The Three-Dimensional Solution Structure of the Reduced High-Potential Iron−Sulfur Protein from *Chromatium vinosum* through NMR†

Lucia Banci,‡ Ivano Bertini,*,‡ Alexander Dikiy,‡ Dieter H. W. Kastrau,‡ Claudio Luchinat,§ and Pornthep Sompornpisut‡

Department of Chemistry, University of Florence, Via Gino Capponi 7, 50121 Florence, Italy, and Institute of Agricultural Chemistry, University of Bologna, Viale Berti Pichat 10, 40127 Bologna, Italy

Received July 29, 1994; Revised Manuscript Received October 17, 1994⊗

ABSTRACT: The ^{1}H NMR assignment of the reduced HiPIP from *Chromatium vinosum* available in the literature [Gaillard, J., Albrand, J.-P., Moulis, J.-M., & Wemmer, D. E. (1992) *Biochemistry 31*, 5632−5639] has been extended up to 85% of the total protein protons. Ninety percent of the nitrogens have been assigned. Then the solution structure has been obtained using as many as 1147 meaningful NOE connectivities. The protein is sizably paramagnetic even though the ground state is a singlet. Nevertheless, the final RMSD values are 0.62 and 1.19 Å for the backbone and the heavy atoms, respectively. These values compare well with those for diamagnetic proteins of the same size. The solution structure is discussed in the light of the available structural information from X-ray data.

The determination of the three-dimensional structure of paramagnetic proteins in solution is still quite difficult since nuclear T_i^{-1} (i = 1,2) values hamper the detection of NMR[1] dipolar and scalar connectivities. A further problem arises from the choice of the force field parameters that are used for metal ions in energy minimization procedures and MD simulations. The former issue has been addressed by us (Banci et al., 1994b) and others (Oh & Markley, 1990; Lecomte et al., 1991; Nettesheim et al., 1992; Xavier et al., 1993) during the last three years. The latter problem has been empirically overcome for zinc-containing proteins (Merz & Kollman, 1989; Hoops et al., 1991) and for open shell metal ions in iron−sulfur proteins (Banci et al., 1992a), hemoproteins (Case & Karplus, 1979; Banci et al., 1994c), and copper-containing proteins (Shen et al., 1990; Banci et al., 1992b).

An attempt to solve the structure of a Fe_2S_2 protein has been reported in the literature (Pochapsky et al., 1994), but the quality of the determined structure suffers from the lack of connectivities, particularly in the vicinity of the metal ions.

We have recently reported the solution structure of the high-potential iron−sulfur protein (HiPIP) from *E. halophila* (HiPIP I) in its reduced form (Banci et al., 1994a), which has been expressed in *E. coli* (Eltis et al., 1994). The quality of the structure is comparable to that of diamagnetic proteins of homologous size (73 amino acids).

We would like to report here the solution structure of the reduced HiPIP from *Chromatium vinosum* having 85 amino acids. The protein is paramagnetic with about 0.5 unpaired electrons per iron (Phillips et al., 1974) and T_1 values of the Cys β-CH_2 protons ranging from 2 to 7 ms (Bertini et al., 1991). The proton assignment available (Gaillard et al., 1992; Bertini et al., 1992) is largely based on the solid state structure of the oxidized species, which is available at a resolution of 2.0 Å (Carter et al., 1974). We have extended the proton assignment and assigned the ^{15}N signals through 3D ^{1}H−^{1}H TOCSY/^{1}H−^{15}N HMQC and ^{1}H−^{1}H NOESY/^{1}H−^{15}N HMQC spectroscopic experiments. By using the obtained assignments, we have proceeded with the structure determination and with the comparison between the solution and the solid state structure. The degree of resolution of the present structure of a paramagnetic protein is comparable to that of diamagnetic proteins of the same size. The comparison between the solution and the solid state structures suggests that the two redox states have very similar structure. Some of the observed differences may be indicative of real, albeit modest, structural rearrangements upon change in redox state.

EXPERIMENTAL PROCEDURES

All chemicals used were of the best quality available. Native HiPIP from *Ch. vinosum* was prepared and purified as described earlier (Bartsch, 1978). ^{15}N enriched protein was isolated from bacteria grown in $(^{15}NH_4)_2(SO_4)$ containing media (0.5 g/L). The purification of the latter was identical to the procedure for the purification of the native protein.

Samples for NMR spectroscopy (3 mM native protein and 1.5 mM ^{15}N labeled protein) were prepared in 50 mM phosphate buffer (either in H_2O or D_2O), pH 5.1. Sample preparation as well as buffer exchange was performed by five repetitions of ultrafiltration with this buffer in a Amicon cell equipped with a YM1 membrane.

† This work is partially supported by "Progetto Finalizzato Biotecnologie" and "Comitato Scienze Agrarie" of CNR, Italy. Coordinates of the RMD family of the structures and of the RMDw structure are being submitted to the Brookhaven Protein Data Bank.

* Please address correspondence to this author at the Department of Chemistry, University of Florence, Via Gino Capponi 7, 50121 Florence, Italy. Phone: (39) 55-275-7549; Fax: (39) 55-275-7555.

‡ University of Florence.

§ University of Bologna.

⊗ Abstract published in *Advance ACS Abstracts,* December 1, 1994.

[1] Abbreviations: NMR, nuclear magnetic resonance; 1(2,3)D, one-(two-, three-) dimensional; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; HMQC, heteronuclear multiple quantum correlation spectroscopy; DG, distance geometry; REM, restrained energy minimization; RMD, restrained molecular dynamics; *Ch. vinosum, Chromatium vinosum; E. halophila, Ectothiorhodospira halophila; E. coli, Escherichia coli.*

0006-2960/95/0434-206$09.00/0 © 1995 American Chemical Society

NMR experiments were performed on a AMX Bruker spectrometer operating at 600.14 MHz proton Larmor frequency. All 2D homonuclear spectra were acquired in phase sensitive TPPI mode. 2D NOESY spectra were collected using the standard pulse sequence RD−90°−t_1−90°−t_m−90°−AQ (Macura & Ernst, 1980), with mixing times of 100, 35, and 15 ms. The repetition time was varied between 550 ms and 1.15s. In the NOESY experiment with mixing time of 15 ms, a trim pulse (2 ms) was applied before the beginning of the pulse sequence. 2D TOCSY spectra were acquired with the RD−90°−t_1−spin lock−AQ pulse sequence (Braunschweiler & Ernst, 1983; Davis & Bax, 1985), where the spin lock field along the x axis is obtained with the MLEV-17 pulse sequence (Bax & Davis, 1985). The repetition time for all TOCSY experiments was 1.1 s, while the spin lock time was set to different values (25, 45, and 75 ms).

All 2D spectra consisted of 4K data points in the F_2 dimension. From 800 to 1024 experiments were recorded in the F_1 dimension, 32−128 scans per experiment. Raw data were processed using a sine squared window function shifted by $\pi/2$ in both dimensions. In order to increase the resolution of cross peaks close to the diagonal, shifts of $\pi/3$ were also used. A polynomial base-line correction was applied in both directions. Data were always zero-filled to obtain 2K × 1K data points matrices. The experiments for the detection of connectivities in the diamagnetic region were performed with a spectral width of 7400 Hz. The spectra were calibrated assuming a chemical shift of 4.91 and 4.79 ppm for the water signal with respect to DSS at 290 and 300 K, respectively.

To obtain NOE constraints for protons experiencing paramagnetic effects, additional experiments were carried out. A series of 1D NOE experiments was acquired saturating the hyperfine shifted and well-resolved signals of the β-CH_2 protons of cysteines. These experiments were acquired by using a previously reported procedure (Banci et al., 1989; Dugad et al., 1990). The water signal was suppressed by applying the SUPERWEFT pulse sequence 180−τ−90°−AQ, where the recycle time was 400 ms and τ was adjusted to values between 80 and 220 ms.

Heteronuclear 2D and 3D experiments with the ^{15}N enriched HiPIP were carried out in phase sensitive TPPI mode at 300 K using a 5 mm reverse detection probe. In the case of 2D spectra, the HMQC pulse sequence was used (Bax et al., 1983a,b), and 2K data points were collected in the F_2 direction with 512 increments in the F_1 direction, 64 scans per experiment. Refocusing time was 5 ms, recycling time 1.5 s. 3D TOCSY-HMQC and 3D NOESY-HMQC spectra (Griesinger et al., 1987, 1989; Kay et al., 1989) were collected with 512 data points in the F_3 (^1H) dimension. Sixty-four increments in the F_2 (^{15}N) and 128 increments in the F_1 (^1H) dimensions were accumulated, with 16 scans per increment. A NOESY mixing time of 100 ms and a TOCSY spin lock time of 30 ms were used. Decoupling of nitrogens during acquisition was achieved by WALTZ-16 pulse sequence (Shaka et al., 1983). Nitrogen chemical shifts were calibrated by assigning a shift of 21.4 ppm to the $(^{15}NH_4)_2(SO_4)$ signal at 300 K with respect to liquid ammonia.

Data processing was performed using a standard Bruker software package. The 2D and 3D spectra were analyzed on IBM RS 6000 computers using the program XEASY (ETH Zürich) (Eccles et al., 1991). For the structure calculation, volumes of assigned NOESY cross peaks were integrated using the routines provided by XEASY. Most of the volumes were determined using lineshape integration. The remaining peaks, for which no lineshape was found, were integrated manually. With the help of the program CALIBA (Güntert et al., 1991), the volumes were then translated into upper distance limits. During preliminary structure calculations, the factors for the volume-to-distance conversion were evaluated by plotting volumes of peaks arising from pairs of protons in rigid parts of the molecule. At all the following steps of the structure calculation, the calibration was then refined using the interatomic distances determined from previous structures. Five different calibration classes were used (Güntert et al., 1991):

1: intraresidue except NH, Hα, Hβ
2: sequential and intraresidue NH, Hα, Hβ
3: medium range
4: long range backbone
5: long range

The obtained experimental distance constraints were then used to calculate the three-dimensional protein structure using the distance geometry (DG) program DIANA (Güntert et al., 1991). For the DG calculations, the default minimization strategy offered by the program was applied using the following parameters: the first and last minimization levels were set to 0 and 85, respectively. The maximal number of target function evaluations was 2000 during all levels except the final three ones. Here, a number of 5000 has proved to yield an optimal convergence for the calculated structures. The number of evaluations of the target function in the first levels was set to 200, for the last three to 500.

In addition to the standard procedure of distance geometry calculations, extensive use was made of the redundant angle strategy (REDAC) (Güntert & Wüthrich, 1991) for the definition of starting conformations. Constraints for dihedral angles were only accepted by the program if in at least 10 structures the local target function for the residue in question was equal to or smaller than 0.4 Å^2; during the last four REDAC cycles the local target function cutoff was reduced to 0.1 Å^2.

