Characterization of a Partially Unfolded High Potential Iron Protein†

Ivano Bertini,*,‡ James A. Cowan,§ Claudio Luchinat,|| K. Natarajan,§ and Mario Piccioli‡

Department of Chemistry, University of Florence, via Gino Capponi, 7, 50121 Florence, Italy, Evans Laboratory, Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210, and Department of Soil Chemistry and Plant Nutrition, University of Florence, P.le delle Cascine 28, 50144 Florence, Italy

Received April 8, 1997⊗

ABSTRACT: A partially unfolded state of the Fe_4S_4-containing high potential iron−sulfur protein from *Chromatium vinosum* has been detected and characterized by NMR spectroscopy following addition of a concentrated solution of guanidinium chloride to the native protein. This intermediate species (i) maintains the polymetallic center, (ii) exhibits a largely collapsed secondary structure, and (iii) undergoes fast cluster decomposition upon oxidation. This information is framed into the knowledge about this class of proteins, and the possible role of this intermediate with respect to the *in vivo* folding/unfolding process is discussed as well its role in the slow hydrolytic degradation characteristic of oxidized HiPIPs.

Protein folding is a fundamental biological problem (Pascher et al., 1996; Dill, 1990; Karplus & Weaver, 1994; Redfield et al., 1994; Johnson & Fersht, 1995; Cohen & Pielak, 1995; Doyle et al., 1996). Among several approaches, NMR has proven to be very successful in yielding information concerning secondary and tertiary structural elements, local mobility, the dynamics of conformational equilibria, and the folding pathways for proteins of moderate size (M_r < 20 000) (Buck et al., 1995, 1996; Neri et al., 1992; Englander & Kallenbach, 1983; Logan et al., 1994; Harding et al., 1991; Feng et al., 1994; Shortle, 1996; Arcus et al., 1995; Smith et al., 1996; Serrano, 1995; Frank et al., 1995; Neira & Rico, 1997). However, less progress has been made in the understanding of folding in the case of metalloproteins. This is particularly true for proteins that carry paramagnetic metal centers, since paramagnetism is considered to be a major obstacle to the application of NMR techniques (Bertini et al., 1993). This attitude has, for instance, precluded investigation of the folding properties of the entire class of iron−sulfur proteins, which is otherwise well characterized (Bertini et al., 1995a, 1996b, and references therein). Many of these proteins do not maintain the native conformation when deprived of the iron−sulfur cluster (Pilon et al., 1992; Li et al., 1996); however, nothing is known concerning either the number of steps required for folding to the native state or the timing and role of cluster assembly on the folding pathway to the native tertiary structure. The possible isolation of a chaperonin for Fe-S proteins (Seaton & Vickery, 1994), if confirmed, makes the problem even more interesting. In the past few years, strategies have been demonstrated that permit application of high-resolution NMR methods to the study of paramagnetic metalloproteins in general (Bertini & Luchinat, 1986, 1996; Xavier et al., 1993; La Mar et al., 1995) and of iron−sulfur proteins in particular (Bertini et al., 1992; Luchinat & Piccioli, 1995). These extend even to a complete solution structure determination (Bertini et al., 1996b; Banci et al., 1994, Pochapsky et al., 1994). We now feel that it is timely to tackle the folding problem of iron−sulfur proteins and that NMR provides a powerful experimental approach. We have chosen to investigate an iron−sulfur protein that contains a single Fe_4-S_4 cluster and belongs to the class of high potential iron−sulfur proteins (HiPIPs) (Beinert, 1990; Carter et al., 1974; Thompson, 1985). These proteins show $[Fe_4\text{-}S_4]^{3+}$ and $[Fe_4\text{-}S_4]^{2+}$ as the biologically relevant redox states (Carter et al., 1974), and their biological function appears to involve electron transport during photorespiration in anaerobic photosynthetic bacteria (Hochkoeppler et al., 1995). Factors that are responsible for the stability of this class of proteins have also been identified (Agarwal et al., 1995; Iwagami et al., 1995; Li et al., 1996; Lui & Cowan, 1994; Bian et al., 1996). In this paper we identify a species that may be a relevant intermediate on the folding/assembly pathway of iron−sulfur proteins and propose that such an intermediate has a key role in mediating the well-known oxidative instability of the entire class of HiPIPs.

EXPERIMENTAL PROCEDURES

Sample Preparation. Reduced, recombinant *Chromatium vinosum* (*C. vinosum* hereafter) HiPIP, including ^{15}N uniformly labeled samples, have been obtained according to a previously reported methodology (Agarwal et al., 1994). Samples for NMR spectroscopy were prepared by dissolving lyophilized protein in a 50 mM phosphate buffer, at pH 5.1 or 6.8. Samples were prepared under an argon atmosphere to avoid the presence of oxygen. A ca. 12 M solution of GdmCl was prepared by dissolving guanidinium chloride (Merck) in a 50 mM phosphate buffer, at pH 5.1 or 6.8. The GdmCl stock solution was deaerated under an argon atmosphere to avoid the presence of oxygen. The oxidized *C. vinosum* HiPIP sample was prepared upon addition to the reduced sample of an excess of ferricyanide solution. The excess of oxidizing agent was eliminated by ultrafiltration under a nitrogen atmosphere. H_2O/D_2O exchange was performed by ultrafiltration under nitrogen with a 50 mM phosphate buffer prepared in D_2O (Sigma Aldrich) at pH 5.1.

† This work was supported by grants from Consiglio Nazionale delle Ricerche, Italy, the Henry and Camille Dreyfus Foundation (Teacher-Scholar Award to J.A.C.), and the National Science Foundation (NYI Award to J.A.C.). J.A.C. and M.P. thank the North Atlantic Treaty Organization for the award of a NATO Collaborative Research Grant.

‡ Department of Chemistry, University of Florence.

§ The Ohio State University.

|| Department of Soil Chemistry and Plant Nutrition, University of Florence.

⊗ Abstract published in *Advance ACS Abstracts,* June 15, 1997.

S0006-2960(97)00810-6 CCC: $14.00

NMR Experiments. All NMR experiments were performed on a Bruker AMX 600 spectrometer. 1D NOE difference spectra were collected using the previously reported methodology (Banci et al., 1989). $^{15}N-^{1}H$ experiments were collected using the HSQC sequence (Bodenhausen & Ruben, 1980). A total of 16 transients were collected for each experiment with a 2048 × 400 data points matrix, using a delay of 2.75 ms for the INEPT transfer (1/4 J_{NH}) and a recycle delay of 850 ms. Longitudinal relaxation times for ^{15}N backbone nuclei were obtained using a refocused INEPT HSQC sequence (Barbato et al., 1992). A total of 16 transients were collected for a 2048 × 170 data points matrix. To build up the antiphase and to refocus the ^{15}N single-quantum coherence, delays of 2.3 and 2.7 ms, respectively, were used, and the recycle delay was 2 s. Decoupling during acquisition was performed through a GARP sequence. The pulse scheme $(\tau-P_{180^{\circ}1H}-\tau)_n$ was applied during the inversion recovery delay to quench cross-correlation effects (Barbato et al., 1992), using a τ delay of 5 ms. Seven time delays were used, 40, 100, 180, 280, 420, 680, and 980 ms, and integrated intensities of peaks were fitted to an exponential decay. Other experimental parameters were as described above.

