 10  $\mu\text{M}$ . The two parent solutions were pressurized by a syringe pump and introduced into the T-shaped mixing nozzle, which is in direct contact with a quartz flow cell (250  $\times$  250  $\mu\text{m}$ ). The details of the characterization of this mixer were described previously (21). The time-resolved resonance Raman and fluorescence spectra were obtained by moving the quartz cell relative to the laser focusing point along the flow direction. The position of the flow cell was controlled by a micrometer on a translational stage, such that an accurate time domain could be calculated from the solution flow rate.

The Raman measurements were made with previously described instrumentation (21). The output at 407 nm from a krypton ion laser (Spectra Physics) was focused to an  $\sim 30 \mu\text{m}$  spot (power  $\sim 100$  milliwatts) on the continuously flowing sample in the flow cell. The scattered light was collected and focused on the entrance slit of a 1.25-m polychromator (Spex) where it was dispersed and then detected by a charge-coupled device camera (Princeton Instruments). The Raman spectrum at each time point was deconvoluted into contributions from HH, HW, 5C, and HM based on the procedure published elsewhere (21).

For the tryptophan fluorescence measurements, the output at 264 nm from a frequency-doubled argon ion laser (Coherent) was focused to an  $\sim 30 \mu\text{m}$  spot on the continuously flowing sample in the observation cell. The fluorescence was collected and focused on the entrance slit of a 0.27-m polychromator (Spex) where it was dispersed and then detected by a charge-coupled device camera (Photometrics).

#### RESULTS AND DISCUSSION

**Equilibrium Unfolding Studies**—GdnHCl-induced unfolding of *cyt c* was monitored by far UV CD and tryptophan fluorescence spectroscopy as a function of denaturant concentration. Far UV CD monitors the  $\alpha$ -helix content of the protein (30). The tryptophan fluorescence (31, 32), on the other hand, probes the average distance between the fluorescent chromophore, Trp-59, and the intrinsic quencher, the heme prosthetic group, which is covalently linked to Cys-14 and Cys-17 through thioether linkages. At pH 5.9, the equilibrium titration curve from CD measurements coincides well with that from the fluorescence measurements with a midpoint of approximately 2.5 M as shown in Fig. 1A.

To gain more structural information, the equilibrium unfolding curves were investigated with resonance Raman scattering. The native HM form converts to HH and HW as the concentration of GdnHCl increases with a preference of HH over HW (Fig. 1B). Based on the populations of these three species, a representation of their free energy changes relative to HW as a function of the GdnHCl concentration was constructed and is shown in Fig. 1C. In the absence of GdnHCl, the relative stabilities of the intermediates are  $\text{HM} > \text{HH} > \text{HW}$  (22). In the presence of 2.4 M GdnHCl, the population of HM is equal to that of HH (Fig. 1B), suggesting that the free energies of HM and HH are identical. Furthermore, the energy level of HW lies above that of HM and HH, because the population of HW is less than that of HM and HH (Fig. 1B). Likewise, the free energies of HM and HW are identical at 2.7 M GdnHCl and that of HH lies below those of HM and HW. It is concluded from Fig. 1C that both the native state, HM, and the off pathway intermediate, HH, are destabilized by GdnHCl with respect to HW. This result is consistent with the observation that the presence of GdnHCl in the refolding solution destabilizes HH and assists the conversion of HH back to the obligatory intermediate HW during the folding reaction (23).

It is intriguing that the transition curve from the resonance Raman scattering measurements occurs slightly earlier (lower GdnHCl concentration) as compared with those from the CD and fluorescence measurements as shown in Fig. 1A. This is most evident in the pre-transition region, where at 2 M GdnHCl the population of HM determined from the resonance Raman measurements has dropped by nearly 20% in contrast to no reduction of the native  $\alpha$ -helical content and 8% reduction of the molecular size based on the CD and fluorescence measure-