The iron−sulfur cluster was included as previously described (Banci et al., 1994a) through the addition of an artificial amino acid to the residue library used by the program DIANA. The artificial residue consists of a cysteinyl residue in which the thiol hydrogen (H) was replaced by an iron atom (Fe) at the proper distance and by adding to the latter, through another covalent bond, the sulfur atom (S) constituting the inorganic sulfide. Bond lengths and angles used in this construction were derived from the average of these parameters observed in known X-ray structures of HiPIPs and cubane-containing model systems (Carter et al., 1974; Carter, 1977; Carney et al., 1988; Holm et al., 1990; Breiter et al., 1991; Rayment et al., 1992; Benning et al., 1994).

The four modified cysteines can mimic the geometry of the cluster if the four iron atoms and the four terminal sulfur atoms (S) are linked through covalent bonds along the edges of the cubane. Four of these bonds belong to the newly defined residues, and eight are treated by DIANA as "special covalent bonds" by imposing the appropriate upper and lower

distance limits. Additional "special covalent bonds" were introduced to maintain the shape of the cube, as previously described (Banci et al., 1994a). A total of 26 links were used. This construction allows us to define a rigid cluster, leaving at the same time undefined the chirality of the peptide folding around it. The correct chirality was spontaneously achieved in all of the accepted structures from the DG calculations.

Stereospecific assignments were obtained with the help of the program GLOMSA (Güntert et al., 1991) in the final stages of the DG structure calculations.

Restrained energy minimization (REM) and restrained molecular dynamics (RMD) calculations on the 15 structures with the lowest target function values belonging to the DG family were carried out using the SANDER module (Pearlman & Case, 1991) of the AMBER 4.0 program package (Pearlman et al., 1991). The force field parameters for all the residues, except those for the atoms of the cluster and the iron-bound cysteines, were the standard AMBER "all atoms" parameters (Weiner et al., 1986). For the iron–sulfur cluster and cysteines bound to it the force field parameters previously reported were used (Banci et al., 1992a, 1993). Each DG structure was energy minimized and then subjected to RMD calculations. These calculations were performed for 36 ps, heating the system from 0 K to the equilibrium temperature of 300 K. The molecules were coupled to a thermal bath at 300 K with a coupling constant of 0.1 ps (Berendsen et al., 1984). The calculations were performed *in vacuo* with a distance-dependent dielectric constant. The time step for RMD calculations was 1.5 fs. The bond lengths in all RMD calculations were kept rigid by using the SHAKE algorithm (van Gunsteren & Berendsen, 1977) whereas they were allowed to vary in all REM calculations. The nonbonded interactions were evaluated with a cutoff of 10 Å, and the pair list for the evaluation of the nonbonded interactions was updated every 20 steps. During the RMD calculations a NOE-weighted potential function was applied using the SANDER module of AMBER. The mixed linear–harmonic potential used in SANDER (Pearlman & Case, 1991) is:

$$E = \begin{cases} k(d_{i,j} - d_{i,j}^{u})^2 \text{ for } d_{i,j}^{u} < d_{i,j} < d_{i,j}^{u} + 0.5 \text{ Å} \\ k(2.0 - d_{i,j})^2 \text{ for } 1.5 \text{ Å} < d_{i,j} < 2.0 \text{ Å} \\ 0 \text{ for } 2.0 \text{ Å} < d_{i,j} < d_{i,j}^{u} \\ k'(d_{i,j} - d_{i,j}^{u} - 0.25) \text{ for } d_{i,j} > d_{i,j}^{u} + 0.5 \text{ Å} \\ k'(1.75 - d_{i,j}) \text{ for } d_{i,j} < 1.5 \text{ Å} \end{cases}$$

where $d_{i,j}^{u}$ represents the upper limit for the distance $d_{i,j}$ between atoms i and j. The force constant k was set to 133.8 kJ mol^{-1} Å^{-2} and k' to 133.8 kJ mol^{-1} Å^{-1}. The $d_{i,j}^{u}$ values used by SANDER were the upper limit distances between protons i and j used by DIANA.

The final 12.0 ps of the RMD simulation of each structure of the family were used to generate 15 average structures with the program CARNAL (Ross, 1994). The average structures were then subjected to REM calculations.

The mean structure of the RMD family was subjected to further RMD calculations in water. The procedure was identical to that for calculations *in vacuo,* except that the protein was solvated by a 10 Å thick shell of water molecules. A fixed dielectric constant was used. The trajectory was calculated for 144 ps, and the last 86 ps were used for structure averaging.

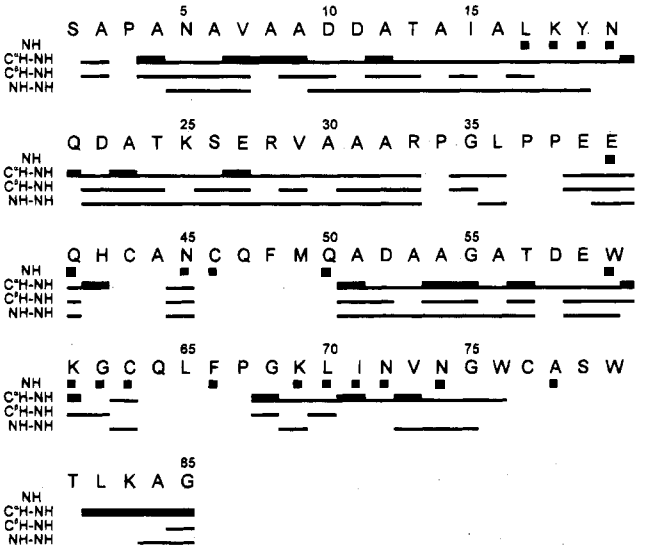

FIGURE 1: Schematic representation of the sequential connectivities involving NH, Hα and Hβ protons in the reduced HiPIP from *Ch. vinosum* obtained through standard 2D experiments. Dipolar NH–Hα connectivities are divided into strong (thick lines) and weak (thin lines). Filled squares indicate slowly exchanging NH protons.

The analysis of NMR spectra and all DG, REM, and RMD calculations (double precision accuracy) were performed on IBM RISC 6000/530 computers.

RESULTS AND DISCUSSION

^{1}H and ^{15}N Assignment. An extensive sequence-specific assignment of the ^{1}H NMR spectrum of reduced HiPIP from *Ch. vinosum* is available in the literature (Gaillard et al., 1992). To observe stronger NOE effects and, therefore, stronger NOE connectivities, we carried out our NOESY experiments at a lower temperature (290 K). Consequently, we adapted the existing assignment to the chemical shift values measured under our experimental conditions. Sequence-specific assignments were achieved by conventional 2D and 3D methods.

The analysis of two-dimensional homonuclear experiments recorded for detecting connectivities in the diamagnetic region (Wüthrich, 1986; Arseniev et al., 1988; Schultze et al., 1988) allowed us to assign the aliphatic spin systems of 69 amino acids. Most of the spin systems were identified from TOCSY spectra recorded with spin lock times of 45 and 75 ms, performed in H_2O and D_2O solutions at 300 K. These assignments were further confirmed by 3D TOCSY-HMQC experiments.

Comparative analysis of the dipolar (mixing times 35 and 100 ms) and scalar connectivities observed in the fingerprint region of the 2D NOESY and TOCSY experiments as well as analysis of 3D NOESY-HMQC and TOCSY-HMQC maps allowed the sequence-specific assignment for the backbone protons of 69 amino acids (Figure 1). CH(i)–NH(i+1) and NH(i)–NH (i+1) connectivities led to the identification of several segments, the largest one being constituted by 31 amino acids (Pro-3–Arg-33). Sequence-specific assignments achieved by 3D experiments were crucial to distinguish degenerate NH resonances. As examples, a few ω_{1-3} planes extracted from the 3D NOESY-HMQC spectrum are presented in Figure 2. Lower values

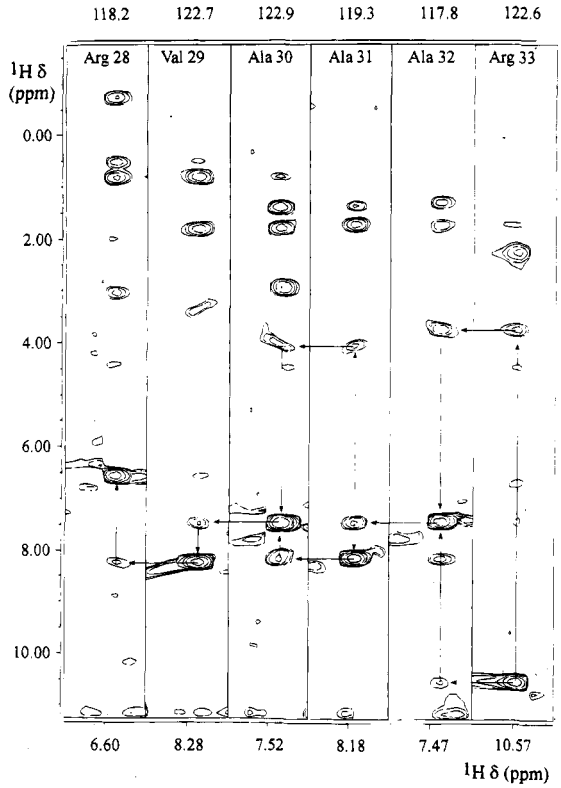

FIGURE 2: F_1−F_3 slices extracted from the 3D ^{1}H−^{1}H NOESY/^{1}H−^{15}N HMQC spectrum of the reduced HiPIP from *Ch. vinosum*, recorded at 300 K, 0.05 M phosphate buffer in H_2O solution, pH 5.1. The sequential assignment for residues 28−33 is indicated by arrows. On the top, the ^{15}N F_2 frequency of each slice is reported.

of pH and temperature give rise to a decrease of the NH proton exchange rate and, therefore, favor the detection of the connectivities in which the NH protons are involved. Indeed, NH−NH connectivities for the segments 4−7, 9−19, 21−33, 35−36, 39−41, 44−45, 50−57, 58−60, 62−63, 68−69, 72−75, and 83−85 are observed at pH 5.1 and 290 K (Figure 1). Besides the sequential backbone connectivities, we also observe in each plane dipolar connectivities between the NH protons and their spatial neighbors. The 3D experiments led also to the assignment of the backbone ^{15}N signals (Table 1). Sequential connectivities have not been observed for the entire protein chain due to the presence of 5 prolines which do not contain the amide NH, and due to the presence of amino acids close to the cluster (Figure 1). The latter are affected by the paramagnetism of the [4Fe-4S] core.

Comparison of spin patterns and backbone sequence-specific assignment has led us to the sequence-specific assignment of 66 residues. In the aromatic region the patterns of three tryptophans (out of three) are evident. NOE's between β-CH_2 protons of the already assigned residues 60 and 76 and Trp aromatic spin systems complete the assignment of the latter residues. By exclusion, the remaining Trp aromatic spin system is assigned to Trp-80.