RESULTS

The Two-State Equilibrium. A concentrated solution of guanidinium chloride (GdmCl hereafter) was added to recombinant, reduced *C. vinosum* HiPIP, at pH 6.8 under an argon atmosphere. For GdmCl concentrations ranging from 0 to 3 M, the NMR signals of the protein show only minor changes, both in the diamagnetic domain and in the region of hyperfine-shifted resonances.

In the range of 3.3−4.4 M GdmCl, a novel set of hyperfine-shifted signals is observed while the hyperfine-shifted signals of the native state were seen to decrease in intensity, as reported in Figure 1. The four most downfield-shifted signals of the new species, which are easily identified during the titration with GdmCl, are reporters of an equilibrium, slow on the NMR time scale, between the native state and a conformational state of the protein that is stabilized by the addition of GdmCl. Given the features described throughout this section, we suggest that formation of this state may be an essential step of the folding pathway. Accordingly, this species is subsequently referred to as an "intermediate" state. The equilibrium is reversible since, upon dilution of the guanidinium solution, the intermediate species converts back to the native species. Following a well-established protocol in protein folding studies by NMR (Evans et al., 1991), saturation transfer experiments have allowed us to correlate the four most shifted signals of the intermediate state to the corresponding signals of the native state. In the case of the present system, the occurrence of well-isolated signals arising from the cluster-coordinated residues allows the use of more sensitive one-dimensional saturation transfer experiments. They are performed upon selective irradiation of signals of interest, using the same experimental approach to collect 1D NOE experiments (Banci et al., 1989). As an example, Figure 2B shows the one-dimensional saturation transfer experiment performed upon selective irradiation of the most downfield-shifted signals of the intermediate state. A peak to signal b of the native species is clearly observed, indicating that the most shifted signal of the intermediate species arises from the same proton which gives rise to signal b of the native species [Cys 63 Hβ_2 (Bertini et al., 1992; Nettesheim et al., 1992)].

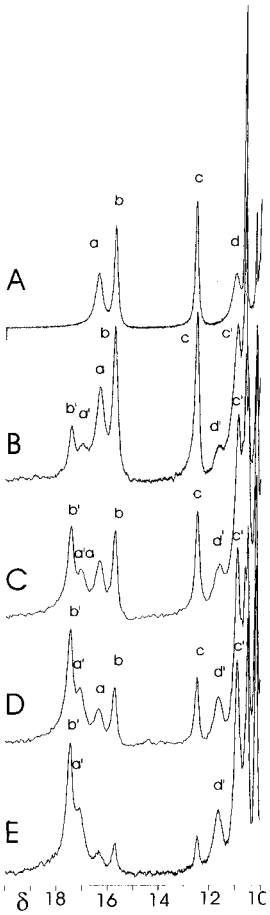

FIGURE 1: 600 MHz ^{1}H NMR spectra of reduced *C. vinosum* HiPIP with increasing GdmCl concentrations: (A) [GdmCl] 0 M, (B) [GdmCl] 3.3 M, (C) [GdmCl] 3.7 M, (D) [GdmCl] 4.0 M, and (E) [GdmCl] 4.4 M, at pH 6.8 and 290 K.

From saturation transfer experiments it is also possible to estimate the exchange rate between the two species. The fractional variation η of signal intensity on signal b of the native species is related to the pseudo-first-order exchange rate constant k by eq 1, where R is the longitudinal relaxation rate of signal b. The fractional variation of intensity of signal b upon saturation of signal b′ is obtained on the basis of eq 2, where $I_{0b} - I_b$ and $I_{0b'}$ are the experimental values as

$$\eta = \frac{k}{R + k} \tag{1}$$

$$\eta = \left[\frac{I_{0b} - I_b}{I_{0b}}\right] = \left[\frac{I_{0b} - I_b}{I_{0b'}}\right]\left[\frac{I_{0b'}}{I_{0b}}\right] \tag{2}$$

obtained from the 1D difference spectra (Ramaprasad et al., 1984) and $I_{0b'}$ and I_{0b} are obtained by integrating peaks in a reference spectrum.

An exchange rate constant k of 6 ± 3 s^{-1} has been obtained from a quantitative analysis of the saturation transfer between signals b and b′ (Cys 63 Hβ_2), by using an R value for signal b of 200 ± 40 s^{-1} (Bertini et al., 1992).

Paramagnetic NMR Signals of the Intermediate State. Under the above conditions, 1D NOEs were performed at 4.4 M GdmCl concentration, corresponding to about 80% of the intermediate state, in order to identify the geminal partner for each of the signals a−d of the intermediate

