Protein Science (2000), *9*:1474–1486. Cambridge University Press. Printed in the USA.

Solution structure of a zinc substituted eukaryotic rubredoxin from the cryptomonad alga *Guillardia theta*

KRISTIAN SCHWEIMER,[1] SILKE HOFFMANN,[1] JÜRGEN WASTL,[2]
UWE G. MAIER,[2] PAUL RÖSCH,[1] AND HEINRICH STICHT[1]

[1]Lehrstuhl für Biopolymere, Universität Bayreuth, D-95440 Bayreuth, Germany
[2]Cell Biology and Applied Botany, Philipps-University Marburg, D-35032 Marburg, Germany

(RECEIVED March 20, 2000; FINAL REVISION June 2, 2000; ACCEPTED June 2, 2000)

Abstract

The rubredoxin from the cryptomonad *Guillardia theta* is one of the first examples of a rubredoxin encoded in a eukaryotic organism. The structure of a soluble zinc-substituted 70-residue *G. theta* rubredoxin lacking the membrane anchor and the thylakoid targeting sequence was determined by multidimensional heteronuclear NMR, representing the first three-dimensional (3D) structure of a eukaryotic rubredoxin. For the structure calculation a strategy was applied in which information about hydrogen bonds was directly inferred from a long-range HNCO experiment, and the dynamics of the protein was deduced from heteronuclear nuclear Overhauser effect data and exchange rates of the amide protons. The structure is well defined, exhibiting average root-mean-square deviations of 0.21 Å for the backbone heavy atoms and 0.67 Å for all heavy atoms of residues 7–56, and an increased flexibility toward the termini. The structure of this core fold is almost identical to that of prokaryotic rubredoxins. There are, however, significant differences with respect to the charge distribution at the protein surface, suggesting that *G. theta* rubredoxin exerts a different physiological function compared to the structurally characterized prokaryotic rubredoxins. The amino-terminal residues containing the putative signal peptidase recognition/cleavage site show an increased flexibility compared to the core fold, but still adopt a defined 3D orientation, which is mainly stabilized by nonlocal interactions to residues of the carboxy-terminal region. This orientation might reflect the structural elements and charge pattern necessary for correct signal peptidase recognition of the *G. theta* rubredoxin precursor.

Keywords: *Guillardia theta*; NMR; rubredoxin; structure; zinc substitution

Rubredoxins are small iron proteins (~6 kDa) that contain one iron atom tetrahedrally coordinated by four cysteines. Despite the fact that all rubredoxins show very similar redox potentials in the range of 0 to −50 mV, numerous different physiological functions have been reported for them. For example, rubredoxin is described to participate in sulfate reduction of *Desulfovibrio gigas* (Gomes et al., 1997), in nitrate reduction of *Clostridium perfringens* (Seki et al., 1988), in hydrogen oxidation of *Azotobacter vinelandii* (Chen & Mortensen, 1992), and in alkane degradation of *Acinetobacter calcoaceticus* (Geissdorfer et al., 1995).

The structures of the prokaryotic rubredoxins from *D. gigas* (Frey et al., 1987), *Desulforibrio desulfuricans* (Stenkamp et al., 1990), *Desulfovibrio vulgaris* (Adman et al., 1991; Dauter et al., 1992; Misaki et al., 1999), *Clostridium pasteurianum* (Watenpaugh et al., 1979; Dauter et al., 1996), and from the hyperthermophilic *Pyrococcus furiosus* (Day et al., 1992; Bau et al., 1998) have been determined by X-ray crystallography. NMR spectroscopy was applied for the structure determination of reduced *C. pasteurianum* rubredoxin (Bertini et al., 1998) and of a zinc-substituted form of *P. furiosus* rubredoxin (Blake et al., 1992). The structures of these prokaryotic rubredoxins are quite similar, generally showing RMSDs of <1.0 Å for the protein backbone. The geometry of the cluster vicinity is maintained by a network of six conserved hydrogen bonds that are formed between backbone amide protons and the sulfur atoms of the iron-ligating cysteines (Sieker et al., 1994). Common structural elements of these rubredoxins include one β-sheet, 3_{10}-helical turns, glycine-containing turns, and a hydrophobic core that is mainly formed by aromatic amino acids. An invariant and mainly hydrophobic region in the cluster vicinity was suggested to provide the common docking and electron exchange region, whereas a more variable region of the protein confers specificity to the interaction with the rubredoxin redox partners (Sieker et al., 1994).