At this point we have assigned 69 amino acids. Now we could have proceeded by obtaining a structural model from our distance geometry calculations and then pursuing the assignment through experimental NOE's. However, further assignments are available in the literature (Gaillard et al., 1992; Bertini et al., 1992) based on the X-ray data obtained on the oxidized protein (Carter et al., 1974). Presumably, the oxidized protein is also a good model for the reduced protein. Therefore, we have proceeded in the same way, while extending or even correcting the previous assignment. Meaningful extension has been reached in the vicinity of the cluster.

The assignment of the previously unassigned amino acids Ser-1, Ala-2, and Ala-78 is achieved by observing interresidue dipolar connectivities and scalar connectivities. In the X-ray structure Hζ1 of Trp-60 is close to the CH_3 of Ala-2 (about 3 Å). This allowed us to assign the methyl group of Ala-2. The position of Hα of Ala-2 was determined from the observation of the connectivity with Nε1H of Trp-60 (2.7 Å). The resonances belonging to the Ser-1 spin system were then assigned by the conventional procedure. The ^{15}N chemical shift of 33.5 ppm in the HMQC spectrum (data not shown), characteristic of the N-terminal amino acid, allowed the identification of the NH of Ser-1. The pattern of Ala-78 was located due to a cross peak between Hα Ala-78 and the close lying Hη proton of Trp-76 (2.5 Å). We have been able to detect also the Hα proton of Gln-64. Indeed, the NHε protons of Gln-50 are connected with both Hγ protons of Gln-64, which, in turn, show cross peaks with the Hα proton of the same residue. The new assignment of NHε protons of Gln-64 was obtained from the ^{1}H−^{15}N 2D HMQC spectrum.

The analysis of the 2D ^{1}H−^{15}N HMQC spectrum (Figure 3) led to the assignment of the side chain NH protons of some amino acids. We located the NH protons and the guanidine protons for both Arg-28 and Arg-33. Gaillard and co-workers (Gaillard et al., 1992) had assigned the signals at 9.61 and 6.55 ppm to the NH geminal protons of Asn-72 based on dipolar and scalar cross peaks between them. However, we were not able to detect heteronuclear connectivities from these signals to any ^{15}N atom. After having assigned all the NH_2 protons lying in the usual (6.5−8.0 ppm) region to Asn and Gln residues, we remained with one unassigned strong NOESY peak between signals with chemical shifts of 7.10 and 4.40 ppm, unusual for geminal NH protons. However, judging from the characteristic strong intensity as well as from the lineshape of this peak, this possibility was not discarded. In the ^{1}H−^{15}N HMQC spectrum, we detected the alignment of these two resonances with a ^{15}N resonance at 112.4 ppm. Therefore, we could unambiguously assign the latter ^{1}H−^{1}H cross peak as originating from two NH_2 protons. The set of dipolar connectivities observed for these two protons (with both β-protons of Asn-72, γ-protons of Gln-21 and Gln-59, and aromatic protons of Trp-60) indicates that they arise from the NH_2 of Asn-72. Other reasonable candidates for the signals at 9.61 and 6.55 ppm should then be found. On the basis of scalar and dipolar connectivities in both H_2O and D_2O solutions, we can assign these signals as arising from the Hε1 (6.55 ppm) and the OHη (9.61 ppm) protons of Tyr-19. The detection of a strong NOESY cross peak (the volume corresponds to an interatomic distance of 1.9 Å) between the OHη of Tyr-19 and the NH of Asn-72 further confirms the validity of our assignment.

Another discrepancy between the published assignment (Gaillard et al., 1992) and our experiments concerns the assignment of Arg-28. We detect sequence-specific connectivities for this part of the protein chain (Hα Glu-27−

Table 1: ^{15}N and ^{1}H Assignments for the Reduced HiPIP from *Ch. vinosum*

residue	N	NH	Hα	Hβ	others
Ser-1	33.5	7.56[a]	4.18[a]	4.01[a], 3.90[a]	
Ala-2	126.2	8.71[a]	4.04[a]	0.99[a]	
Pro-3			4.23	2.10[2], 1.59[1]	$^{\gamma}$H 1.39[a,2], 1.05[a,1] $^{\delta}$H 2.23[a,2], 1.29[a,1]
Ala-4	111.6	8.91	3.97	1.37	
Asn-5	111.0	8.26	4.74	3.14[1], 2.66[2]	N$^{\delta}$H 7.28[1], 6.60[2]; N$^{\delta}$ 111.7
Ala-6	122.1	6.98	4.58	0.82	
Val-7	125.0	8.46	3.44	1.18	$^{\gamma}$H$_3$ 0.24[1], 0.00[2]
Ala-8	134.4	8.83	4.53	1.57	
Ala-9	121.5	8.57	4.14	1.47	
Asp-10	111.7	8.05	4.63	2.98, 2.58	
Asp-11	123.0	7.51	4.52	3.04[2], 2.90[1]	
Ala-12	130.3	9.12	4.08	1.57	
Thr-13	125.1	8.81	4.15	4.43	$^{\gamma}$H$_3$ 1.14
Ala-14	126.3	8.13	4.38	1.78	
Ile-15	118.6	8.55	3.91	1.87	$^{\gamma}$H 1.20[1]; $^{\gamma}$H$_3$ 1.00; $^{\delta}$H$_3$ 0.88
Ala-16	124.2	7.95	4.20	1.64	
Leu-17	116.5	8.33	4.44	2.38, 2.11	$^{\gamma}$H 2.10; $^{\delta}$H$_3$ 1.02[a,1], 1.31[a,2]
Lys-18	116.7	7.92	4.04	2.08[2], 2.36[1]	$^{\gamma}$H 1.41[1]; $^{\delta}$H 1.79; $^{\epsilon}$H 3.10
Tyr-19	117.9	8.49	5.13	3.04[1], 2.62[2]	$^{\delta}$H 6.51; $^{\epsilon}$H 6.55; O$^{\eta}$H 9.61[a]
Asn-20	128.9	8.63	4.12	2.27, 2.25[a]	N$^{\delta}$H 7.72, 6.66; N$^{\delta}$ 114.3
Gln-21	124.7	8.32	3.98	2.45[1], 2.17[2]	$^{\gamma}$H 2.08, 2.21; N$^{\epsilon}$H 7.14[1], 7.90[2]; N$^{\epsilon}$ 112.8
Asp-22	118.1	8.75	4.61	2.74, 2.36	
Ala-23	127.3	9.80	3.86	1.62	
Thr-24	137.8[c]	8.40	4.09	4.23	$^{\gamma}$H$_3$ 1.27
Lys-25	121.7	7.61	4.46	2.10, 1.66	$^{\gamma}$H 1.36; $^{\delta}$H 1.21; $^{\epsilon}$H 2.94
Ser-26	115.9	7.19	4.56	5.76[2], 3.76[1]	
Glu-27	119.5	8.97	4.45	1.96	$^{\gamma}$H 2.43, 2.47
Arg-28	118.2	6.61	3.10[d]	0.55[d], −0.73[d]	$^{\gamma}$H 0.56, 0.87; $^{\delta}$H 2.04[d,1], 3.02[d,2] N$^{\epsilon}$H 6.78[a]; N$^{\eta}$H 6.26[a], 6.94[a], 6.85[a]; N$^{\epsilon}$ 86.2; N$^{\eta}$ 68.9, 78.2
Val-29	122.7	8.33	3.37	1.83	$^{\gamma}$H$_3$ 0.83[1], 0.82[a,2]
Ala-30	122.9	7.60	4.05	1.41	
Ala-31	119.3	8.24	4.10	1.75	
Ala-32	117.8	7.48	3.77	1.33	
Arg-33	122.6	10.64	4.50	2.54[2], 2.29[1]	$^{\gamma}$H 2.12[2]; $^{\delta}$H 3.57[a,1], 3.45[a,2]; N$^{\epsilon}$H 7.80[a]; N$^{\eta}$H 6.02[a], 7.04[a], N$^{\epsilon}$ 85.1; N$^{\eta}$ 70.1, 74.0
Pro-34			4.63	2.08[1], 2.35[a,2]	$^{\delta}$H 3.84, 3.94
Gly-35	109.0	8.74	4.76, 3.62		
Leu-36	118.9	7.33	4.85	1.48, 1.45[a]	$^{\gamma}$H 2.00; $^{\delta}$H$_3$ 1.14[a], 1.10[a]
Pro-37			4.63	2.46, 1.60	$^{\gamma}$H 1.97[a,1], 2.10[2]; δH 3.84[2], 3.47[1]
Pro-38			3.69	1.95[a], 2.27[a]	$^{\gamma}$H 1.53[a], 1.95[a]; $^{\delta}$H 3.93[1], 3.75[2]
Glu-39	112.2	8.95	3.63	1.97[2], 1.86[1]	$^{\gamma}$H 2.06, 2.19
Glu-40	116.4	7.83	4.34	2.36, 1.77	$^{\gamma}$H 2.16, 2.21
Gln-41	121.9	7.41	3.76	1.74[1], 0.23[2]	$^{\gamma}$H 1.86[2], −0.10[1]; N$^{\epsilon}$H 6.62, 6.78; N$^{\epsilon}$ 110.1
His-42	117.4	9.02	6.23	4.01[1], 3.93[2]	$^{\delta}$H 7.71[a]; $^{\epsilon}$H 8.85[a]
Cys-43	128.9[c]	10.14	4.71[c]	16.20[b,2], 7.19[b,1]	
Ala-44	113.7	7.82	4.45	1.25	
Asn-45	113.0	8.15	5.10	3.76[1], 2.85[2]	N$^{\delta}$H 7.82, 7.22; N$^{\delta}$ 110.3
Cys-46	127.0	8.28	4.30[b]	10.05[b,1], 11.01[b,2]	
Gln-47	136.1	8.88	4.04	0.70[1]	$^{\gamma}$H 1.70[1], 1.61[2]; N$^{\epsilon}$H 6.71[a,2], 6.82[a,1]; N$^{\epsilon}$ 111.9
Phe-48	115.8[c]	7.80[c]	3.57[c]	2.29, 3.45	$^{\delta}$H 7.04; $^{\epsilon}$H 6.72
Met-49					$^{\gamma}$H 2.74[a], 2.94[a]; $^{\epsilon}$H 2.20[a]
Gln-50	128.5	8.51	4.67	1.85[2], 1.64[1]	$^{\gamma}$H 2.25[1], 2.19[2]; N$^{\epsilon}$H 6.83[a,2], 7.60[a,1]; N$^{\epsilon}$ 117.7
Ala-51	129.5	9.12	4.09	1.45	
Asp-52	114.9	8.43	4.75	2.74, 2.67	
Ala-53	122.2	6.61	4.15	1.30	
Ala-54	126.3	8.66	4.14	1.32	
Gly-55	111.8	8.88	4.14[2], 3.62[1]		
Ala-56	120.7	7.68	4.46	1.62	
Thr-57	115.6	9.80	4.85	4.89	$^{\gamma}$H$_3$ 1.40
Asp-58	118.9	8.74	4.41	2.75[1], 2.71[2]	
Glu-59	117.0	8.06	4.34	1.79, 1.65	$^{\gamma}$H 2.01
Trp-60	121.1	7.24	5.63	3.14[1], 3.03[2]	$^{\delta}$H 7.27; $^{\epsilon}$H 7.22[a], N$^{\epsilon}$H 10.18; $^{\zeta}$H 7.44, 7.53; $^{\eta}$H 7.27; N$^{\epsilon}$ 129.3
Lys-61	120.1	8.94	4.89	1.97[2], 1.62[1]	$^{\gamma}$H 1.68[2], 2.20[a,1]; $^{\delta}$H 2.25[a]; $^{\epsilon}$H 3.04[a,2], 2.93[a,1]; N$^{\zeta}$H 7.58[a]
Gly-62	105.4	7.85	4.46[1], 3.63[2]		
Cys-63	129.7	9.25	3.70[b]	5.27[b,1], 15.78[b,2]	
Gln-64	131.3[c]	8.82[c]	4.07[a]	2.30[c]	$^{\gamma}$H 2.50[a], 2.35[a]; N$^{\epsilon}$H 7.59[a], 6.89[a]; N$^{\epsilon}$ 113.2
Leu-65	127.8[c]	7.53[c]			
Phe-66	117.8	7.27	4.89	2.74[2], 2.32[1]	$^{\delta}$H 7.43; $^{\zeta}$H 7.93
Pro-67			4.42	2.30, 1.96	$^{\gamma}$H 2.10; $^{\delta}$H 3.58[1], 3.36[2]
Gly-68	117.0	9.04	4.17[2], 3.80[1]		
Lys-69	120.1	7.67	5.12	1.90, 1.54	$^{\gamma}$H 1.34; $^{\epsilon}$H 3.00
Leu-70	120.1	7.78	5.07	1.62, 1.13	$^{\gamma}$H 1.28; $^{\delta}$H$_3$ 0.70[1], 0.55[2]
Ile-71	115.4	9.56	4.57		$^{\delta}$H$_3$ 0.997[a]