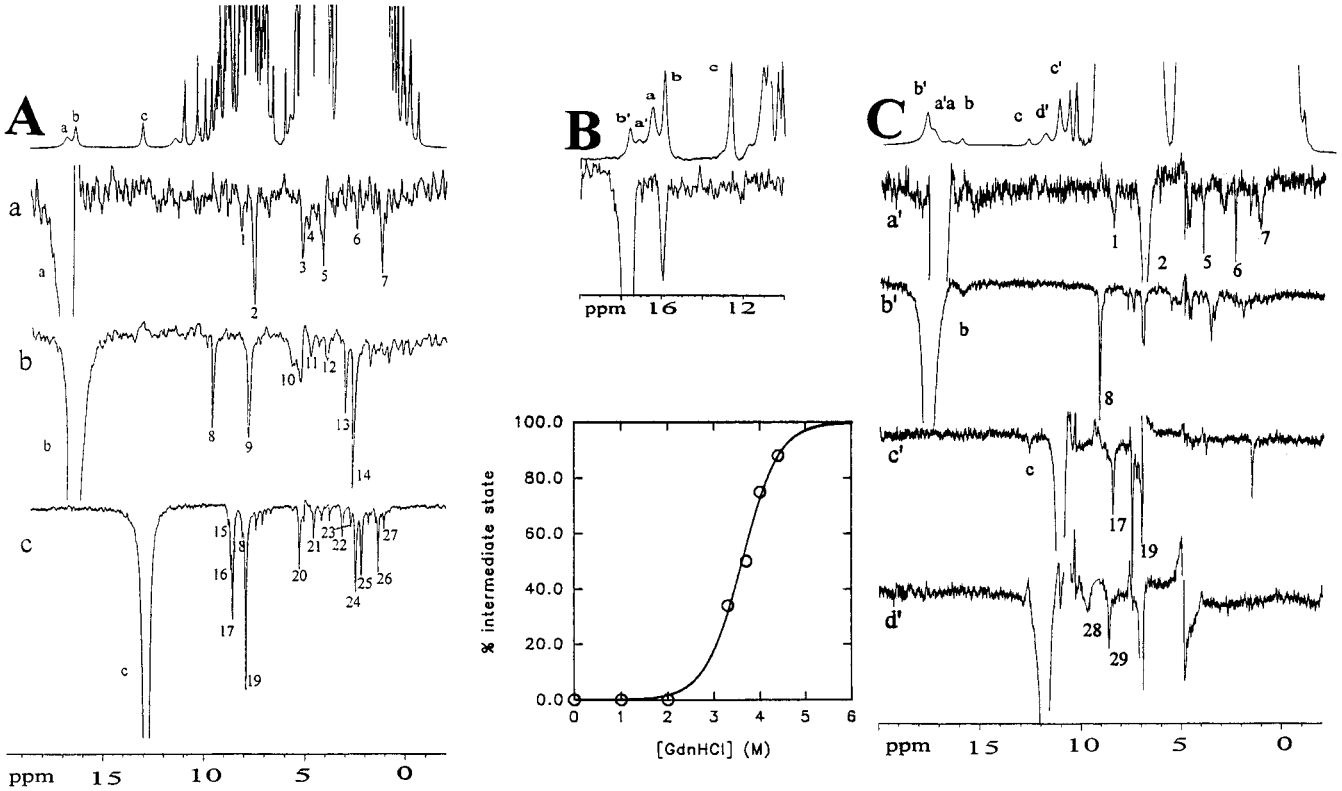

FIGURE 2: 600 MHz ^{1}H NMR spectra of (A) native HiPIP, (B) a mixture of native and intermediate states, and (C) >90% intermediate state for reduced *C. vinosum* HiPIP, obtained at pH 6.8 and 298 K. Spectrum B was obtained in 3.3 M guanidinium chloride and spectrum C in 4.4 M guanidinium chloride. 1D NOE difference spectra observed in the diamagnetic region by saturating some of the hyperfine-shifted signals are reported for the native state (A) and intermediate state (C). Each trace was obtained upon selective irradiation of the signals indicated with lower case letters. The saturation transfer difference spectra reported in (B) was obtained using the same methodology used from 1D NOE difference spectra. A fit of the relative intensity of the signals from the native and intermediate states as a function of guanidinium chloride concentration is shown in (D).

Table 1: Chemical Shifts Observed at 290 K, pH 6.8, in Reduced *C. vinosum* HiPIP in the Native and Intermediate States

signal	native	intermediate	assignment
a	16.25	16.96	Cys 43 Hβ2
b	15.70	17.41	Cys 63 Hβ2
c	12.47	10.93	Cys 77 Hβ1
d	10.85	11.63	Cys 46 Hβ2
a′	7.19	6.62	Cys 43 Hβ1
b′	5.27		Cys 63 Hβ1
c′	7.79	7.31	Cys 77 Hβ2
d′	10.05	9.54	Cys 46 Hβ1
c″	8.23	8.27	Cys 77 Hα

species. The resulting assignment is summarized in Table 1. Seven out of eight Cys Hβ signals, as well as one Cys Hα, could be unambiguously assigned in the intermediate species. Only the Cys 63 Hβ_1 signal was not identified in the intermediate species. Indeed, such a resonance is expected to be rather broad and possibly hidden below the solvent signal. Although the observed changes are larger than those observed, for example, when passing from WT to mutants both in *C. vinosum* (Agarwal et al., 1995; Babini et al., 1996) and *Ectothiorhodospira halophila* HiPIP I (Iwagami et al., 1995; Bertini et al., 1996a), they clearly indicate that no cysteine is detached from iron coordination in the intermediate species. Still, there is clear evidence that structural rearrangements have occurred around the cluster.

Further information on the nature of the structural rearrangement in the proximity of the cluster derives from a comparison of the 1D NOE difference spectra collected on the intermediate state with those of the native protein. In Figure 2, panels A and C show the 1D NOE difference spectra observed in the diamagnetic region by saturating some of the hyperfine-shifted signals. The complete assignment of 1D NOEs on the native state has been already reported elsewhere (Bentrop et al., 1996). We will briefly point out those crucial interresidue NOEs that are meaningful to investigate structural rearrangements of the intermediate state around the cluster. It turns out that three of the long-range NOEs from Cys 43 Hβ2 (signal a, a′) in native HiPIP (5, 6, and 7) remain in the intermediate state. They have been assigned to Hα Val 73, ϵCH$_3$ Met 49, and δCH$_3$ Ile 71. Their detection in the intermediate state implies that the protein scaffold, and side-chain orientations, is essentially conserved in the proximity of Cys 43. For Cys 63 Hβ2 (signal b, b′) the three strong NOEs (9, 13, 14) due to Phe 66 are missing in the intermediate state, while the intraresidue NOE (8) is retained. This means that the position of Phe 66 is substantially changed. It is also likely that both the backbone torsional angles, and the χ_1 angles, for Gly 62 and Cys 63 have changed. Similar behavior is also observed for Cys 77 Hβ1 (signal c, c′). While the intraresidue signals (17, 19) are conserved, there is no evidence of long-range NOEs to Tyr 19 (16, 20, 22, 23). Of the several NOEs observed in the aliphatic region from Cys 77 Hβ1, only one strong signal is observed in the intermediate state, and this does not correspond to any of the NOEs observed in the native state. It is therefore likely that a general structural rearrangement in this part of the protein has removed Tyr 19 from close proximity to the cluster and that the side chains of other neighboring groups (for example, Leu 17) have

experienced a reorientation. No direct comparison is possible for the 1D NOEs from Cys 46 as, under the present conditions, WT signal d is too close to the diamagnetic signals to be investigated via NOE.

Overall, information from 1D NOEs shows that in the proximity of the cluster there are only minor rearrangements close to Cys 43, while much larger modifications that involve structural changes of aromatic residues close to the cluster occur around Cys 63 and Cys 77. This is consistent with the variation in chemical shifts of Cys βCH_2 signals when passing from native to an intermediate state (Table 1). Indeed, the βCH_2 proton signals from Cys 63 and Cys 77 experience larger variations than those from Cys 43 and Cys 46. It is worth recalling that aromatic residues close to the cluster are believed to play a crucial role in stabilizing HiPIPs (Agarwal et al., 1995; Li et al., 1996; Soriano et al., 1996; Bian et al., 1996).