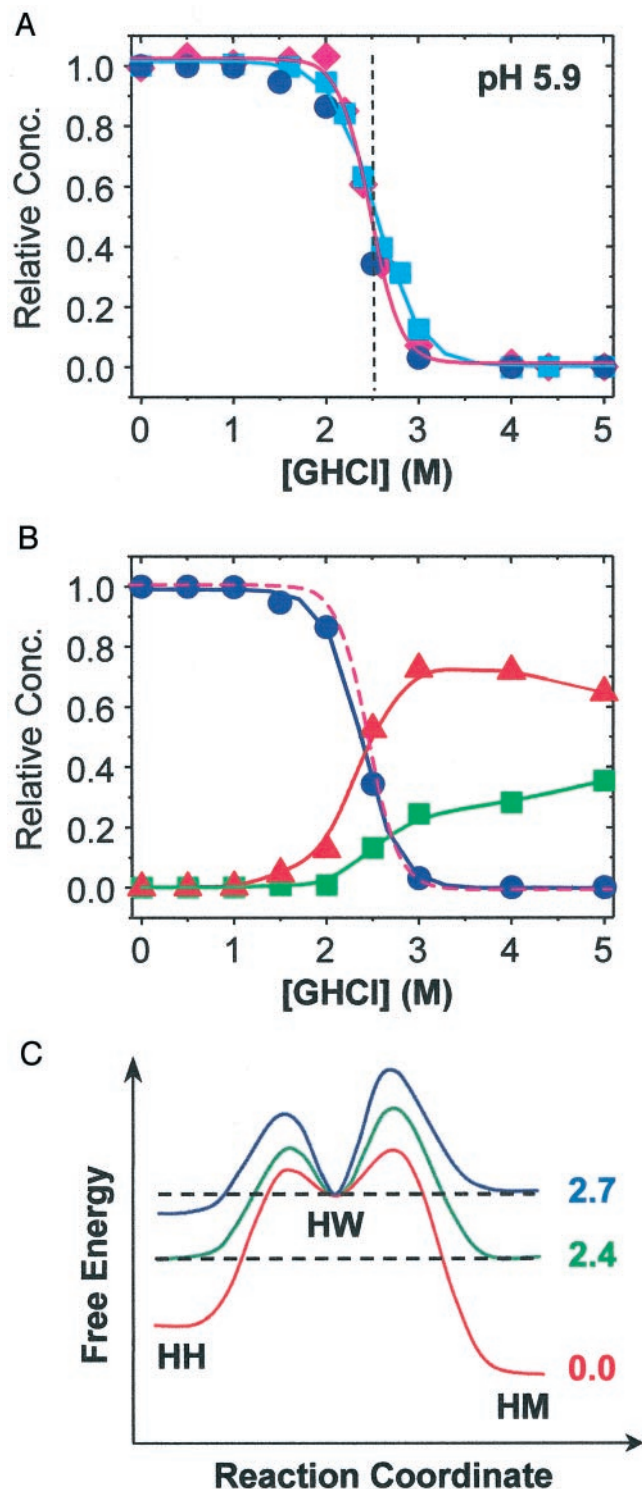


FIG. 1. *A* and *B*, equilibrium titration curves for the unfolding of cytochrome *c* by GdnHCl (GdnHCl) at pH 5.9. In *A*, the populations represented by the cyan squares were determined from the intensity of the tryptophan fluorescence at 344 nm, those represented by the magenta diamonds were determined from CD measurements at 222 nm, and those represented by the dark blue circles were determined from resonance Raman measurements by following the formation of HM. In *B*, the dark blue circles, the green squares, and the red triangles are the populations of the HM, HW, and HH forms, respectively, determined from resonance Raman measurements. The magenta line in *B* is the CD titration curve reproduced from *A* to guide the eye. *C*, pictorial illustration of the relative free energy changes of HM and HH with respect to HW as a function of the GdnHCl concentration. The red, green, and dark blue curves correspond to 0.0, 2.4, and 2.7 M GdnHCl concentrations, respectively.

ments, respectively. The difference between these physical measurements suggests that under low GdnHCl concentrations, the polypeptide backbone structure changes, associated with the changes in heme ligation states, are localized and the thermodynamic intermediates are compact with most of the native secondary structure preserved. Similar conclusions were drawn from equilibrium hydrogen exchange experiments (29). The merger of the HM curve with those from CD and fluorescence measurements in the post-transition region implies that the residual secondary and tertiary structure in HH and HW is eventually lost in the presence of high concentrations of denaturant.

At pH 3.0, the denaturation transition curves measured by CD and tryptophan fluorescence appear much broader than that of pH 5.9 with the transition midpoint shifted from 2.5 to 2.0 M (Fig. 2A). It is plausible that the positive charges distributed over the polypeptide chain resulting from the protonation of certain amino acid residues introduce electrostatic repulsion between polypeptide segments. They destabilize the molecule and loosen the structure. In contrast to the cooperative nature of the sharp transition observed at pH 5.9, the perturbation induced by GdnHCl in one part of the molecule is not readily transmitted to other parts of the molecule. This results in a broader transition with a midpoint shifted to a lower GdnHCl concentration as shown in Fig. 2A.

The change of the ligation states in response to the GdnHCl concentration is much more complex under acidic conditions than that at the higher pH as shown in Fig. 2B. The population of HM from the resonance Raman measurements deviates substantially from the CD and fluorescence titration curves in the pre-transition region. Only one unfolding intermediate, HW, was detected at a GdnHCl concentration of 1.5 M or less. The conversion from HM to HW is extremely efficient; approximately 20% of HM is converted to HW in the presence of only 0.1 M GdnHCl. The sensitivity of the iron-methionine bond to a small amount of GdnHCl cannot be attributed to a direct weakening of this bond under acid conditions. We postulate that the perturbation in the  $\omega$ -loop structure, which holds Met-80 in a favorable position for ligation, introduced by protonation destabilizes the iron-methionine bond.

At concentrations above 1.5 M GdnHCl, where the population of HW reaches a maximum, another folding intermediate, 5C, becomes visible at the expense of HW. The population of 5C levels off at 2.5 M GdnHCl. As the GdnHCl concentration increases from 2.5 to 4.0 M, the ratio of HW/5C stays constant, whereas the amount of secondary structure and the compactness of the protein continue to decline (Fig. 2A). This suggests that the structures of both HW and 5C are loosened steadily by the addition of denaturant. The polypeptide structures associated with HW and 5C are not totally unfolded until the concentration of GdnHCl is greater than 4.0 M at which point more conversion from HW to 5C is observed.