Table 1: (Continued)

residue	N	NH	Hα	Hβ	others
Asn-72	121.3	7.44	4.54	2.71^{2}, 1.73^{1}	$N^{\delta}H$ $7.10^{d,1}$, $4.40^{d,2}$; N^{δ} 112.4
Val-73	122.3	8.49	3.84	2.17	$^{\gamma}H_3$ 1.00, 0.97[a]
Asn-74	118.1	9.07	5.18	3.16^{1}, 2.75^{2}	$N^{\delta}H$ 7.94^{1}, 7.16^{2}; N^{δ} 117.0
Gly-75	112.7	8.62	4.32^{1}, 3.95^{2}		
Trp-76	108.5	8.32	3.95	3.01^{1}, 1.61[c]	$^{\delta}H$ 6.84; $^{\epsilon}H$ 7.23, $N^{\epsilon}H$ 9.96; $^{\zeta}H$ 6.92, 7.57; $^{\eta}H$ 6.73; N^{ϵ} 129.3
Cys-77	121.4[c]	7.81[c]	8.23[b]	$12.60^{b,1}$, $7.79^{b,2}$	
Ala-78	114.3[c]	7.71[a]	3.64[a]	1.30[a]	
Ser-79	118.3[c]	8.39[c]			
Trp-80	112.2	7.57	6.61[c]	3.10	$^{\delta}H$ 9.02; $^{\epsilon}H$ 7.22[a], $N^{\epsilon}H$ 10.56; $^{\zeta}H$ 7.17, 7.38; $^{\eta}H$ 6.95; N^{ϵ} 133.0
Thr-81	112.0	7.05	3.94	3.11	$^{\gamma}H_3$ 0.79
Leu-82	129.9	8.16	3.61	1.63^{2}, 1.43^{1}	$^{\gamma}H$ 1.31[a]; $^{\delta}H_3$ 1.13[a], 0.99[a]
Lys-83	127.9	8.25	3.97	1.62, 1.31	
Ala-84	131.6	8.51	4.35	1.35	
Gly-85	116.3	8.08	3.77		

[a] The chemical shifts of ^{15}N resonances measured at 300 K. New assignments achieved following the conventional methods. [b] Assignments taken from Bertini et al. (1992). [c] These assignments are tentative. [d] Different assignments compared to Gaillard et al. (1992). [1,2] refer to stereospecifically assigned protons. The other assignments agree with the previous work of Gaillard et al. (1992).

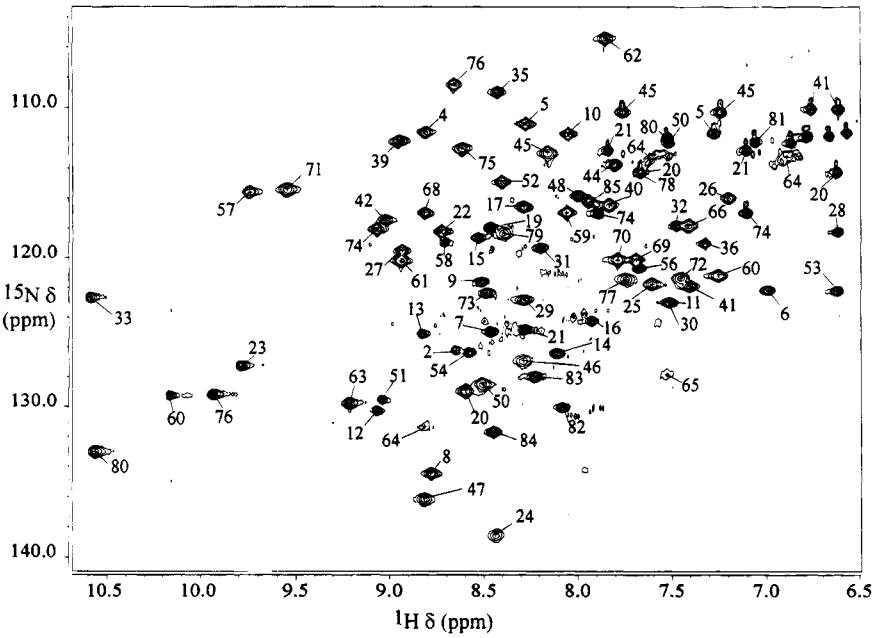

FIGURE 3: 300 K $^{1}H-^{15}N$ HMQC spectrum of the reduced HiPIP from *Ch. vinosum* in 0.05 M phosphate buffer in H_2O solution. The $N^{1}H-^{15}N$ cross peaks are labeled according to their assignment in the amino acid sequence.

NH Arg-28 and Hα Arg-28 and NH Val-29). By analyzing the intensity of the scalar intraresidue peaks as well as of both the intraresidual and interresidual dipolar connectivities, we assign the Hβ protons of Arg-28 at 0.55 and −0.73 ppm, and the Hδ protons at 3.02 and 2.04 ppm. In other words, we reverse the assignment for the β and δ protons.

In the present work 63 additional proton resonances have been assigned [denoted by superscript *a* (footnote) in Table 1] and 7 resonances needed revision [denoted by superscript *d* (footnote) in Table 1].

Summarizing, the NMR experiments allowed us to perform an almost complete assignment of ^{1}H and ^{15}N resonances in the spectra of reduced HiPIP from *Ch. vinosum*. ^{1}H spin systems belonging to 83 amino acids out of 85 were completely or partially assigned by both homonuclear and heteronuclear experiments. About 85% of the ^{1}H resonances were assigned by appropriate experiments. The missing assignments may be due to either degeneracy of resonances or paramagnetic effects. Only the amino acids Leu-65 and Ser-79 remain completely unassigned. As it appears from inspection of the X-ray structure (Carter et al., 1974) and previous NMR investigations on HiPIP from *Ch. vinosum*, these amino acids are very close to the iron−sulfur cluster and are expected to experience a strong enhancement of the nuclear relaxation rates. Moreover, the NH protons of these two amino acids are involved in hydrogen bonding with the sulfur atoms of Cys-63 and Cys-77, respectively (Backes et al., 1991). A tentative assignment of the NH proton of Ser-79 (Table 1) was obtained from the analysis of a 1D NOE difference spectrum obtained by saturation of Hβ1 of Cys-77, and of a $^{1}H-^{15}N$ HMQC spectrum (Figure 3). In the 1D spectrum there is a NOE at 8.39 ppm, while in the $^{1}H-^{15}N$ HMQC spectrum a broad unassigned cross peak with low intensity at the same proton chemical shift is observed. Also, a tentative assignment of NH and ^{15}N of Leu-65 can be achieved from the $^{1}H-^{15}N$ HMQC spectrum (Table 1 and Figure 3).

By analyzing the ^{15}N NMR backbone assignment, we note that, out of 85 amino acids, 71 ^{15}N signals were assigned by conventional methods, 8 were tentatively assigned, and 6 remained unassigned. Out of these 6 amino acids, 5 are prolines which do not give rise to connectivities in the $^{1}H-$

Table 2: Number of Constraints Constituting the Upper Distance Limits File[a]

class no.		
1	intraresidue (except NH, Hα, Hβ)	344 (59 methyl peaks)
2	sequential and intraresidue NH, Hα, Hβ	459 (62 methyl peaks)
3	medium range	105 (14 methyl peaks)
4	long range backbone	27
5	long range	554 (195 methyl peaks)

[a] They are divided according to the classes described in the Experimental Procedures.

^{15}N HMQC spectrum. It should be emphasized that in the ^{1}H−^{15}N HMQC spectrum strong cross peaks even for nuclei close to the cluster are observed, which in ^{1}H TOCSY spectra give rise to weak and broad connectivities. This is due to the large value of the coupling constant between hydrogen and nitrogen ($J_{N\text{-}H} \sim 100$ Hz) and also to the lower sensitivity of ^{15}N to paramagnetic effects. Additionally, we have also assigned ^{15}N resonances belonging to the amino acid side chains using either the correlation patterns, as in the case of Gln's and Asn's, or the characteristic chemical shift values for Trp and Arg side chain nitrogen resonances.

The ^{1}H and ^{15}N NMR resonance assignment of reduced HiPIP from *Ch. vinosum* is reported in Table 1.

Structure Determination. From the NOESY spectra it was possible to detect 1461 cross peaks. The more intense integral of each two symmetric peaks was used for the conversion into upper distance limits. The amount of constraints for each calibration class is reported in Table 2. The calibration was in best agreement for the volume-to-distance correlation when assuming the intensities to be inversely proportional to the power of 4 of the corresponding upper distance limits in the case of classes 1 and 5, to the power of 6 for class 2, and to the power of 5 for classes 3 and 4 (Güntert et al., 1991). An additional 28 connectivities were determined from 1D NOE difference spectra recorded upon saturation of the hyperfine shifted signals. The intensities of these connectivities were measured and properly scaled to those of the other already measured 1461 connectivities. The constraints between the atoms of the cluster are imposed as described in the Experimental Section.