Recently, the first eukaryotically encoded rubredoxins have been identified by analyzing the genome sequences of *Arabidopsis thaliana* and of the *Guillardia theta* nucleomorph. *G. theta* is a

Reprint requests to: Heinrich Sticht, Lehrstuhl für Biopolymere, Universität Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany; e-mail: Heinrich.Sticht@uni-bayreuth.de.

Abbreviations: COSY, correlation spectroscopy; DSS, 4,4-dimethyl-4-silapentane sodium sulfate; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; RMSD, root-mean-square deviation.

representative of the cryptomonads, which are unicellular biflagellate algal cells. A characteristic feature of cryptomonads is the presence of a so-called nucleomorph—a very small eukaryotic genome (520–660 kb) organized in three tiny chromosomes (Rensing et al., 1994). Phylogenetic analysis has shown that nucleomorphs are remnant nuclei of former free-living red algae that have been engulfed by another eukaryotic cell and established as a phototrophic symbiont (Gilson et al., 1997; McFadden et al., 1997).

G. theta and *A. thaliana* rubredoxins contain a core-sequence of ~55 amino acids that exhibits significant sequence homology to prokaryotic rubredoxins including the conserved pattern of iron-ligating cysteines (Fig. 1). Additional structural features not present in prokarytic rubredoxins include an amino-terminal signal sequence and a carboxy-terminal membrane anchor that is linked by a stretch of ~20 amino acids to the core-sequence.

Structural investigation of *G. theta* rubredoxin thus should be helpful in identifying the physiological function of eukaryotic rubredoxins (e.g., by verifying surface complementarity to putative redox partners) and in addressing the question whether the residues extending the core-sequence adopt a defined three-dimensional (3D) fold.

This latter question is of particular interest for residues 57–62 of the precursor, which represent the putative recognition/cleavage site of the signal peptidase (J. Wastl, unpubl. obs.), and for residues 106–120 of the precursor that are expected to adopt an α-helical structure from secondary structure prediction using the PhD software (Rost & Sander, 1993).

For that reasons, we determined the solution structure of a 70 residue *G. theta* rubredoxin comprising residues 57–126 of the precursor from multidimensional heteronuclear NMR data. To avoid paramagnetic effects arising from the presence of an iron atom, a zinc-substituted form was used for structural investigation. Iron- and zinc-substituted forms of *C. pasteurianum* rubredoxin are structurally almost indistinguishable as evidenced from X-ray data (Dauter et al., 1996).

Results and discussion

Resonance assignment and NOE analysis

In the well-resolved ^{1}H, ^{15}N HSQC spectrum of *G. theta* rubredoxin (Fig. 2) all expected resonances (62 backbone amides, four pairs of glutamine and asparagine side-chain amides, and one arginine H^{ϵ} resonance) could be observed without signal overlap. The $^{1}H^{N}$, ^{15}N, $^{13}C^{\alpha}$, and $^{13}C^{\beta}$ resonances could be automatically assigned sequentially using an in-house written search algorithm based on inter- and intraresidual C^{α} and C^{β} chemical shifts taken from the CBCA(CO)NH and HNCACB spectra for sequential linking of amide resonances and amino acid type determination. Furthermore, using the HBHA(CO)NH and HCCH-COSY data, all H^{α} and H^{β} could be assigned. For longer aliphatic side chains, complete carbon and proton assignments were made from the 3D H(C)CH-COSY and C(CO)NH spectra, while aromatic proton resonances were assigned from homonuclear 2D spectra and from