Based on the populations of HM, HW and 5C shown in Fig. 2B, a pictorial illustration of their free energy changes relative to HW as a function of the GdnHCl concentration, was constructed (Fig. 2C). In the absence of GdnHCl, HM is more stable than HW, and HW is more stable than 5C (23). In the presence of 1.2 M GdnHCl, the population of HM is equal to that of HW (Fig. 2B), suggesting that the free energies of HM and HW are identical. In addition, the energy level of 5C lies above that of HM and HW, because the population of 5C is less than that of HM and HW. Analogously, the free energies of HM and 5C are identical in the presence of 2.0 M GdnHCl and that of HW lies below those of HM and 5C. The free energies of HW and 5C are identical in the presence of 2.3 M GdnHCl, and that of HM lies above those of HW and 5C. It can be concluded from

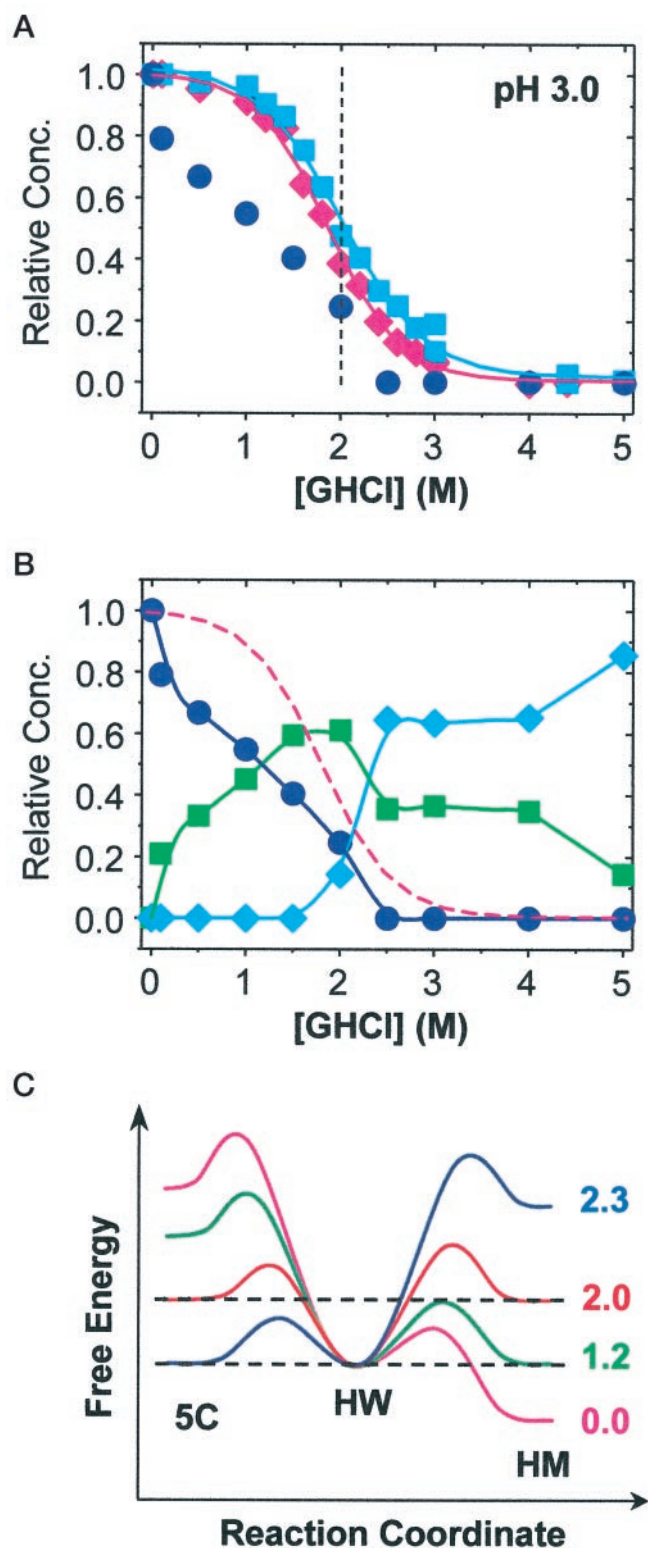


FIG. 2. *A* and *B*, equilibrium titration curves for the unfolding of cyt *c* by GdnHCl (GHCl) at pH 3.0. In *A*, the populations represented by the cyan squares were determined from the intensity of the tryptophan fluorescence at 344 nm, those represented by the magenta diamonds were determined from CD measurements at 222 nm, and those represented by the dark blue circles were determined from resonance Raman measurements by following the formation of HM. In *B*, the dark blue circles, the green squares, and the cyan diamonds are the populations of the HM, HW, and 5C forms, respectively, determined from resonance Raman measurements. The magenta line in *B* is the CD titration curve reproduced from *A* to guide the eye. *C*, pictorial illustration of the relative free energy changes of HM and 5C with respect to HW as a