Out of the 1489 upper distance limits, the DG program found 342 to be irrelevant for the determination of the structure. This means that either no conformation of the peptide chain could give rise to violations of these constraints, or that they belong to fixed proton−proton distances. Thus, 1147 distance constraints were used in all the DG, as well as in the following REM and RMD, calculations. With an average of 17.2 NOE's per residue as input for the DG calculation and 13.4 accepted experimental constraints per residue, we provide thereby the prerequisite for the calculation of a very well-resolved three-dimensional structure. For the residues without a stereospecific assignment, pseudo atoms were used. The total number of NOE's per residue is shown in Figure 4. The filled bars give the number of NOE's per residue that were found to be meaningful by the DG program for the structure calculation.

In the beginning of the structure calculation, 4 times 50 random structures were generated in four subsequent DIANA runs. The angle constraints obtained using the REDAC strategy after each run were used as input for the following.

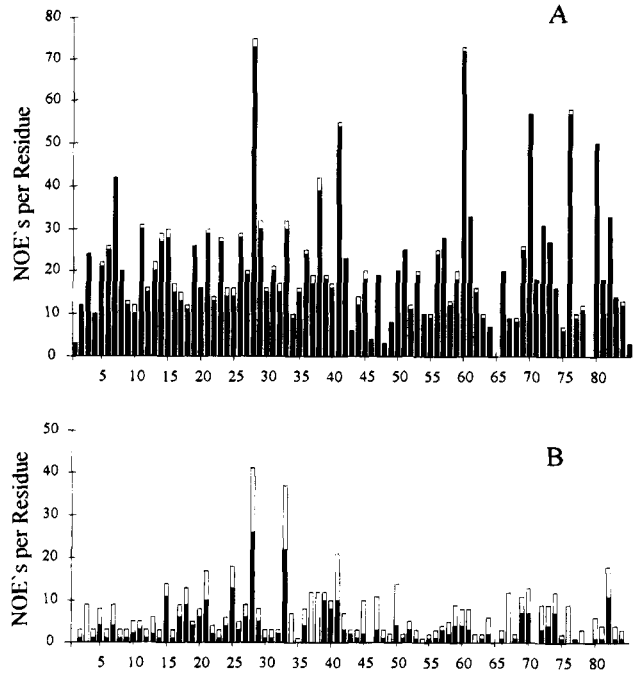

FIGURE 4: Number of inter- (A) and intraresidual (B) NOE's per residue identified in the NMR spectra of the reduced HiPIP from *Ch. vinosum*. The total height of each column represents the amount of NOE's which were used as input for the DG calculations; the filled part of the column indicates the numbers of NOE's which were found to be meaningful by the DG calculations.

After these initial runs, 314 angle constraints were obtained, as well as 12 stereospecific assignments by using GLOMSA. Another 231 random structures were calculated, again using the previously obtained angle constraints except in the very last run. Out of these 231, only the 27 structures with a target function lower than 2.3 Å^2 were used to produce new angle constraints. After this calculation, by using GLOMSA, the number of stereospecific assignments was increased to 31. Again, 315 random structures were generated, out of which 34 had a target function lower than 2.5 Å^2 and were used to produce angle constraints. Then, out of 238 generated random structures, 50 structures with a target function lower than 2.2 Å^2 were used to improve the angle constraints, and finally, out of 99 newly calculated random structures, 22 structures with a target function lower 2.0 Å^2 were taken for the production of more refined angle constraints. So, in total more than 1000 random structures have been generated before the final four REDAC cycles were started. Another GLOMSA run now allowed the obtainment of 37 stereospecific assignments. Out of the previously calculated 99 random structures, the 50 with the smallest target function were then used for 4 subsequent REDAC cycles. By using GLOMSA again, the final number of 57 stereospecific assignments out of 126 was reached. The last REDAC cycle was performed in the presence of these stereospecific assignments but without angle constraints. The stereospecifically assigned protons are reported in Table 1.

The 15 structures generated by DG calculations which have the lowest target function (<0.9 Å^2) have no consistent violations, have no residual violation exceeding 0.3 Å, and experience mean global RMSD values of 0.73 ± 0.12 and 1.13 ± 0.12 Å for the backbone and the heavy atoms, respectively. The distribution of the RMSD values per

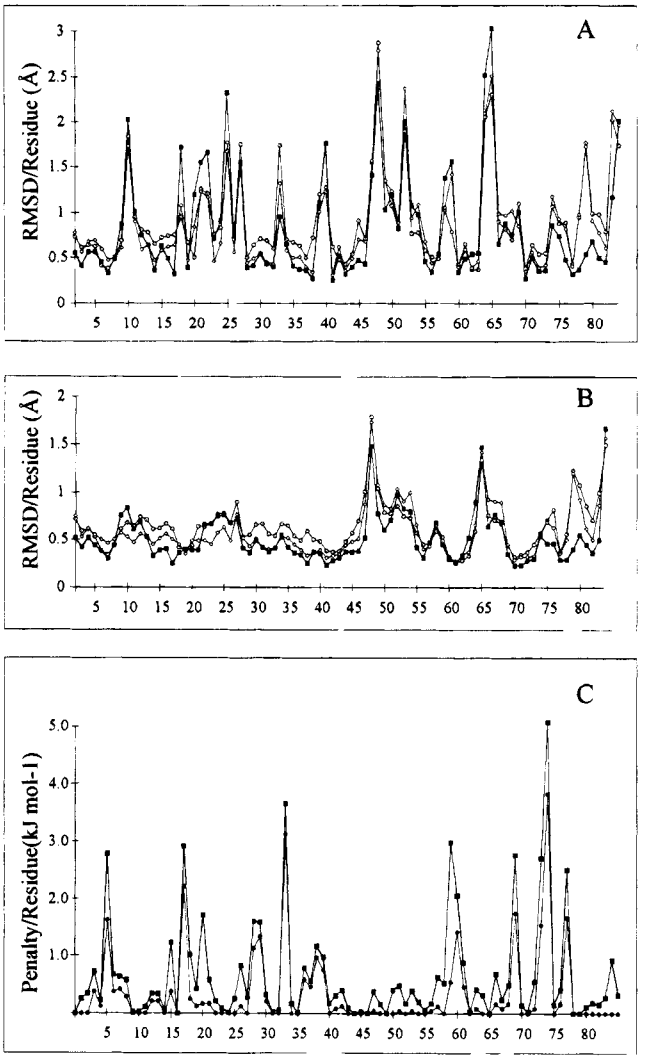

FIGURE 5: Diagrams of the RMSD per residue for the 15 accepted structures of the reduced HiPIP from *Ch. vinosum* after DG (○), REM (◇) and RMD (■) calculations. RMSD values for the heavy atoms and the backbone are presented in panels A and B, respectively. Panel C shows the average (■) as well as the minimum (●) value of the penalty function per residue reported by SANDER after RMD.

residue for both the backbone and the heavy atoms is shown in Figure 5. The DG family was then submitted to REM calculations. The mean global RMSD values for residues 2−82 in the REM family were 0.69 ± 0.11 and 1.14 ± 0.11 Å for the backbone and side chain atoms, respectively (Figure 5). The REM family was finally subjected to RMD calculations, time averaging, and, again, energy minimization, thus obtaining the final RMD family. The distribution of the RMSD per residue for the RMD family is shown in Figure 5. Going from the DG to the RMD family, the average RMSD values for the backbone decreases from 0.73 ± 0.12 to 0.62 ± 0.12 Å, while for the heavy atoms we detect an increase from 1.13 ± 0.12 to 1.19 ± 0.18 Å. The decrease for the backbone is sizable for the following regions: Ala-2−Ala-8, Thr-13−Tyr-19, and Ser-26−Gln-47 as well as for Val-73−Leu-82. On the other hand, an improvement for residues 48−66 is less prominent or even absent.

Average RMSD values for backbone (BB) and for all heavy atoms for the DG, REM, and RMD families together with the average energy values for the last two families are reported in Table 3.

Stereo drawings of the backbone for the DG, REM, and RMD families are shown in Figure 6. It becomes clear that the conformation of the RMD family is essentially the same as those of the DG and REM families; the decrease in RMSD is also evident from the figure.

Evaluation of the Structure Quality. By inspection of the backbone RMSD values (Figure 5B), clearly two different regions can be located. Part of the protein on the N-terminal (Ala-2−Cys-46) region shows, already after the DG calculations, an almost homogeneous distribution of the RMSD values together with a remarkable small average value of 0.46 Å. In contrast to this, two residues with RMSD values larger than 1.0 Å (Phe-48 and Leu-65) are found later in the sequence; likewise, significantly high average RMSD's (0.69 Å) are observed in the C-terminal half of the sequence (Phe-48−Gly-68). The C-terminal region of the protein, in the range from Phe-48 to Asn-72, is characterized by three stranded, twisted, antiparallel β-sheets (Carter et al., 1974). The two maxima of the RMSD for the whole protein are found in this region for residues Phe-48 and Leu-65. They are poorly defined in terms of experimental NOE constraints. Neither REM nor RMD calculations were able to provide a significant decrease of the disorder of the structures in this area.

However, a lack of experimental constraints alone cannot give rise to the overall poor definition in this part of the structure. This becomes obvious when one compares the average number of 20.3 NOE's per residue in the segment spanning from Phe-48 to Asn-72 against the 20.7 NOE's per residue for the region Ala-2 to Cys-46. One possible explanation for the different quality of the structure in the two protein segments is that not only is the number of the experimental constraints important but rather their quality (i.e., origin). For instance, our structural data, in agreement with the X-ray structure, show that the amino acid region from Gln-50 to Glu-59 forms an external loop. The average number of interresidue NOE constraints for this region of the protein is 17.6. On the other hand, almost all these

Table 3: Average RMSD Values for Backbone (BB) and for All Heavy Atoms (HA) for the DG, REM, and RMD Families[a]

		DG	REM	RMD	RMDw
RMSD for residues Ala-2−Leu-82 (Å)	BB	0.73 ± 0.12	0.69 ± 0.11	0.62 ± 0.12	
	HA	1.13 ± 0.12	1.14 ± 0.11	1.19 ± 0.18	
target function (Å^2)		0.56			
target function (kJ mol^{-1})			55.98	57.07	44.27
total energy (kJ mol^{-1})			−5701.0	−7259.0	−6648.0
deviation from ideal bond distances (Å)					0.007
deviation from ideal bond angles (deg)					1.55

[a] For the last two families the average target functions and energy values are also reported. RMDw refers to the restrained energy minimized time average of the long RMD run in water.