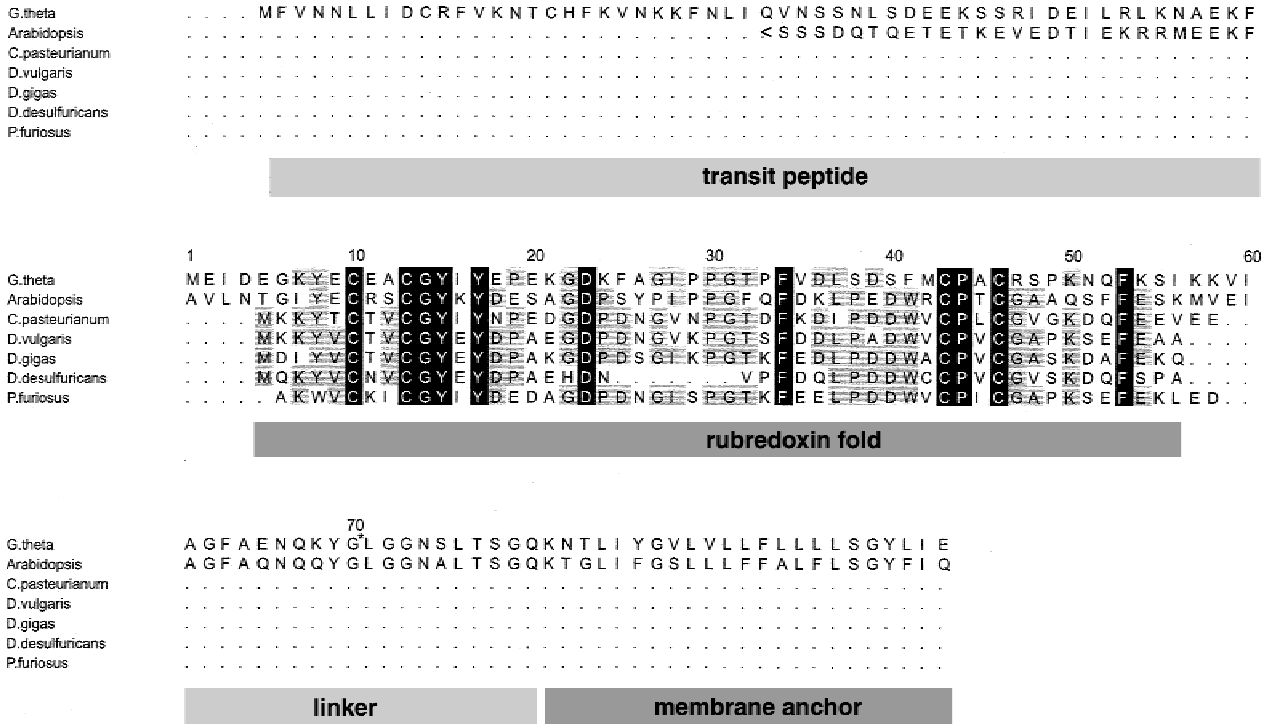

Fig. 1. Multiple sequence alignment of the eukaryotic rubredoxins from *G. theta*, *A. thaliana*, and the structurally characterized prokaryotic rubredoxins from *C. pasteurianum*, *D. vulgaris*, *D. gigas*, *D. desulfuricans*, and *P. furiosus*. Conserved amino acids in all sequences are shaded black, conserved amino acids in more than 50% of the compared sequences are shaded gray. The boundaries of the recombinant nucleomorph rubredoxin used for structure determination are indicated with *. The first 65 amino acids of the hypothetical *Arabidopsis* rd protein are not displayed due to missing homology (marked with <). Possible functions of parts of the nucleomorph rd are shown below. The alignment was generated using the programs ClustalW (Higgins et al., 1992) and Alscript (Barton, 1993).