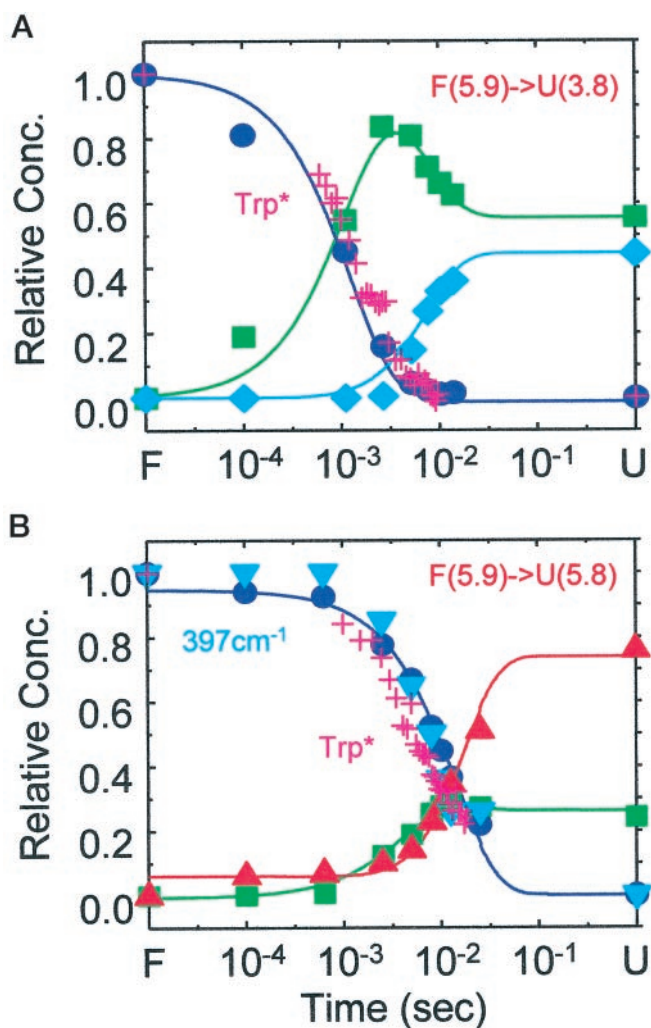
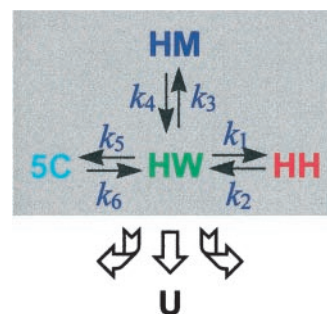


FIG. 3. Time course for the populations of the various heme-ligand coordinated states during the unfolding of cyt *c* determined from tryptophan fluorescence and resonance Raman scattering measurements. The folded protein at pH 5.9 was unfolded at pH 3.8 (*A*) and at pH 5.8 (*B*) in 4.0 M GdnHCl. The magenta crosses were determined from the intensity of the tryptophan fluorescence. The green squares, the red triangles, the dark blue circles, and the cyan diamonds are the populations of the HW, HH, HM, and 5C forms, respectively, determined from the resonance Raman measurements. The cyan wedges in *B* represent the intensity of the 397-cm<sup>-1</sup> Raman line that is sensitive to the structure of the heme pocket. The solid lines are fits to the data using the model presented in Scheme 2. The rate constants are listed in Table I.



Scheme 2

function of the GdnHCl concentration. The magenta, green, red, and dark blue curves correspond to 0.0, 1.2, 2.0, and 2.3 M GdnHCl concentrations, respectively.

TABLE I  
Cytochrome *c* folding and unfolding rate constants

The rate constants were determined by fitting the time-resolved data for folding and unfolding reactions of cytochrome *c* under various pH conditions as indicated in the parentheses.

	$k_1$ (s <sup>-1</sup> )	$k_2$ (s <sup>-1</sup> )	$k_3$ (s <sup>-1</sup> )	$k_4$ (s <sup>-1</sup> )	$k_5$ (s <sup>-1</sup> )	$k_6$ (s <sup>-1</sup> )
F(5.9) → U(5.8)	170	60	0	70	0	0
F(5.9) → U(3.6)	0	0	0	800	65	81
U(5.8) → F(5.0)	700	35	350	0	0	0
U(3.6) → F(5.0)	160	50	300	8	0	1500

Fig. 2C that in contrast to the native HM state, 5C is stabilized by GdnHCl with respect to HW.