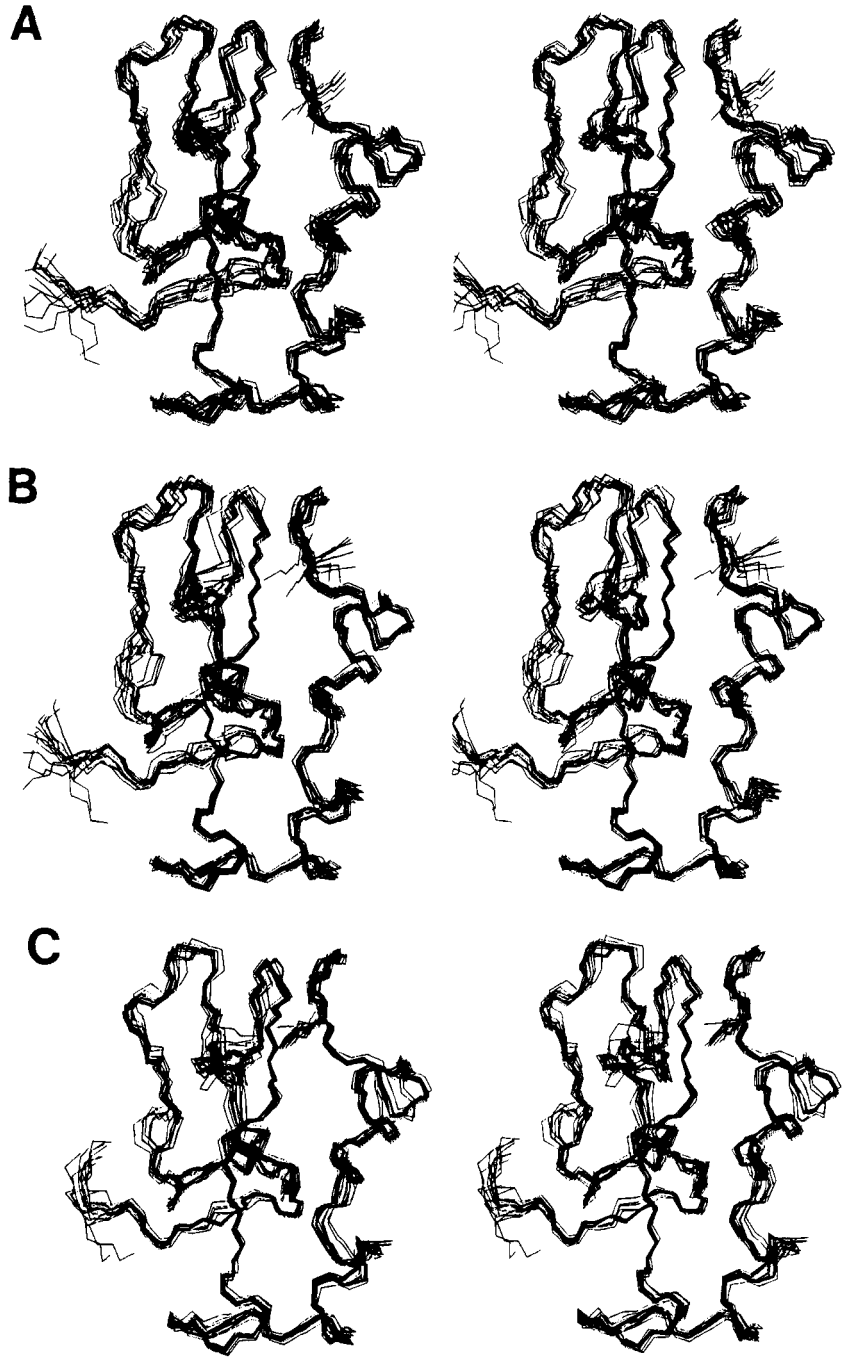

FIGURE 6: Stereo drawings of the 15 accepted structures of the reduced HiPIP from *Ch. vinosum* obtained after DG (A), REM (B), and RMD (C) calculations.

constraints are among the protons of the residues belonging to the loop itself. Therefore, while this part is well-defined in terms of local NOE constraints, they are not sufficient to prevent disorder with respect to other parts of the structure.

Beside the β-sheet region, RMD calculations are able to largely decrease the RMSD values compared to the DG family for residues Val-73−Leu-82. Whereas after the DG calculations the RMSD in this region is almost as high as for Phe-48 to Gly-68, after RMD calculations, it becomes as small as it is for the N-terminal part. For example, at the REM level, Ser-79 had the maximal local RMSD value in this part of the protein. After the RMD calculations the RMSD value for Ser-79 decreases from 1.26 to 0.40 Å, becoming even smaller than the average RMSD value for the whole backbone. This different behavior of the two regions during the structure calculation can be explained as due to an involvement of residues Phe-48−Gly-68 in an antiparallel β-sheet in one case, while the second region (Val-73−Leu-82), in analogy with the N-terminal half of the protein (Ala-2−Cys-46), does not form a similar secondary structure.

As can be seen from Figure 5A, although for most residues the local RMSD values after RMD calculations become smaller in comparison with both DG and REM calculations, the increase of the local RMSD values for the heavy atoms of some residues causes a slight increase of the average RMSD value for the RMD family with respect to the DG and REM families.

The division of the protein into two regions (N-terminal and C-terminal) according to the RMSD values, which was

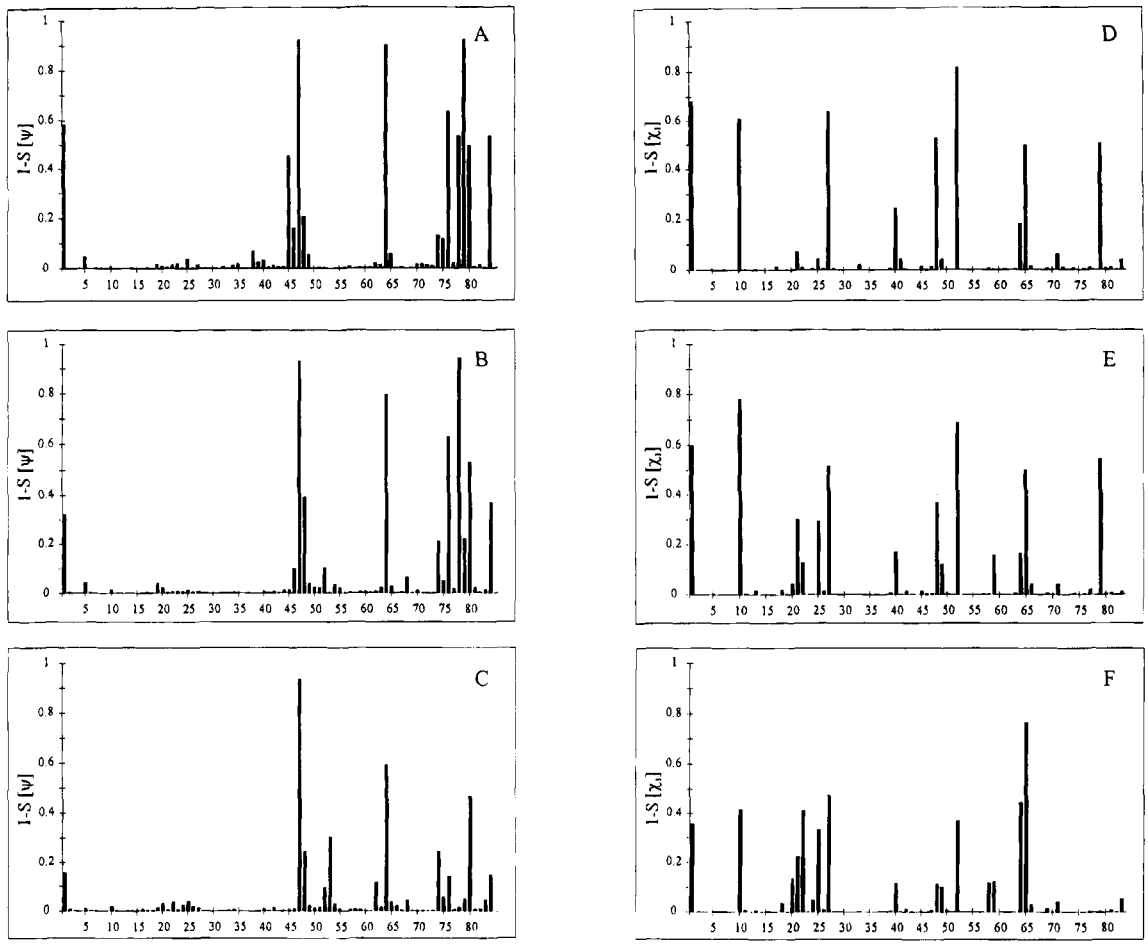

FIGURE 7: Diagrams of the (1 − S) parameter for the 15 accepted structures of the reduced HiPIP from *Ch. vinosum* for the DG, REM, and RMD families. Histograms of the (1 − S) parameter values of the ψ angles and of the χ_1 angles in the DG, REM, and RMD families are presented in panels A, B, C, D, E, and F, respectively.

clearly observed for backbone atoms, is still observed for heavy atoms, although it becomes less evident. The average RMSD values for the two regions Ala-2−Cys-46 and Phe-48−Gly-68 are 0.75 and 1.12 Å, respectively. The most noticeable improvement of the RMSD values for heavy atoms after RMD was observed in the region Val-73−Leu-82.

We can conclude that, apart from some local differences, the overall behavior of RMSD for the heavy atoms essentially parallels that of the backbone atoms. Therefore, the explanation of the behavior of the local RMSD values discussed for the backbone atoms is valid also for heavy atoms. Later, we will return to the discussion of the possible reasons for the increase of some local heavy atoms RMSD values.

In Figure 5C, the averaged and minimal values of the distance penalty function for the RMD calculations are plotted against the residue number. These plots give an overview of the NOE violations induced by the energy requirements of the RMD calculations. These violations are absent in the DG structures. We would like first to emphasize that the overall penalty function values are relatively small. Comparing these data to the values of RMSD for the RMD family, we can divide them into three classes. When a low value of the RMSD is present in correspondence with a low value of the penalty function, it means that the residue is well-defined in terms of experimental NOE's, and the latter are in agreement with energy requirements. The case when a low penalty function is in correspondence with a high RMSD value implies that either the residue is poorly defined by NOE constraints or the latter are not restrictive enough and RMD calculations find several conformations satisfying both distance and energy requirements. When a high penalty function is present with a low RMSD value, it means that RMD calculations provide a unique conformation with low energy, but NOE constraints are violated. The unfavorable case where a high RMSD corresponds to a high penalty value was not met in our investigation.

In order to give further elements to estimate the quality of the structures, we can also consider the order parameter S (Hyberts et al., 1992). This parameter ranges from 0 (randomly oriented structures) to 1 (perfectly superimposable structures). Here we report $1 - S$, which is thus a measure of the extent of disorder for the structures within one family. This parameter, evaluated from the deviations of the ψ and χ_1 dihedral angles of the structures, is reported in Figure 7, ψ and χ_1 being the $N_i-C_i-CO_i-N_{i+1}$ and the $N_i-C_i-C_i-X_i$ dihedral angles, respectively. In the case of the angles for the backbone (Figure 7A−C) we observe that the RMD family is more ordered with respect to both the DG and REM families. Sizable decreases of the 1 − S values in the RMD family are observed for the following amino acids: Phe-48, Gln-64, and Trp-76−Ala-84. Therefore, these regions become much better defined after RMD and are characterized by a low penalty function values. On the contrary, Asp-52 and Ala-53 become more disordered after RMD calculations.