Free Energy Analysis. The inset of Figure 2 shows the relative intensity of the signals from the native and intermediate states as a function of GdmCl concentration. These data were fitted to eq 3, where f_{int} is the molar fraction of the intermediate state, ΔG^{H_2O} is the free energy difference between the native and the intermediate state in the absence of denaturant, [GdmCl] is the concentration of denaturant, and m is a constant that relates to the increase in the degree of exposure of the protein during the perturbation. This equation holds in the assumption of only two states being present, each insensitive to changes in [GdmCl] (linear baseline approximation) (Doyle et al., 1996). Equation 3 is equivalent to the standard eq 4 relating the free energy

$$f_{int} = \frac{\exp(m[\text{GdmCl}]/RT)}{\exp(\Delta G^{H_2O}/RT) + \exp(m[\text{GdmCl}]/RT)} \quad (3)$$

$$\Delta G^{D} = \Delta G^{H_2O} - m[\text{GdmCl}] \quad (4)$$

difference between native and intermediate states to the concentration of denaturant (Pace, 1986), where ΔG^{D} is the free energy of the perturbation at a specific concentration of denaturant defined by [GdmCl]. From these data, $\Delta G^{H_2O} = 22 \pm 3$ kJ mol^{-1} and $m = 6 \pm 1$ kJ mol^{-1} were estimated from a linear regression analysis. The magnitude of ΔG^{H_2O} compares well with typical values obtained for normal denaturation of moderately sized proteins, indicating substantial unfolding. The magnitude is also consistent with the fact that this equilibrium has not previously been observed in variable temperature experiments, lying at considerably higher energies than the Boltzmann factor RT around room temperature (2−3 kJ mol^{-1}). The magnitude of m is, however, significantly smaller than typical values observed in protein folding/unfolding experiments (Johnson & Fersht, 1995), indicating incomplete exposure of the polypeptide chain to the solvent and consistent with cysteine ligation to the iron−sulfur cluster.

Protein Unfolding Monitored through $^1H-^{15}N$ Correlation Experiments and ^{15}N Relaxation Measurements. Increasing amounts of GdmCl were added to a freshly dissolved solution of ^{15}N-enriched, reduced *C. vinosum* HiPIP, at pH 5.2. The effect of adding GdmCl was followed by monitoring the changes in the HSQC spectrum of the protein, according to the common practice in protein folding experiments (Frank et al., 1995; Neri et al., 1989). Figure 3 reports the HSQC spectra collected as a function of GdmCl concentration.

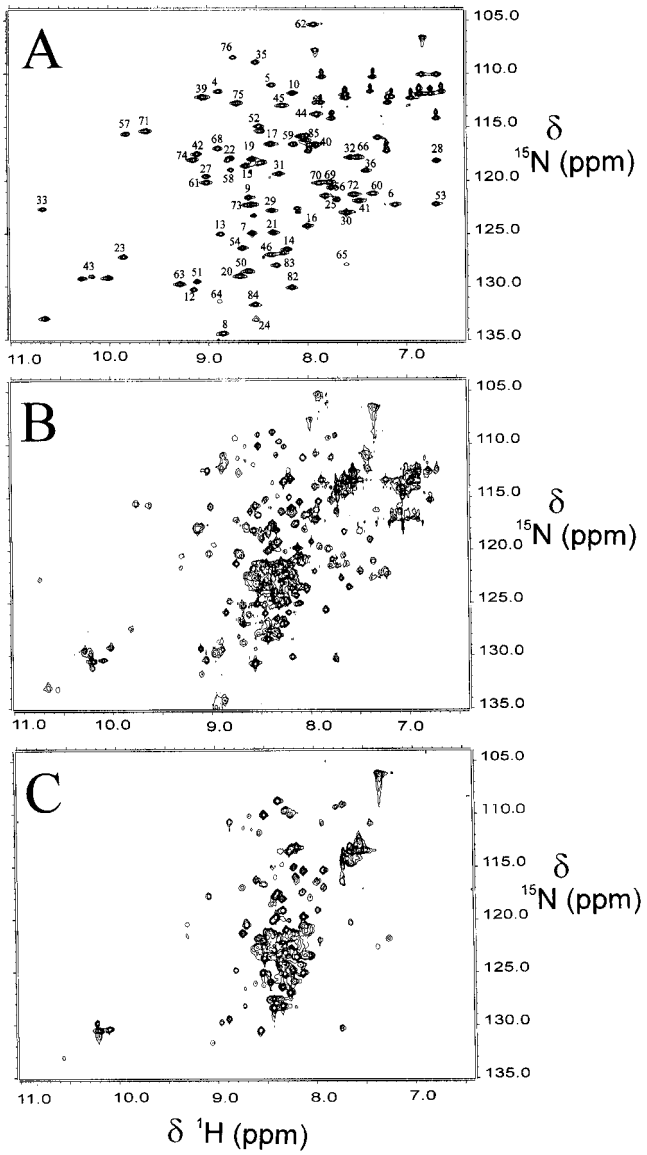

FIGURE 3: 600 MHz (^{1}H) ^{1}H−^{15}N HSQC spectra of (A) native HiPIP, (B) a 1:1 mixture of native and intermediate states obtained at 3.7 M GdmCl, and (C) >90% intermediate state for reduced *C. vinosum* HiPIP, obtained at 4.4 M GdmCl at pH 6.8 and 298 K. For most cross peaks in (A) a safe sequence-specific assignment is available (Banci et al., 1995). These cross peaks are labeled according to the sequential numbering of the corresponding NH groups. The peak assigned to Gln 47 is left out of the spectral window.

While only small variations in the chemical shifts of some signals of the native state (Figure 3A) are observed throughout the GdmCl titration, as can be clearly observed in the 50%−50% mixture shown in Figure 3B, significant changes are observed in the intermediate state. The HSQC spectrum of the fully formed intermediate state is shown in Figure 3C. The HSQC spectrum, except for a few cross peaks, is definitely reminiscent of what is expected for a protein in an unfolded state. The observed reduction of proton chemical shift dispersion as well as the nitrogen shifts, which are largely grouped in the limits of random coil systems (Szilagyi, 1995), indeed confirms that the intermediate species is largely unfolded.

^{15}N-labeled samples allowed us to easily monitor shifts and relaxation properties of the signals from amide groups. Although a firm assignment of signals of the intermediate state could not be performed, the T_1 values of the resonances

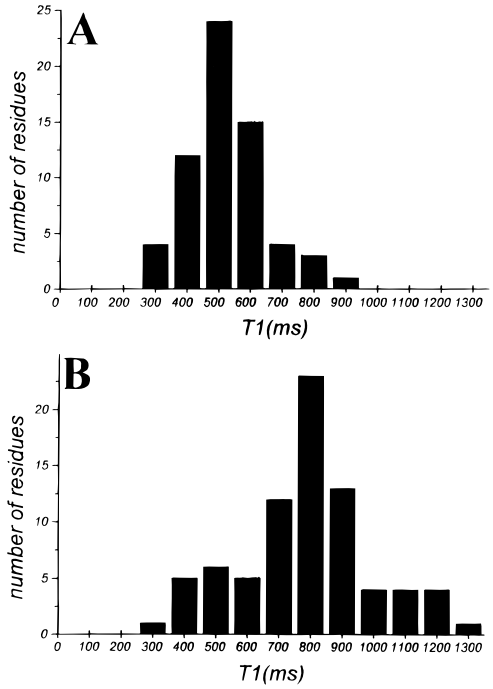

FIGURE 4: Distribution of ^{15}N T_1 values for the NH groups in native (A) and intermediate (B) states of reduced *C. vinosum* HiPIP. The measurements were made at 600 MHz (^1H) and 298 K, using a refocused INEPT HSQC sequence (Barbato et al., 1992).

could be determined and compared to those obtained for the signals of native *C. vinosum* HiPIP.