Although both HH and 5C are off pathway intermediates, their response to denaturant is dramatically different as demonstrated in Figs. 1C and 2C. This difference can be attributed to the nature of the peptide backbone conformations associated with the two species. HH is a misligated state with the native His-18 and non-native His-33 as the two axial ligands. His-26 is also a potential heme ligand, but it was shown recently that His-33 is the primary ligand in the misligated state (33). The coordination of His-33 to the heme stitches the wrong parts of the peptide fragments together and traps the protein in a misfolded structure. It has been demonstrated that small amounts of GdnHCl accelerate the conversion of HH to the productive folding intermediate, HW (23). This is consistent with the destabilization effect of GdnHCl on HH reported here. On the other hand, the peptide backbone structure of 5C is rather floppy as compared with that of HH, because the intramolecular cross-linking between both His-18 and Met-80 through the heme is broken. The 5C intermediate is more stable in the presence of higher GdnHCl concentrations.

**Unfolding Kinetics**—Fig. 3 shows the unfolding kinetics of cyt *c* at pH 3.8 and 5.8 monitored by resonance Raman scattering. The time-dependent populations of the various folding intermediates in these two reactions are strikingly similar to those observed in the folding reactions. At pH 3.8, HM steadily converts to HW with a half-life of just under a millisecond (Fig. 3A). HW reaches a maximum (~80%) at 5 ms and decays to ~60% at 40 ms. 5C appears at about 1 ms and reaches steady state at approximately 40 ms. The transient population of HW suggests that HW may be the precursor of 5C. At pH 5.8, the unfolding reaction is much slower than that at pH 3.8. HM gradually converts to HH and HW (Fig. 3B). The reaction reaches steady state at about 80 ms.

To gain more insight into the unfolding mechanism, several kinetic models were tested to account for the time-dependent data shown in Fig. 3. The model, which best fits the data, is illustrated in Scheme 2, which surprisingly follows the reverse sequence of the folding reaction (Scheme 1). In this scenario, the unfolding reaction first produces the HW heme coordination state by replacing one of the two native ligands, Met-80, with a water molecule. It is then followed by either dissociation of His-18 to generate 5C (pH 3.8) or the coordination of a “wrong” histidine ligand, His-33, in the replacement of the water molecule to produce HH (pH 5.8). This model confirms the prediction that HW is the common precursor for 5C mentioned earlier. The *solid lines* in Fig. 3 are fits to the data with the kinetic model shown in Scheme 2. The resulting rate constants are listed in Table I. The agreement between the calculations and the data confirms the validity of the model.

The relaxation of the tertiary structure surrounding the heme during the unfolding of cyt *c* at pH 5.8 (Fig. 3B) was monitored by the decrease in the 397-cm<sup>-1</sup> line in the low

frequency Raman spectrum. It has been shown that the heme prosthetic group in cyt *c* is highly distorted because of the local environment of the native polypeptide fold (34, 35). This tension is totally released as the polypeptide chain unfolds. The Raman line at 397 cm<sup>-1</sup>, assigned as a bending mode of the sulfur atom of Cys-14/17, the  $\beta$ -carbon atom of the heme, and the atom bridging these two (36), was established as a marker line for the formation of the native tertiary structure of the heme pocket, because the heme is covalently linked to the polypeptide backbone through these two cysteine residues (22). Interestingly, the time-dependent changes of the 397-cm<sup>-1</sup> line coincide well with the decay of the native HM structure. This result suggests that the dissociation of Met-80 from the heme is concerted with the disruption of the heme pocket, the same type of behavior as has been observed during folding (22).

The global structure changes associated with the ligand exchange reactions were investigated further with tryptophan fluorescence spectroscopy. In the native state, the fluorescence from Trp-59, the only tryptophan residue in cyt *c*, is totally quenched by the spatially nearby heme prosthetic group. The majority of the fluorescence intensity from Trp-59 can be recovered as cyt *c* unfolds because of the increase in the distance between Trp-59 and the heme that is anchored at residue positions 14 and 17. Tryptophan fluorescence therefore provides a qualitative estimate of the global size of the protein. The recovery of the fluorescence intensity during the unfolding reactions of cyt *c* at pH 3.8 and 5.8 agrees well with the disappearance of HM as shown in Fig. 3.

Several conclusions can be drawn from these experiments. First, the fluorescence data shown in Fig. 3 suggest that the global molecular size of cyt *c* is not completely relaxed until the ligand exchange reaction reaches steady state. In the presence of GdnHCl, at concentrations of 4 M or more, the polypeptide chain of cyt *c* is totally unfolded and is free from conformational constraints. The potential heme ligands (His-18 or His-33), covalently linked to the polypeptide chain, can freely exchange with one another or a solvent water molecule with rates >10,000 s<sup>-1</sup> evident from the pH jump experiments (data not shown). These rates are much faster than those observed for the unfolding reactions (Table I) suggesting that the polypeptide backbone is in a semi-compact form during the ligand exchange phase, such that the reactions are constrained by residual tertiary structure just as that observed during the folding of cyt *c*. Second, the Met-80 off rate ( $k_4$ ) is much faster at low pH than that at high pH (800 *versus* 70 s<sup>-1</sup>). This is because of destabilization of the native polypeptide structure at low pH by the electrostatic repulsion resulting from the protonation of certain amino acid residues as discussed earlier. Third, the His-33 on rate ( $k_1$ ) is 0 at pH 3.8 but 170 s<sup>-1</sup> at pH 5.8. This contrast is because of the lack of unprotonated His-33 available for heme coordination at low pH as demonstrated in the folding reaction (22). On the other hand, the formation rate of 5C ( $k_5$ ) is 0 at pH 5.8 because of the instability of 5C at this pH. Lastly, analogous to the folding reaction, as may be seen in Scheme 2, HW is an obligatory unfolding intermediate leading to the formation of the HH and 5C forms.