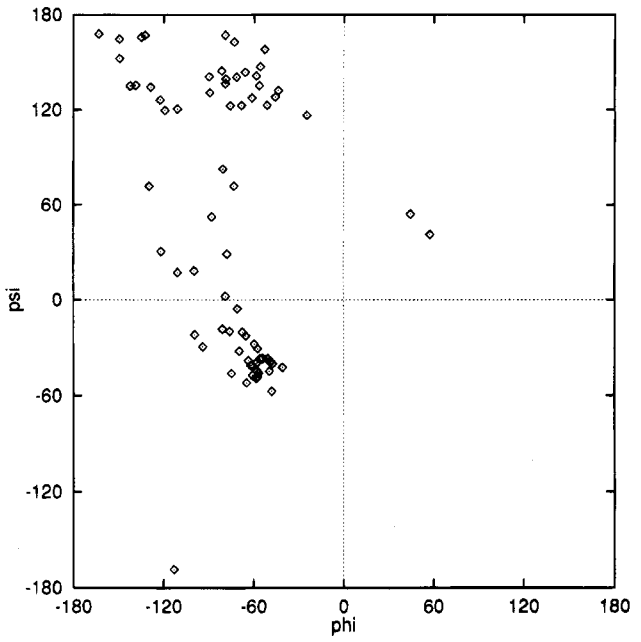

FIGURE 8: Ramachandran plot of the ψ and ϕ dihedral angles for the RMDw structure (without prolines and glycines) of the reduced HiPIP from *Ch. vinosum*.

Since, the penalty functions are also small for these residues, we conclude that for these amino acids all conformations produced as a result of RMD calculations are able to satisfy both distance and energy requirements.

The 1 − S parameters evaluated for the χ_1 dihedral angles for the three levels of refinement are shown in Figure 7D−F. In the N-terminal part of the protein we observe that in the RMD family, in comparison with the DG and REM families, the 1 − S parameters substantially decrease for a few residues, like Ser-1 and Asp-10 (that experience low penalty function values), while they become larger, especially with respect to DG calculations, for the side chains of Asn-20, Gln-21, Asp-22, and Lys-25 (again showing low penalty functions). This feature coincides with the increase of local RMSD values for the heavy atoms of the same residues in the RMD family, which was already mentioned above. These residues are well determined by NOE connectivities (Figure 4). Therefore, the observed disorder can be explained as mobility of the charged side chains placed on the surface of the protein and thereby gaining additional rotational freedom. As can be seen from Figure 7, the remaining side chains of the N-terminal part of the protein are well ordered. Peak values of the 1 − S parameter for the C-terminal part of the protein are observed for the side chains of Asp-52, Gln-64, and Leu-65. The charged Asp-52 residue belongs to the external loop (see before), and therefore, its side chain can be arranged in different conformations. The poor definition of Gln-64 and Leu-65 by NOE constraints is a reason for the high 1 − S values for these amino acids. Note the significant decrease of disorder for the side chains of Phe-48 and Ser-79, achieved after RMD calculations. All the above residues display low values of penalty functions. A high penalty function is displayed by residues Asn-5, Leu-17, Arg-33, Glu-59, Lys-69, Asn-74, and Cys-77. These are clear examples of conflict between experimental data and energetic requirements. The values of the individual NOE's involving these residues were further checked and confirmed.

A final comment is due to the side chains of the aromatic residues that are placed relatively close to the iron−sulfur cluster. These residues have been proposed to play a role in the process of electron transfer (Carter et al., 1974; Breiter et al., 1991). The HiPIP from *Ch. vinosum* has seven aromatic amino acids: Tyr-19, His-42, Phe-48, Trp-60, Phe-66, Trp-76, and Trp-80. All the aromatic protons of Tyr-19, His-42 (except its NH), and all tryptophans were assigned. These amino acids are well-defined in terms of NOE's per residue (Figure 4) and, as a result, have a rather small RMSD value (0.48 Å, Figure 5A) for the heavy atoms in the RMD family as well as 1−S parameters for the χ_1 dihedral angles close to zero at every degree of refinement. The aromatic rings of the two phenylalanines are pointing toward the [4Fe-4S] core, and some proton signals are significantly broadened; as an example one proton per residue escaped detection. Nevertheless, in the case of the aromatic ring of Phe-66 we have found several connectivities (Figure 4), which allowed us to obtain, after RMD calculations, an RMSD value for heavy atoms of 0.65 Å and a relatively small 1 − S parameter (Figure 7). The small number of dipolar connectivities from the aromatic protons of Phe-48 result in a high RMSD value for the heavy atoms of this residue (2.43 Å). In this case, however, RMD calculations sizably decrease the corresponding 1 − S parameter (Figure 7F).

Comparison of the X-ray and Solution Structures. The mean structure of the RMD family was subjected to more

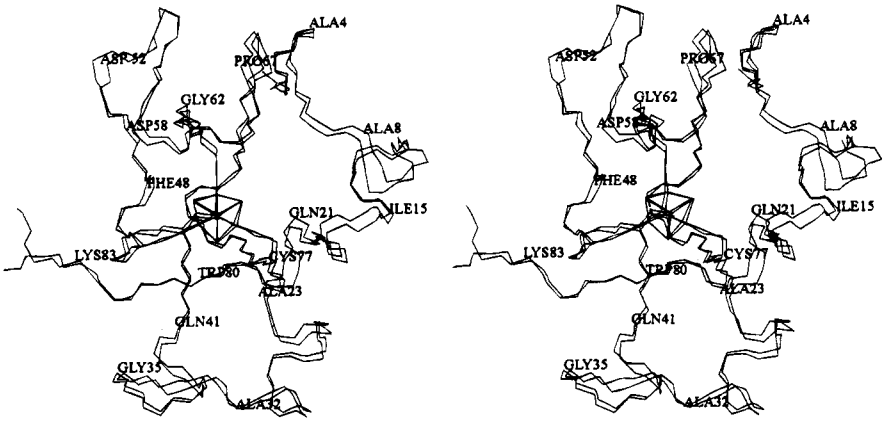

FIGURE 9: Stereo drawings of the backbone atoms of the average RMDw structure of the reduced HiPIP from *Ch. vinosum* with the X-ray structure of the oxidized species.

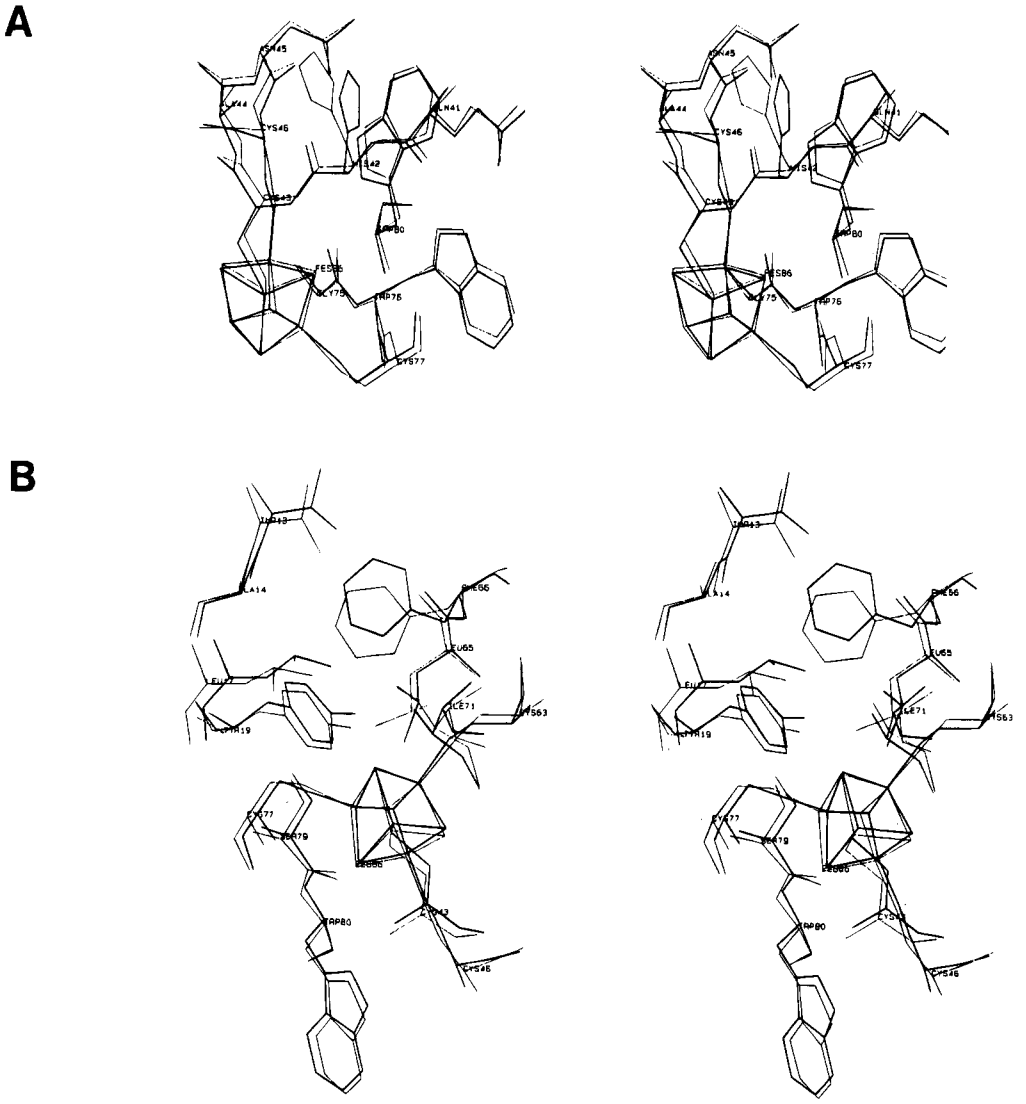

FIGURE 10: Stereo drawings of selected regions of the structures showing the orientations of His-42 (A) and Phe-66 (B) side chains in the RMDw (thick line) and X-ray (thin line) structures of the HiPIP from *Ch. vinosum*.

extended RMD calculations in water to take into account interactions of water molecules with the protein. The structure obtained in this way (RMDw) was compared to the RMD family. The average RMSD values of the RMDw structure with respect to the RMD structures are 0.62 ± 0.12 and 1.19 ± 0.18 Å for the backbone and heavy atoms, respectively. These values are identical to those within the RMD family. The similarity between the refined RMDw structure and the structures of the RMD family allows us to consider the former as a further refinement of the RMD family.