Figure 4 shows the distribution of ^{15}N T_1 values for the NH groups in native (A) and intermediate (B) states of reduced *C. vinosum* HiPIP. The two distributions are clearly different by statistical criteria [at the 0.05 level for the standard independent t-test (Ott, 1988)]. The most representative values are around 500 and 800 ms for the two species, respectively. The statistical analysis has been performed using only those T_1 values that could be obtained with standard errors smaller than 20%. As longer T_1 values in unfolding experiments are often related to the enhanced mobility of the protein region bearing the NH group (Neri et al., 1989; Farrow et al., 1995), it appears that most residues in the intermediate state show an increased mobility with respect to the native state. Still, there are a few signals of the HSQC spectrum that do not change appreciably, neither in shifts nor in relaxation, when passing to the intermediate state. They most probably belong to residues in those parts of the protein in which tertiary structure of the native protein has been retained.

HSQC in D_2O Solution. To monitor differential changes in the solvent accessibility of the amide region with increasing concentrations of GdmCl, HSQC experiments were performed on samples freshly dissolved in D_2O. We expected that the exchange between H_2O and D_2O would quench the intensity of signals of the intermediate state because of the increased solvent accessibility due to the "open" configuration of the protein in the intermediate state (Buck et al., 1996). Since the native and intermediate forms interconvert over a time scale that is slow with respect to the chemical shift differences, but fast with respect to the duration of the H_2O/D_2O exchange NMR experiment, the quenching of the HSQC signals of the intermediate species should also be transferred to the signals of the native form. Therefore, the addition of GdmCl to D_2O solutions in 2 mM

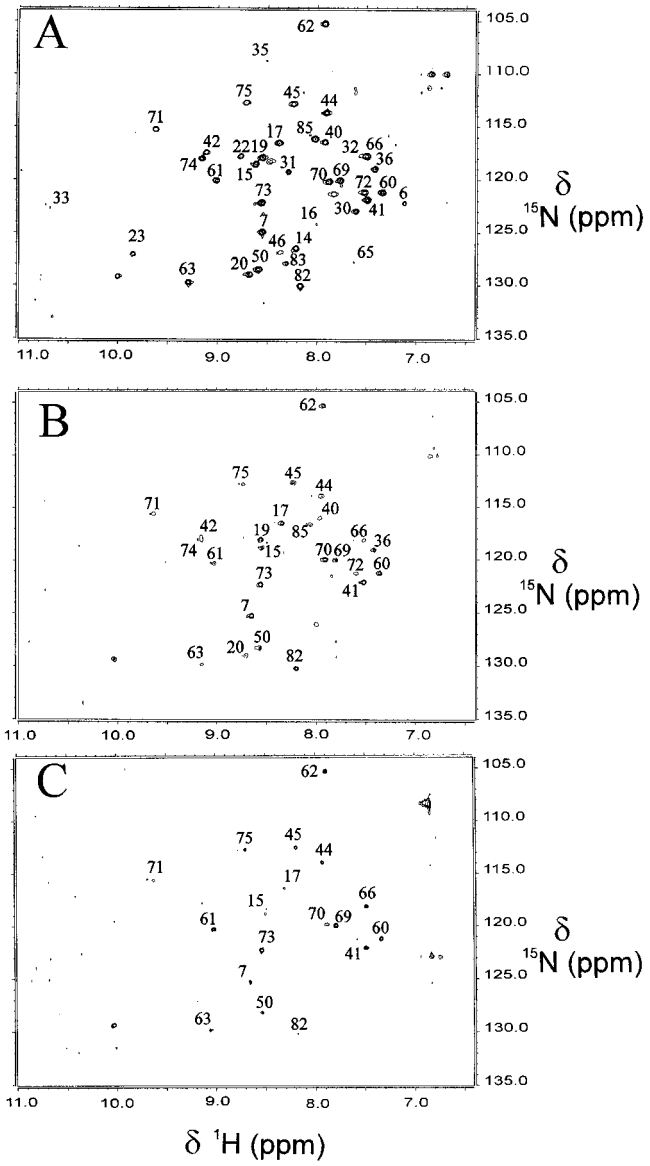

FIGURE 5: 600 MHz (^1H) $^1H-^{15}N$ HSQC spectra, obtained on reduced *C. vinosum* HiPIP at pH 6.8 and 298 K, for (A) native in D_2O, (B) native in D_2O in the presence of 2.0 M guanidinium chloride, and (C) native in D_2O in the presence of 3.0 M guanidinium chloride. The signals in (A) are labeled as in Figure 3. Spectrum A was obtained after 6 h of D_2O exchange. A kinetic profile (not shown) for the first 6 h demonstrates that the missing signals in (A), with respect to the water sample (Figure 3A), are exchanged either immediately or within the first hour, whereas the signals that are still apparent in (A) have not exchanged appreciably over the first 6 h. Spectrum B was obtained after an additional 6 h from the initial addition of 2.0 M guanidinium chloride and spectrum C after a further 6 h (18 h total) from the subsequent addition of guanidinium chloride, raising its concentration to 3.0 M. Again, the missing signals in (B), with respect to (A), are either exchanged immediately or within the first hour, whereas the signals that are still apparent in (C) have not exchanged appreciably over the first 6 h after the second addition of guanidinium chloride. The peak assigned to Gln 47 is left out of the spectral window, and it is observed in all the three experiments (A−C).

GdmCl, still containing more than 98% of the native state, should directly result in a decrease of those signals for which the solvent accessibility in the intermediate state is increased.

The HSQC spectrum of the native protein in D_2O is shown in Figure 5A. Residues which are completely exchanged after 2 h are localized on the first part of the sequence and in the region 51−59. In these regions, corresponding to the

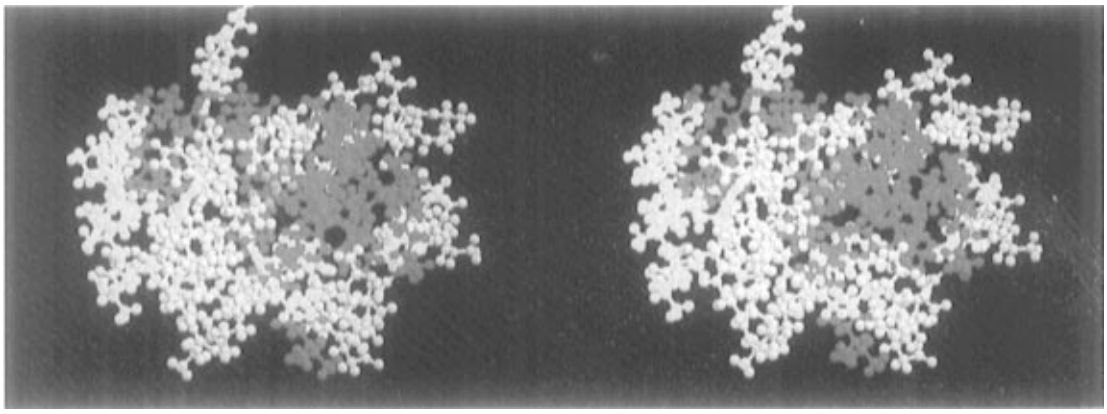

FIGURE 6: Stereoview of the solution structure of native reduced *C. vinosum* HiPIP (Banci et al., 1995), color-coded to show the different exchange properties of NH protons between native and intermediate states: gray, NH protons readily exchanged in D_2O in the native state; yellow, additional NH protons that become exchangeable through the intermediate state; red, NH protons that exchange slowly in both the native and the intermediate states. The cluster atoms are colored in blue. It can be noted that the nonexchangeable backbone NH protons are all grouped on the side of Cys 43 and Cys 46 (right-hand side), which remain shielded from solvent. In contrast, most of the NH protons that become exchangeable through the intermediate state are on the side of Cys 63 and Cys 77 (left-hand side and bottom), thus exposing this side of the cluster to solvent.