**Folding Kinetics**—For comparison, Fig. 4 shows the folding reactions of cyt *c* under similar conditions as those applied in Fig. 3. The *solid lines* are the fitted data with the model illustrated in Scheme 1 with the rate constants listed in Table I. The His-18 on rate ( $k_6$ ) and the His-33 on rate ( $k_1$ ) are faster than those observed under unfolding conditions because of the difference in the concentration of denaturant in the solution. It is important to point out that although the folding conditions are the same for both reactions shown in Fig. 4, the overall folding rate slows down by more than an order of magnitude

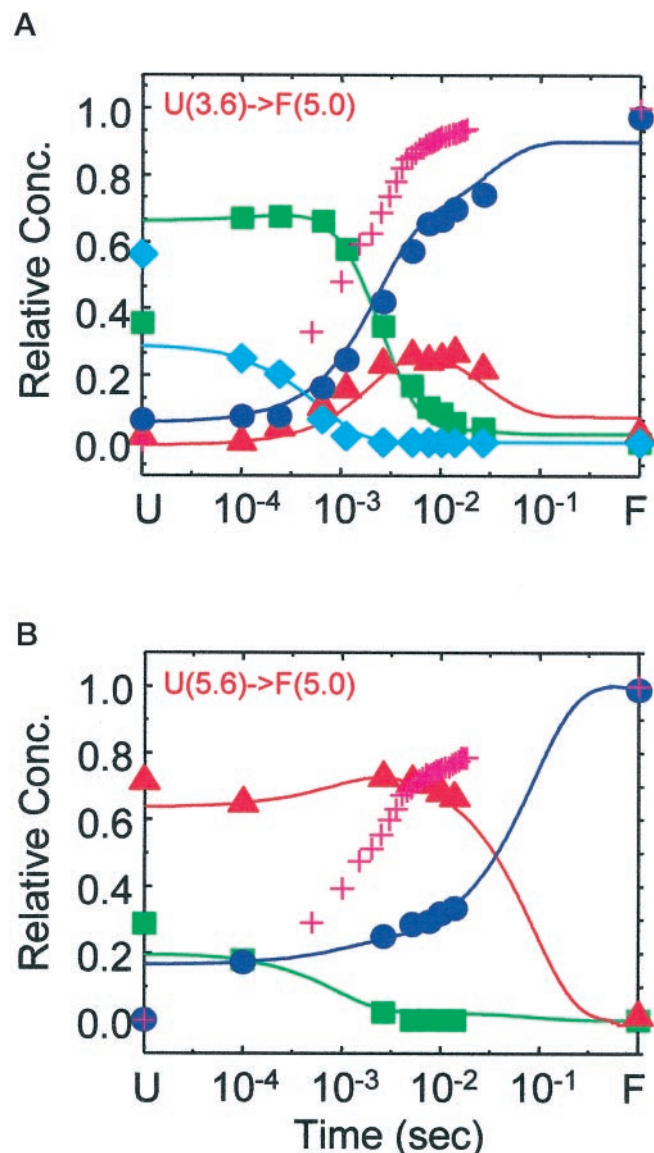


FIG. 4. Time course for the populations of the various heme-ligand coordinated states during the folding of cyt *c* determined from tryptophan fluorescence and resonance Raman scattering measurements. The protein was unfolded at pH 3.6 (A) and pH 5.6 (B) in 4.4 M GdnHCl and refolded at pH 5.0 to a final GdnHCl concentration of 0.7 M. The symbols are defined in the legend for Fig. 3.

when the pH of the unfolded protein is raised from 3.6 to 5.9. This is a result of the kinetic trapping effect of the misfolded HH state, which occupies more than 60% of the population during the nascent phase of folding as shown in Fig. 4B. On the other hand, HH is only transiently populated at approximately 10 ms following the initiation of folding at pH 3.8. This comparison confirms the role of HH as an off pathway intermediate and a deep kinetic trap.

It is intriguing that in contrast to the reaction observed during unfolding, the folding kinetics measured by tryptophan fluorescence spectroscopy deviate significantly from those measured by the resonance Raman scattering as shown in Fig. 4. The loss of the fluorescence intensity appears much earlier than the formation of the native HM state. This is in good agreement with the model illustrated in Scheme 1, in which the protein collapsed to a semi-compact form during the nascent phase prior to the ligand exchange phase. The collapsing of the peptide backbone reduces the average size of the molecule before the ligand exchange equilibrium can take place; *i.e.* the