The values of energy and target function for the RMDw structure are reported in Table 3, together with the root mean square deviation of the bond distances (0.007 Å) and of the bond angles (1.55°) from ideal values.

Figure 8 shows the Ramachandran plot, which reports the values of the ψ and ϕ dihedral angles of the amino acids constituting the protein, the ϕ angle being the $CO_{i-1}-N_i-C_i-CO_i$ backbone dihedral angle. As can be seen, the dihedral angles ψ and ϕ fall within the low energy regions, thus witnessing the good quality of the structure.

We are now going to compare the solution structure of the reduced protein with the X-ray structure of the oxidized species. Indeed, the X-ray structure of the reduced protein is also available (C. W. J. Carter, unpublished). However, whereas the solid state structures appear very similar to one another, the available refinement of the reduced protein is poorer.

The comparison of the backbones of the RMDw and X-ray structures (Figure 9) shows that the overall agreement between the two structures is very good. For large parts of the chain, the two structures are almost superimposable. The average RMSD between the two structures is 0.61 Å for the backbone atoms, while it is 1.16 Å for the heavy atoms. These values are almost the same as the RMSD values within the structures of the RMD family.

We also note that elements of the secondary structure present in the X-ray structure are maintained in solution. From the NMR data we observe two regions of α-helices spanning from Ala-12 to Ala-16 and from Arg-28 to Ala-32. The initial observation of slowly exchangeable amide protons, together with strong $C\alpha H(i)-NH(i+1)$ sequential connectivities in the 50−74 region and long range dipolar connectivities (for instance, between protons belonging to spin systems of Lys-61 and Ile-71, Gln-50 and Gln-64, Trp-60 and Leu-70, etc.), already suggested that the secondary

structure in this part of the protein is a twisted antiparallel β-sheet. This was then confirmed by the final RMDw solution structure.

Deviations between the solution and X-ray structures are detected in the regions Ala-8−Asp-11, Gln-21−Asp-22, Leu-36, Glu-39, His-42, Met-49−Gln-50, Asp-52−Gly-55, Phe-66, and Val-73. The regions 8−11, 21−22, and 52−55 are characterized by relatively high RMSD and 1 − S values for the side chains. On the other hand, the other residues showing deviations from the X-ray structure are rather well-defined in terms of NOE constraints (Figure 4) and, in contrast to the previous three regions, have small RMSD and 1 − S parameters (Figures 5 and 7). This may suggest that in the latter cases we detect real differences between the solid state and solution structures. In particular, we would like to comment on the deviations of the aromatic rings of His-42 and Phe-66 (Figure 10). These residues are placed in the vicinity of the [4Fe-4S] cluster and can be involved in electron transfer. As we are comparing the solution structure of the reduced protein with the solid state structure of the oxidized species, we may suggest that the position of the side chains of these amino acids changes on passing from the reduced to the oxidized form.

It is noteworthy to stress again how similar are the overall structures in the solid state and in solution for these electron transfer proteins (Banci et al., 1994a). Actually, this is consistent with the idea of electron transfer proteins, whose structures should be rather rigid and should imply no major rearrangements between oxidized and reduced states. This becomes more noticeable the more the quality of the solution structure approaches that of the X-ray structure. In the present case the high quality solution structure does not leave much room for discussion. In a way, this observation encourages us to proceed toward the solution structure of the oxidized form, which is in progress in our lab. Small structural variations, or more precisely, small variations on the part of the protein that can be easily monitored through NOE's, may provide significant hints to the understanding of electron transfer processes.

ACKNOWLEDGMENT

A.D. thanks the International Centre for Genetic Engineering and Biotechnology (UNIDO) and D.H.W.K. thanks the C.E.C. Human Capital and Mobility Program for research postdoctoral fellowships. P.S. thanks the Italian Ministry of Foreign Affairs for a Ph.D. grant.

SUPPLEMENTARY MATERIAL AVAILABLE

A listing of the 1147 experimental NOE and NOESY intensities used for the structure calculation is available (16 pages). Ordering information is given on any current masthead page.

REFERENCES

Arseniev, A., Schultze, P., Worgotter, E., Braun, W., Wagner, G., Vasak, M., Kagi, J. H., & Wüthrich, K. (1988) *J. Mol. Biol. 201*, 637−657.

Backes, G., Mino, Y., Loehr, T. M., Meyer, T. E., Cusanovich, M. A., Sweeney, W. V., Adman, E. T., & Sanders-Loehr, J. (1991) *J. Am. Chem. Soc. 113*, 2055−2064.

Banci, L., Bertini, I., Luchinat, C., Piccioli, M., Scozzafava, A., & Turano, P. (1989) *Inorg. Chem. 28*, 4650−4656.

Banci, L., Bertini, I., Carloni, P., Luchinat, C., & Orioli, P. L. (1992a) *J. Am. Chem. Soc. 114*, 10683−10689.

Banci, L., Carloni, P., La Penna, G., & Orioli, P. L. (1992b) *J. Am. Chem. Soc. 114*, 6994−7001.

Banci, L., Bertini, I., Capozzi, F., Carloni, P., Ciurli, S., Luchinat, C., & Piccioli, M. (1993) *J. Am. Chem. Soc. 115*, 3431−3440.

Banci, L., Bertini, I., Eltis, L. D., Felli, I., Kastrau, D. H. W., Luchinat, C., Piccioli, M., Pierattelli, R., & Smith, M. (1994a) *Eur. J. Biochem. 225*, 703−714.

Banci, L., Bertini, I., & Luchinat, C. (1994b) in *Methods in Enzymology* (James, T. L., & Oppenheimer, N. J., Eds.) Vol. 239, pp 485−514, Academy Press, Orlando, FL.

Banci, L., Carloni, P., & Gori Savellini, G. (1994c) *Biochemistry 33*, 12356−12366.

Bartsch, R. G. (1978) *Methods Enzymol. 53*, 329.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983a) *J. Am. Chem. Soc. 105*, 7188−7190.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983b) *J. Magn. Reson. 55*, 301−315.

Benning, M. M., Meyer, T. E., Rayment, I., & Holden, H. M. (1994) *Biochemistry 33*, 2476−2483.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684−3690.

Bertini, I., Briganti, F., Luchinat, C., Scozzafava, A., & Sola, M. (1991) *J. Am. Chem. Soc. 113*, 1237−1245.

Bertini, I., Capozzi, F., Ciurli, S., Luchinat, C., Messori, L., & Piccioli, M. (1992) *J. Am. Chem. Soc. 114*, 3332−3340.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.

Breiter, D. R., Meyer, T. E., Rayment, I., & Holden, H. M. (1991) *J. Biol. Chem. 266*, 18660−18667.

Carney, M. J., Papaefthymiou, G. C., Spartalian, K., Frankel, R. B., & Holm, R. H. (1988) *J. Am. Chem. Soc. 110*, 6084−6095.

Carter, C. W. J. (1977) in *Iron−Sulfur Proteins* (Lowenberg, W., Ed.) p 157, Academic Press, New York.

Carter, C. W. J., Kraut, J., Freer, S. T., Xuong, N.-H., Alden, R. A., & Bartsch, R. G. (1974) *J. Biol. Chem. 249*, 4212−4215.

Case, D. A., & Karplus, M. (1979) *J. Mol. Biol. 132*, 343−368.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820−2821.

Dugad, L. B., La Mar, G. N., Banci, L., & Bertini, I. (1990) *Biochemistry 29*, 2263−2271.

Eccles, C., Güntert, P., Billeter, M., & Wüthrich, K. (1991) *J. Biomol. NMR 1*, 111−130.

Eltis, L. D., Iwagami, S. G., & Smith, M. (1994) *Protein Eng.* (in press).

Gaillard, J., Albrand, J.-P., Moulis, J.-M., & Wemmer, D. E. (1992) *Biochemistry 31*, 5632−5639.

Griesinger, C., Sörensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 73*, 574−579.

Griesinger, C., Sörensen, O. W., & Ernst, R. R. (1989) *J. Magn. Reson. 84*, 14−63.

Güntert, P., & Wüthrich, K. (1991) *J. Biomol. NMR 1*, 447−456.

Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.

Holm, R. H., Ciurli, S., & Weigel, J. A. (1990) in *Progress in Inorganic Chemistry: Bioinorganic Chemistry.* (Lippard, S. J., Ed.) *Vol. 38*, pp 1−74, John Wiley & Sons, New York.

Hoops, S. C., Anderson, K. W., & Merz, K. M., Jr. (1991) *J. Am. Chem. Soc. 113*, 8262−8270.

Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1-I*, 736−751.

Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72−84.

Lecomte, J. T. J., Unger, S. W., & La Mar, G. N. (1991) *J. Magn. Reson. 94*, 112−122.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95.

Merz, K. M., Jr., & Kollman, P. A. (1989) *J. Am. Chem. Soc. 111*, 5649−5658.

Nettesheim, D. G., Harder, S. R., Feinberg, B. A., & Otvos, J. D. (1992) *Biochemistry 31*, 1234−1244.

Oh, B.-H., & Markley, J. L. (1990) *Biochemistry 29*, 3993−4004.

Pearlman, D. A., & Case, D. A. (1991) *SANDER*, University of California, San Francisco.

Pearlman, D. A., Case, D. A., Caldwell, J. C., Siebel, G. L., Singh, U. C., Weiner, P., & Kollman, P. A. (1991) in *AMBER 4.0*, University of California, San Francisco.

Phillips, W. D., McDonald, C. C., Stombaugh, N. A., & Orme-Johnson, W. H. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 140−143.

Pochapsky, T. C., Mei Ye, X., Ratnaswamy, G., & Lyons, T. A. (1994) *Biochemistry 33*, 6424−6432.

Rayment, I., Wesemberg, G., Meyer, T. E., Cusanovich, M. A., & Holden, H. M. (1992) *J. Mol. Biol. 228*, 672−686.

Ross, W. S. (1994) *CARNAL*, Department of Pharmaceutical Chemistry, University of California, San Francisco (unpublished).

Schultze, P., Worgotter, E., Braun, W., Wagner, G., Vasak, M., Kagi, J. H., & Wüthrich, K. (1988) *J. Mol. Biol. 203*, 251.

Shaka, A. J., Keeler, J., Frenkiel, T., & Freeman, R. (1983) *J. Magn. Reson. 52*, 335−338.

Shen, J., Wong, C. F., Subrasmaniam, S., Albright, T. A., & McCammon, J. A. (1990) *J. Comput. Chem. 11*, 346−350.

van Gunsteren, W. F., & Berendsen, H. J. C. (1977) *Mol. Phys. 34*, 1311−1327.

Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 287−303.

Wüthrich, K. (1986) *NMR of Protein and Nucleic Acids*, Wiley, New York.

Xavier, A. V., Turner, D. L., & Santos, H. (1993) in *Methods in Enzymology* (Vallee, B. L., & Riordan, J. F., Eds.) Academic Press, San Diego, CA.

BI941745I