more external part of the protein, exchange with solvent is such that HSQC signals are already quenched in the native state (i.e., before the addition of GdmCl).

At 2 M GdmCl concentration (Figure 5B), there are several more resonances of the native state that disappear or significantly decrease in intensity, as expected from the increased solvent exposure of the intermediate state. They are 6, 14, 16, 22−24, 30−35, 46, 64, 65, and 83. Some of the above signals completely disappear, while others show a significant decrease in intensity (signals 34 and 64 are not shown in Figure 5A because they are barely detectable from the noise, while signal 24 is not shown because it is outside of the spectral region shown in Figures 2 and 5). Overall, only a few signals remain in the region 1−39, and at least three of them are located close to position 19, which, as extensively discussed by several authors (Stout, 1982; Agarwal et al., 1995; Iwagami et al., 1995; Li et al., 1996), is supposed to play an important role in driving the stability of the protein. From residue 40 thereon, almost no signal shows distinct behavior on passing from 0 to 2 M GdmCl, with the exception of Cys 46 and residues 64 and 65 that lie between Cys 63 and Phe 66, which is another aromatic residue close to the cluster. The successive addition of GdmCl (Figure 5C) basically reinforces this picture.

DISCUSSION

In a solution of 4.4 M GdmCl, NMR solution spectra provide clear evidence for a distinct partially unfolded state of reduced HiPIP (Figures 2C and 3C). This state is characterized by several unique features, which include (i) an intact $[Fe_4\text{-}S_4]^{2+}$ cluster regularly coordinated by the four original cysteine ligands (cysteines 43, 46, 63, and 77), as evidenced by the presence of all their hyperfine-shifted βCH_2 protons (Figure 2C); (ii) conformationally flexible regions, demonstrated by the clustering of many HSQC peaks in chemical shift ranges typical of random coil structures (Figure 3C), by a significant overall increase in ^{15}N longitudinal relaxation times (Figure 4), and by the loss of many peptide NH signals in D_2O solution; (iii) substantial solvent exposure of about half of the cluster cubane surface, on the side coordinated by cysteines 63 and 77, as shown by the loss of long-range 1D NOEs from the βCH_2 protons of these residues to the groups that shield this side of the cluster in the native state (compare panels A and C of Figure 2); (iv) a relatively large degree of unfolding but, at the same time, substantial retention of tertiary structure, as suggested by a thermodynamic analysis of the NMR data at variable GdmCl concentrations (Figure 2B inset); and (v) a dynamic interconversion with the native state, with a first-order rate constant of about 10 s^{-1} (Figure 2B).

The increased conformational flexibility of the intermediate state is further illustrated by the data summarized in Figure 6, which shows the regions of the protein backbone where the peptide NHs exchange readily with D_2O in the native state (gray), the additional regions where the peptide NHs become readily exchangeable in the intermediate state (yellow), and the core of the protein that maintains nonexchangeable peptide NH in both the native and intermediate states (red). Figure 6 shows that the essential structural features of the intermediate state are a native-like conformation of two stretches of sequence: one (40−47) encompassing Cys 43 and Cys 46, and the other (60−77), encompassing Cys 63 and Cys 77 and protecting the Cys 43−Cys 46 cluster side from solvent. Figure 6 also confirms the exposure of the cluster upon mobilization of the stretches of sequence (from Ala 14 to Thr 24 and from Ala 30 to Ala 32) that protect the cluster from solvent on the Cys 63−Cys 77 side. Within these regions, only Ile 15 and Leu 17 NH are still observable in D_2O solution at 3 M GdmCl, although with a much reduced intensity (Figure 5C).

The $[Fe_4\text{-}S_4]^{3+}$ center, at variance with the reduced $[Fe_4\text{-}S_4]^{2+}$ center, is unstable with respect to the hydrolysis products (Pilon et al., 1992; Li et al., 1996; O'Sullivan & Millar, 1985; Ciurli et al., 1990). The oxidized $[Fe_4\text{-}S_4]^{3+}$ cluster in HiPIPs is protected by a hydrophobic protein environment, which excludes solvent and prevents hydrolysis of the electrophilic cluster (Agarwal et al., 1995; Iwagami et al., 1995). Nevertheless, oxidized HiPIP solutions do have a slight tendency to slowly decompose the hydrolysis products of the cluster and then act as reducing agents toward the oxidized protein molecules [a dismutation process known as "cannibalism" (Holm et al., 1990)]. This tendency might be ascribed to imperfect protection of the cluster, due for instance to modest structural differences between the reduced

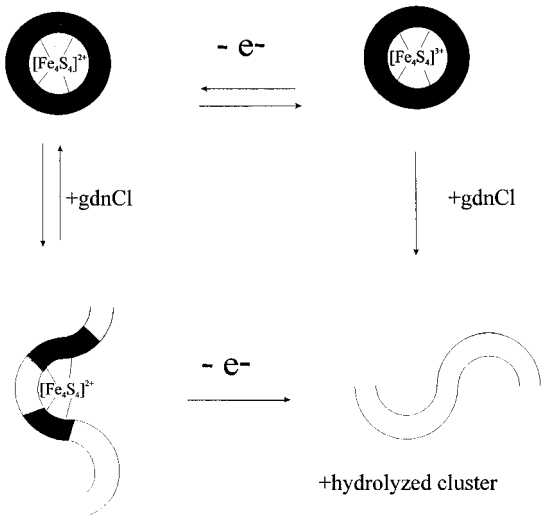

FIGURE 7: Schematic pathway of the folding/unfolding process of *C. vinosum* HiPIP in different oxidation states. While there is a reversible equilibrium between the reduced native state and the reduced intermediate state driven by the presence of GdmCl, no intermediate state can be detected in the oxidized form. The oxidation of the intermediate reduced state and the loss of the cluster in the oxidized form upon addition of GdmCl are both irreversible processes.

and oxidized states which, however, have not been detected at the present level of resolution of the available solution structures (Bertini et al., 1995b, 1996b; Banci et al., 1994). The present work indicates a distinctly different possibility: namely, that the oxidized form could decompose as a result of an equilibrium with the intermediate state described herein. To test this hypothesis, 2 M GdmCl (i.e., an amount that generates less than 2% of the reduced intermediate species, Figure 2 inset) was added to a sample of oxidized HiPIP, in the absence of oxygen. Immediate bleaching of the solution was observed and disappearance of the hyperfine-shifted ^{1}H NMR signals typical of the oxidized species (data not shown). This observation suggests that, although the fraction of intermediate state decreases to undetectable levels upon lowering the concentration of GdmCl, even the tiny amounts of intermediate state in solution in the absence of guanidinium chloride (of the order 0.01%, as estimated from ΔG^{H_2O}, Figure 2B inset) will be able to mediate the slow hydrolytic decomposition of oxidized HiPIP, as experimentally observed.