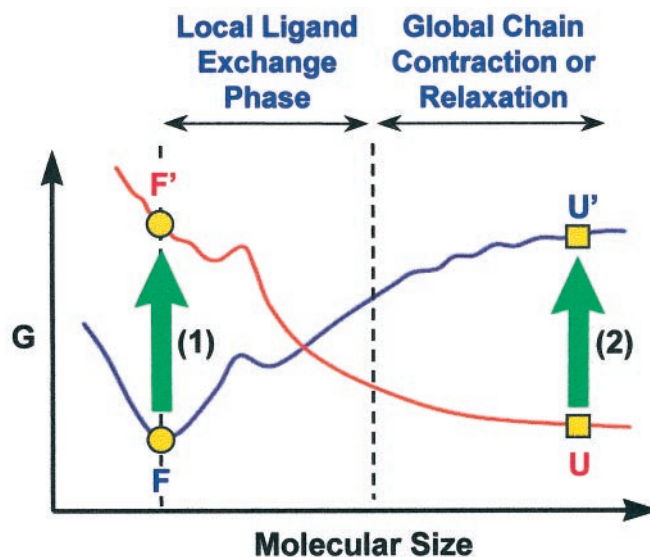


FIG. 5. Schematic illustration of the free energy (*G*) surfaces of cyt *c* during folding (blue curve) and unfolding (red curve). U and U' represent the unfolded state under denaturing and native conditions, respectively, and F and F' represent the folded state under native and denaturing conditions, respectively. The initiation of folding from U to U' and unfolding from F to F' is achieved by rapid exchange of the solvents as indicated by the green vertical arrows. For cyt *c*, the refolding and unfolding reactions share a common ligand exchange phase as indicated.

native structure cannot be realized through ligand exchange reactions until the protein collapses to a semi-compact form. Conversely, the coincidence between the fluorescence and Raman data during unfolding (Fig. 3) suggests that the Met-80 heme linkage cannot be broken until the polypeptide architecture is loosened.

**Concluding Remarks**—The folding of cyt *c* follows a biphasic mechanism as illustrated in Scheme 1. In the nascent phase of folding, the unfolded protein collapses to a semi-compact form in a kinetically controlled process within 100  $\mu$ s (15). It is followed by a thermodynamically controlled heme ligand exchange phase culminating in the native conformation (15). In this work, an analogous biphasic mechanism with a reversed sequence was observed during unfolding (Scheme 2). Unlike the fast ligand exchange rate ( $>10,000$  s<sup>-1</sup>) for the totally unstructured polypeptide, the ligand exchange rates observed during the unfolding reactions are comparable to those observed during the folding reactions (Table I). These results suggest that the ligand exchange phase occurring during unfolding takes place while the polypeptide chain is still in a relatively compact state such that the conformations of the various intermediates are partially constrained by the residual intramolecular tertiary interactions. This proposal is confirmed by the results from tryptophan fluorescence measurements, which indicate that only at the termination of the ligand exchange phase does the protein expand fully into a random coil-like unfolded state.

Based on this and prior studies (15, 21–23), the free energy changes during the folding and unfolding reactions of cyt *c* can be illustrated pictorially in Fig. 5. In this cartoon, the free energy curves plotted against the global molecular size under native (0.0 M GdnHCl) and denaturing conditions (4.4 M GdnHCl) are shown as the dark blue curve and the red curve, respectively. Under native conditions, the energy of the unfolded protein (U') is higher than that of the folded form (F) because of the exposure of the hydrophobic residues to the solvent and the loss of the stabilization energy gained from inter amino acid contacts (hydrogen bonding, electrostatic in-

teractions, hydrophobic-hydrophobic interactions, etc.). As the unstructured polypeptide chain reduces its molecular size, the associated free energy is reduced accordingly (*dark blue curve*). It reaches an energy minimum when the polypeptide chain assumes its native conformation. When the size of the polypeptide molecule is further reduced, the energy starts to climb because of the repulsion imposed by steric interactions between the spatially neighboring amino acid residues. This free energy curve is analogous to the attractive potential energy curve of a typical diatomic molecule, in which the energy minimum of the potential curve is located at the natural bond length of the given molecule. For simplicity, in this energy diagram, the roughness of the potential energy surface is ignored.

On the other hand, under denaturing conditions, the unfolded polypeptide molecule (U) is more stable than its "folded" counterpart (F') because of the high entropy of the unstructured polypeptide backbone and the stabilization energy gained from denaturant-amino acid interactions. Unlike the unique folded state (F) observed under native conditions, there is no well defined folded state (F') under these conditions. As the size of the polypeptide chain is reduced, the energy increases because of the loss of entropy by the exclusion of the denaturant from the protein molecules. As the size of the molecule is further reduced, additional reduction of the stabilization energy is introduced by the repulsion between amino acid residues. The potential energy curve of cyt *c* under denaturing conditions therefore is analogous to the repulsive potential energy curve of a diatomic molecule.