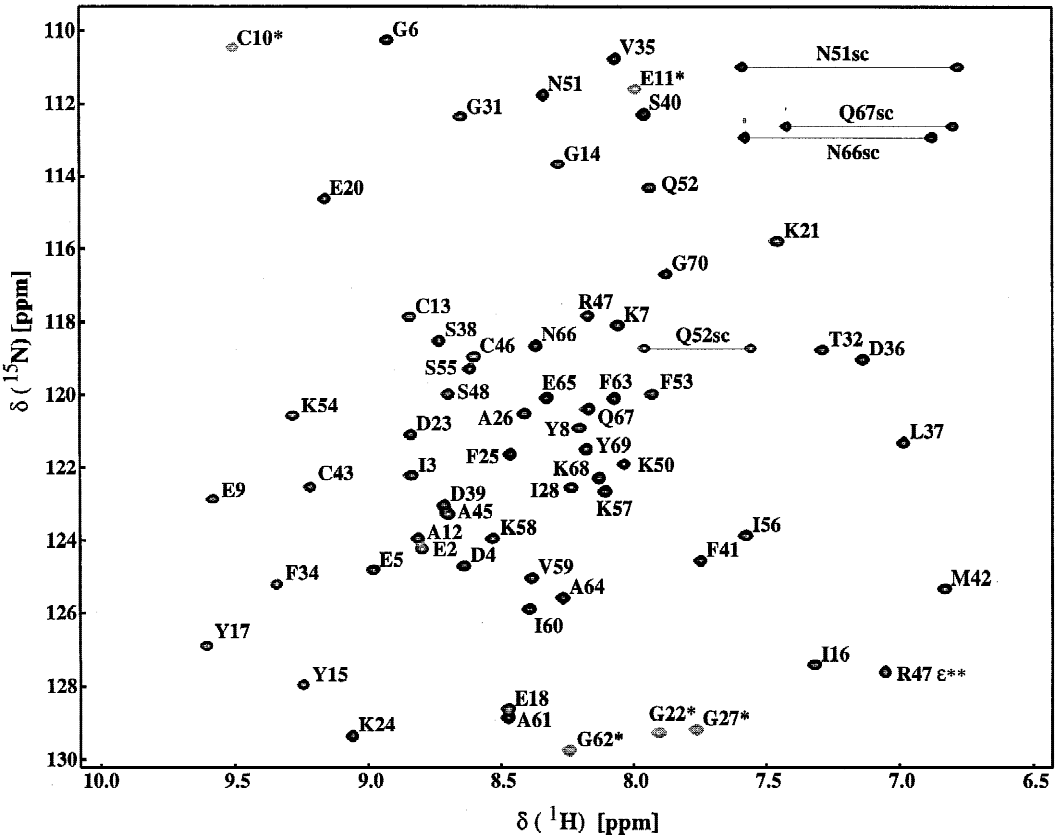

Fig. 2. 600 MHz [^{1}H,^{15}N]-HSQC spectrum of ^{13}C/^{15}N labeled *G. theta* rubredoxin at 25 °C that was recorded with ^{13}C decoupling during the ^{15}N evolution by a 180° composite pulse. Resonances are labeled with the corresponding sequence positions. Side-chain NH_2 resonances are connected. Aliased resonances are marked with an asterisk.

the ^{13}C-edited 3D-NOESY. Based on these types of experiments, all side-chain resonances could be assigned with exception of three phenylalanine H^{ζ} resonances and some lysine resonances (H^{γ}, H^{δ}) for which a strong signal overlap was present in the spectra. Y17 H^{η} (9.20 ppm) and T32 $H^{\gamma 1}$ (5.89 ppm) exchange slowly on the NMR timescale under the experimental conditions, and could be unambiguously assigned using characteristic NOEs observed in the 2D NOESY spectrum.

Analysis of the NOE pattern (Fig. 3) revealed the presence of a short triple-stranded antiparallel β-sheet (G14–E20, K7–A12, and Q52–I56), which also exists in the structures of prokaryotic rubredoxins. Additional NOEs typical for β-sheets indicate that residues E2–D4 and K57–V59 are aligned in an antiparallel fashion. Although K57 is next to I56 of the triple-stranded sheet, residues E5 and G6 of the opposite strand have to be bulged out based on the NOE pattern observed (Fig. 3).

Most of the long-range NOEs that are critical for the determination of the 3D structure are observed for the hydrophobic amino acids Y8, Y15, Y17, I28, F34, L37, F41, and F53 suggesting the presence of a hydrophobic core, that is mainly formed by aromatic amino acids (Fig. 4). In addition, an unusually high number of long-range NOEs is present for the highly conserved K50.