The overall behavior of the system is sketched in Figure 7. While the intermediate state can be observed and characterized when the native protein is in the reduced form, its exposure to oxygen leads to irreversible loss of the cluster, leading the system to the same irreversibly denatured state reached upon addition of GdmCl to the oxidized form of the protein. The scheme of Figure 7 is only partially analogous to the Born−Haber cycle adopted by McLendon and co-workers (Bixler et al., 1992) and by Pielak and co-workers (Hilgen-Willis et al., 1993) to address the different stability of ferro- and ferricythochrome *c*. While in the case of cythochromes there is a range of GdmCl concentrations in which protein folding can be triggered by an electron transfer event (Pascher et al., 1996), in *C. vinosum* HiPIP it is not possible to isolate a partially unfolded, oxidized form of the protein.

The relevance of the present data to the actual folding/assembling pathway of iron−sulfur proteins is difficult to assess at this stage. Indeed, the mechanism of formation of Fe-S clusters is far from being understood, in the sense of how the iron ions at the required oxidation state are provided and assembled with sulfide ions. However, we believe that the present data do provide hints for the mechanism of formation of the protein. There are no extended elements of secondary structure in HiPIPs, which would energetically contribute to the actual shape of the protein. On the contrary, the present work shows the role of the Fe-S bonds and of the presence of the cluster as a major contribution to the stabilization of the protein. Indeed, in high concentrations of GdmCl the structural feature which remains is provided by the cluster and the coordinated cysteines. When the oxidized cluster is not protected anymore by the hydrophobic groups as it is in the native state, the cluster falls apart and the protein undergoes denaturation. From model chemistry we learn that the reduced $[Fe_4\text{-}S_4]^{2+}$ cluster assembles even in protic solvents (Berg & Holm, 1982), whereas this is not true for the oxidized $[Fe_4\text{-}S_4]^{3+}$ cluster. When, furthermore, attempts are made to mutate cysteines to serines through site-directed mutagenesis in *Escherichia coli*, only the C77S mutant is relatively stable (Babini et al., 1996; Agarwal et al., 1996), whereas the C43S and C46S mutants cannot be isolated and the C63S mutant is very unstable (unpublished results from our laboratory). This, again, is consistent with the observation that the protein by itself has not a stable tertiary structure independent of the cluster. Finally, a behavior analogous to *C. vinosum* HiPIP is shown by the HiPIP I from *E. halophila* (unpublished results from the Florence laboratory). Therefore, even if this working hypothesis needs to be confirmed, the formation of the cysteine-coordinated cluster with a still largely unfolded protein around is a very reasonable candidate intermediate in the *in vivo* folding of the protein.

CONCLUDING REMARKS

Upon addition of GdmCl we have identified a partially unfolded state that is a possible intermediate in the unfolding pathway of a cluster-containing protein. It has been shown that this intermediate state is characterized by a floppy structure in which a large share of the polypeptide chain is in a random coil state and most of the long-range NOEs are lost with the exception of the protein region approaching the cluster in the proximity of Cys 43 and Cys 46. The lack of NOEs is consistent with short correlation times due to the mobility of the unfolded regions and/or to increases in distances. These movements expose the $[Fe_4S_4]^{3+}$ cluster to solvent and lead to hydrolytic degradation. Accordingly, we propose that protein dynamics controls the cluster stability in the two oxidation states.

If then we consider the possibility of obtaining the Cys/Ser mutants (Babini et al., 1996; Agarwal et al., 1996) and the stability in protic solvents (Berg & Holm, 1982) of the $Fe_4S_4^{2-}$ polymetallic centers, the intermediate appears as a candidate in the folding process of the protein.

REFERENCES

Agarwal, A., Tan, J., Eren, M., Tevlev, A., Lui, S. M., & Cowan, J. A. (1994) *Biochem. Biophys. Res. Commun. 197*, 1357−1362.
Agarwal, A., Li, D., & Cowan, J. A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9440−9444.
Agarwal, A., Li, D., & Cowan, J. A. (1996) *J. Am. Chem. Soc. 118*, 927−928.