When the unfolding reaction is initiated, shown as the *green arrow* (1) in Fig. 5, the protein is rapidly promoted from the F to the F' state. If the initiation of the reaction is fast enough with respect to the conformational changes of the protein molecule, the entire course of the unfolding reaction can be followed from the F' to the U state (*red curve*). Likewise, the folding reaction from the U' to the F state (*dark blue curve*) can be initiated in a similar fashion indicated as the *green arrow* (2) in Fig. 5. The unfolding kinetics reported here demonstrate that the ligand exchange phase appears when the protein molecule is still in a semi-compact form, and it is followed by further relaxation of the polypeptide chain into a random coil-like structure. In contrast, the folding reaction follows a reversed biphasic process, in which molecular collapse occurs prior to the ligand exchange phase, again located in the *shaded area*. One appealing feature in Fig. 5 is the presence of a crossing point between the *dark blue curve* and the *red curve*, in

which the energies of the polypeptide molecules are the same in the absence and in the presence of denaturant. In other words, the folding and refolding reactions must share a common intermediate with a characteristic molecular size, which could exhibit a molten globule-like structure.

## REFERENCES

- Laurents, D. V., and Baldwin, R. L. (1998) *Biophys. J.* **75**, 428–434
- Brooks, C. L., III, Gruebele, M., Onuchic, J. N., and Wolynes, P. G. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 11037–11038
- Shakhnovich, E. I. (1998) *Fold. Des.* **3**, R110–R111
- Dill, K. A., and Chan, H. S. (1997) *Nat. Struct. Biol.* **4**, 10–19
- Karplus, M. (1997) *Fold. Des.* **2**, S69–S75
- Shakhnovich, E. I. (1996) *Fold. Des.* **1**, R50–R54
- Matthews, C. R. (1993) *Annu. Rev. Biochem.* **62**, 653–683
- Fersht, A. R. (1997) *Curr. Opin. Struct. Biol.* **7**, 3–9
- Pitsyn, O. B. (1998) *Biochemistry* **63**, 367–373
- Prusiner, S. B. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13363–13383
- Kelly, J. W. (1998) *Curr. Opin. Struct. Biol.* **8**, 101–106
- Baldwin, R. L. (1995) *J. Biomol. NMR* **5**, 103–109
- Lazaridis, T., and Karplus, M. (1997) *Science* **278**, 1928–1931
- Onuchic, J. N., Luthey-Schulten, Z., and Wolynes, P. G. (1997) *Annu. Rev. Phys. Chem.* **48**, 545–600
- Yeh, S. R., Han, S., and Rousseau, D. L. (1998) *Acc. Chem. Res.* **31**, 727–736
- Telford, J. R., Wittung-Stafshede, P., Gray, H. B., and Winkler, J. R. (1998) *Acc. Chem. Res.* **31**, 755–763
- Shastri, M. C. R., Sauder, J. M., and Roder, H. (1998) *Acc. Chem. Res.* **31**, 717–725
- Englander, S. W., Sosnick, T. R., Mayne, L. C., Shtilerman, M., Qi, P. X., and Bai, Y. (1998) *Acc. Chem. Res.* **31**, 737–744
- Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W. A. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 11860–11864
- Veeraraghavan, S., Rodriguez-Ghidarpour, S., MacKinnon, C., McGee, W. A., Pierce, M. M., and Nall, B. T. (1995) *Biochemistry* **34**, 12892–12902
- Takahashi, S., Yeh, S. R., Das, T. K., Chan, C. K., Gottfried, D. S., and Rousseau, D. L. (1997) *Nat. Struct. Biol.* **4**, 44–50
- Yeh, S. R., Takahashi, S., Fan, B., and Rousseau, D. L. (1997) *Nat. Struct. Biol.* **4**, 51–56
- Yeh, S. R., and Rousseau, D. L. (1998) *Nat. Struct. Biol.* **5**, 222–228
- Shastri, M. C., and Roder, H. (1998) *Nat. Struct. Biol.* **5**, 385–392
- Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A., and Hofrichter, J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1779–1784
- Sosnick, T. R., Mayne, L., and Englander, S. W. (1996) *Proteins* **24**, 413–426
- Colon, W., Elove, G. A., Wakem, L. P., Sherman, F., and Roder, H. (1996) *Biochemistry* **35**, 5538–5549
- Bhuyan, A. K., and Udgaonkar, J. B. (1998) *Biochemistry* **37**, 9147–9155
- Xu, Y., Mayne, L., and Englander, S. W. (1998) *Nat. Struct. Biol.* **5**, 774–778
- Johnson, W. C., Jr. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17**, 145–166
- Tsong, T. Y. (1976) *Biochemistry* **15**, 5467–5473
- Elove, G. A., Chaffotte, A. F., Roder, H., and Goldberg, M. E. (1992) *Biochemistry* **31**, 6876–6883
- Colon, W., Wakem, L. P., Sherman, F., and Roder, H. (1997) *Biochemistry* **36**, 12535–12541
- Bushnell, G. W., Louie, G. V., and Brayer, G. D. (1990) *J. Mol. Biol.* **214**, 585–595
- Jordan, T., Eads, J. C., and Spiro, T. G. (1995) *Protein Sci.* **4**, 716–728
- Hu, S., Morris, I. K., Singh, J. P., Smith, K. M., and Spiro, T. G. (1993) *J. Am. Chem. Soc.* **115**, 12446–12458