Arcus, V. L., Vuilleumier, S., Freund, S. M. V., Bycroft, M., & Fersht, A. R. (1995) *J. Mol. Biol. 254*, 305−321.
Babini, E., Bertini, I., Borsari, M., Capozzi, F., Dikiy, A., Eltis, L. D., & Luchinat, C. (1996) *J. Am. Chem. Soc. 118*, 75−80.
Banci, L., Bertini, I., Luchinat, C., Piccioli, M., Scozzafava, A., & Turano, P. (1989) *Inorg. Chem. 28*, 4650−4656.
Banci, L., Bertini, I., Eltis, L. D., Felli, I. C., Kastrau, D. H. W., Luchinat, C., Piccioli, M., Pierattelli, R., & Smith, M. (1994) *Eur. J. Biochem. 225*, 715−725.
Banci, L., Bertini, I., Dikiy, A., Kastrau, D. H. W., Luchinat, C., & Sompornpisut, P. (1995) *Biochemistry 34*, 206−219.
Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., & Bax, A. (1992) *Biochemistry 31*, 5269−5278.
Beinert, H. (1990) *FASEB J. 4*, 2483−2491.
Bentrop, D., Bertini, I., Capozzi, F., Dikiy, A., Eltis, L. D., & Luchinat, C. (1996) *Biochemistry 35*, 5928−5936.
Berg, J. M., & Holm, R. H. (1982) in *Iron Sulfur Proteins* (Spiro, T. G., Ed.) pp 1−66, Wiley-Interscience, New York.
Bertini, I., & Luchinat, C. (1986) *NMR of paramagnetic molecules in biological systems*, Benjamin/Cummings, Menlo Park, CA.
Bertini, I., & Luchinat, C. (1996) *NMR of Paramagnetic Substances*, Elsevier, Amsterdam.
Bertini, I., Capozzi, F., Ciurli, S., Luchinat, C., Messori, L., & Piccioli, M. (1992) *J. Am. Chem. Soc. 114*, 3332−3340.
Bertini, I., Turano, P., & Vila, A. J. (1993) *Chem. Rev. 93*, 2833−2932.
Bertini, I., Ciurli, S., & Luchinat, C. (1995a) in *Structure and Bonding*, Vol. 83, pp 1−54, Springer-Verlag, Berlin and Heidelberg.
Bertini, I., Eltis, L. D., Felli, I. C., Kastrau, D. H. W., Luchinat, C., & Piccioli, M. (1995b) *Chem.—Eur. J. 1*, 598−607.
Bertini, I., Borsari, M., Bosi, M., Eltis, L. D., Felli, I. C., Luchinat, C., & Piccioli, M. (1996a) *JBIC, J. Biol. Inorg. Chem. 1*, 257−263.
Bertini, I., Couture, M. M. J., Donaire, A., Eltis, L. D., Felli, I. C., Luchinat, C., Piccioli, M., & Rosato, A. (1996b) *Eur. J. Biochem. 241*, 440−452.
Bian, S., Hille, C. R., Hemman, C., & Cowan, J. A. (1996) *Biochemistry 35*, 14544−14552.
Bixler, J., Bakker, G., & McLendon, G. (1992) *J. Am. Chem. Soc. 114*, 6938.
Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185.
Buck, M., Schwalbe, H., & Dobson, C. M. (1995) *Biochemistry 34*, 13219−13232.
Buck, M., Schwalbe, H., & Dobson, C. M. (1996) *J. Mol. Biol. 257*, 669−683.
Carter, C. W. J., Kraut, J., Freer, S. T., & Alden, R. A. (1974) *J. Biol. Chem. 249*, 6339−6346.
Ciurli, S., Carrie, M., Weigel, J. A., Carney, M. J., Stack, T. D. P., Papaefthymiou, G. C., & Holm, R. H. (1990) *J. Am. Chem. Soc. 112*, 2654−2664.
Cohen, D. S., & Pielak, G. J. (1995) *J. Am. Chem. Soc 117*, 1675−1677.
Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
Doyle, D. F., Parikh, S., Alcazar-Roman, L., Cole, J. L., & Pielak, G. J. (1996) *Biochemistry 35*, 7403−7411.
Englander, S. W., & Kallenbach, N. R. (1983) *Q. Rev. Biophys. 16*, 521−655.
Evans, P. A., Topping, K. D., Woolfson, D. N., & Dobson, C. M. (1991) *Proteins: Struct., Funct., Genet. 9*, 248−266.
Farrow, N. A., Zhang, O., Forman-Kay, J. D., & Kay, L. E. (1995) *Biochemistry 34*, 868−878.
Feng, Y. Q., Sligar, S. G., & Wand, A. J. (1994) *Nat. Struct. Biol. 1*, 30−35.
Frank, M. K., Clore, G. M., & Gronenborn, A. M. (1995) *Protein Sci. 4*, 2605−2615.
Harding, M. M., Williams, D. H., & Woolfson, D. N. (1991) *Biochemistry 30*, 3120−3128.
Hilgen-Willis, S., Bowden, E. F., & Pielak, G. J. (1993) *J. Inorg. Biochem. 51*, 649−653.
Hochkoeppler, A., Ciurli, S., Venturoli, G., & Zannoni, D. (1995) *FEBS Lett. 357*, 70−74.
Holm, R. H., Ciurli, S., & Weigel, J. A. (1990) *Prog. Inorg. Chem. 38*, 1−74.
Iwagami, S. G., Creagh, A. L., Haynes, C. A., Borsari, M., Felli, I. C., Piccioli, M., & Eltis, L. D. (1995) *Protein Sci. 4*, 2562−2572.
Johnson, C. M., & Fersht, A. R. (1995) *Biochemistry 34*, 6795−6804.
Karplus, M., & Weaver, D. L. (1994) *Protein Sci. 3*, 650−668.
La Mar, G. N., Chen, Z., & de Ropp, J. S (1995) in *Nuclear Magnetic Resonance of Paramagnetic Macromolecules* (La Mar, G. N., Ed.) pp 55−74, Kluwer Academic Publishers, Dordrecht.
Li, D., Agarwal, A., & Cowan, J. A. (1996) *Inorg. Chem. 35*, 1121−1125.
Logan, T. M., Theriault, Y., & Fesik, S. W. (1994) *J. Mol. Biol. 236*, 637−648.
Luchinat, C., & Piccioli, M. (1995) in *NMR of Paramagnetic Macromolecules, NATO ASI Series* (La Mar, G. N., Ed.) pp 1−28, Kluwer Academic Publishers, Dordrecht.
Lui, S.-M., & Cowan, J. A. (1994) *J. Am. Chem. Soc. 116*, 4483−4484.
Neira, J. L., & Rico, M. (1997) *Folding Des. 2*, R1−R11.
Neri, D., Wider, G., & Wüthrich, K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4397−4401.
Neri, D., Billeter, M., Wider, G., & Wüthrich, K. (1992) *Science 257*, 1559−1563.
Nettesheim, D. G., Harder, S. R., Feinberg, B. A., & Otvos, J. D. (1992) *Biochemistry 31*, 1234−1244.
O'Sullivan, T., & Millar, M. M. (1985) *J. Am. Chem. Soc. 107*, 4096−4097.
Ott, L. (1988) *Introduction to statistical methods and data analysis*, P. W. S.Kent Publishing Co., Boston.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
Pascher, T., Chesick, J. P., Winkler, J. R., & Gray, H. B. (1996) *Science 271*, 1558−1560.
Pilon, M., Rietveld, A. G., Weisbeek, P. J., & de Kruijff, B. (1992) *J. Biol. Chem. 267*, 19907−19913.
Pochapsky, T. C., Mei Ye, X., Ratnaswamy, G., & Lyons, T. A. (1994) *Biochemistry 33*, 6424−6432.
Ramaprasad, S., Johnson, R. D., & La Mar, G. N. (1984) *J. Am. Chem. Soc. 106*, 3632−3635.
Redfield, C., Smith, R. A. G., & Dobson, C. M. (1994) *Nat. Struct. Biol. 1*, 23−29.
Seaton, B. S. & Vickery, L. E. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2066−2070.
Serrano, L. (1995) *J. Mol. Biol. 254*, 322−333.
Shortle, D. R. (1996) *Curr. Opin. Struct. Biol. 6*, 24−30.
Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., & Dobson, C. M. (1996) *J. Mol. Biol. 255*, 494−506.
Soriano, A., Li, D., Bian, S., Agarwal, A., & Cowan, J. A. (1996) *Biochemistry 35*, 12479−12486.
Stout, C. D. (1982) in *Metal Ions in Biology* (Spiro, T. G., Ed.) John Wiley and Sons, New York.
Szilagyi, L. (1995) *Prog. Nucl. Magn. Reson. Spectrosc. 27*, 325−443.
Thompson, A. J. (1985) in *Metalloproteins* (Harrison, P., Ed.) p 79, Verlag Chemie, Weinheim, FRG.
Xavier, A. V., Turner, D. L., & Santos, H. (1993) *Methods Enzymol. 227*, 1−16.

BI970810W