For the 10 carboxyterminal residues only sequential NOEs were observed. Further, all amide protons of the corresponding residues showed fast exchange with the solvent as visible from the strong cross peak at the water frequency in the ^{15}N-NOESYHSQC, indicating that these residues are unstructured in solution. The $^3J(H^N,H^{\alpha})$ couplings constants for these residues are in the 6–8 Hz range, indicative of rotameric averaging.

Characterization of internal flexibility

Additional information about the internal flexibility of the protein was obtained from measurements of the heteronuclear $\{^1H\}^{15}N$ NOE (Fig. 5A) and of the hydrogen exchange rates using the New Mexico experiment for fast exchanging amides (Fig. 5B) and H/D exchange experiments for slow exchanging amides (Fig. 5C), respectively.

The 10 carboxyterminal residues (A61–G70) exhibit amide proton exchange rates faster than 1 s^{-1} and a negative $\{^1H\}^{15}N$ NOE typical for highly flexible, unstructured termini of polypeptide chains (Kay et al., 1989).

For residues 7–57 the heteronuclear NOE is always larger than 0.65, with an average of 0.72, indicating a highly restricted internal motion of the NH bond vector, consistent with a rigid core fold. The flanking residues show a gradual decrease of the heteronuclear NOE, indicating an increasing flexibility toward the termini. The exchange rates of the amide protons exhibit the same pattern. Exchange rates slower than 10^{-2} s^{-1} were found only in the region 7–57, whereas the flanking residues have always faster exchange rates. As expected from the observation of direct interactions between residues and E2–D4 and K57–V59, there is a

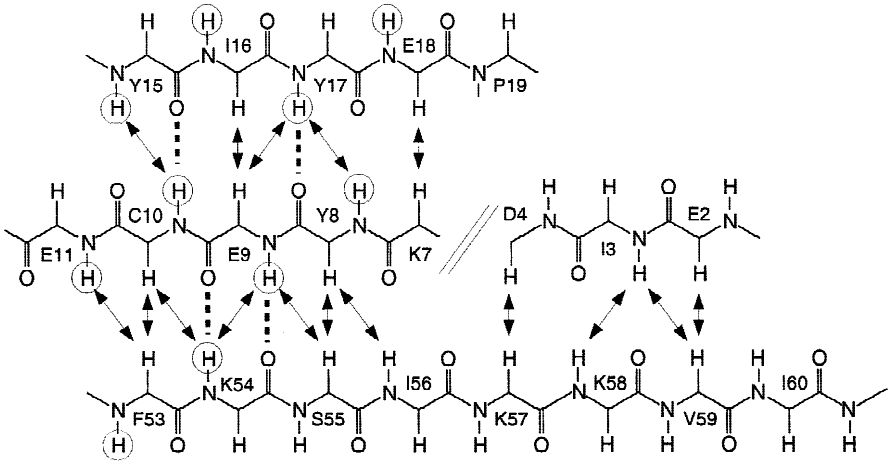

Fig. 3. Schematic representation of the extended regions in *G. theta* rubredoxin. Interstrand NOEs and hydrogens bonds unambiguously measured from a long-range 2D H(N)CO spectrum are indicated by dashed lines. Slowly exchanging backbone amide protons are marked by open circles.

correlation in the dynamic behavior of these two stretches (Fig. 5A,B).

The most slowly exchanging amide protons ($k_{ex} < 10^{-4}$ s^{-1}) are found in the vicinity of the cluster ligating cysteines (C10, C13, C43, C46) and the region of the triple-stranded β-sheet.

The contradictory exchange rates for T32 and R47, showing a fast and a slow rate depending on the type of the experiments (Fig. 5B,C), may be attributed to an exchange relayed NOE causing a signal in the New Mexico experiments that cannot be distinguished from a chemical exchange signal (Mori et al., 1997).

Identification of hydrogen bonds

Slow hydrogen exchange rates obtained from exchange experiments in D_2O (Fig. 5C) can generally be used for defining hydrogen bonds. They suffer, however, from the limitation that the acceptor of the hydrogen bond is not known a priori and can only be deduced by inspection of initial structures that were calculated without hydrogen bond restraints.

In the present study, this limitation was partially overcome by the acquisition of a long-range 2D H(N)CO spectrum that allowed the observation of five correlations caused by the ^{3h}J(N,CO) scalar

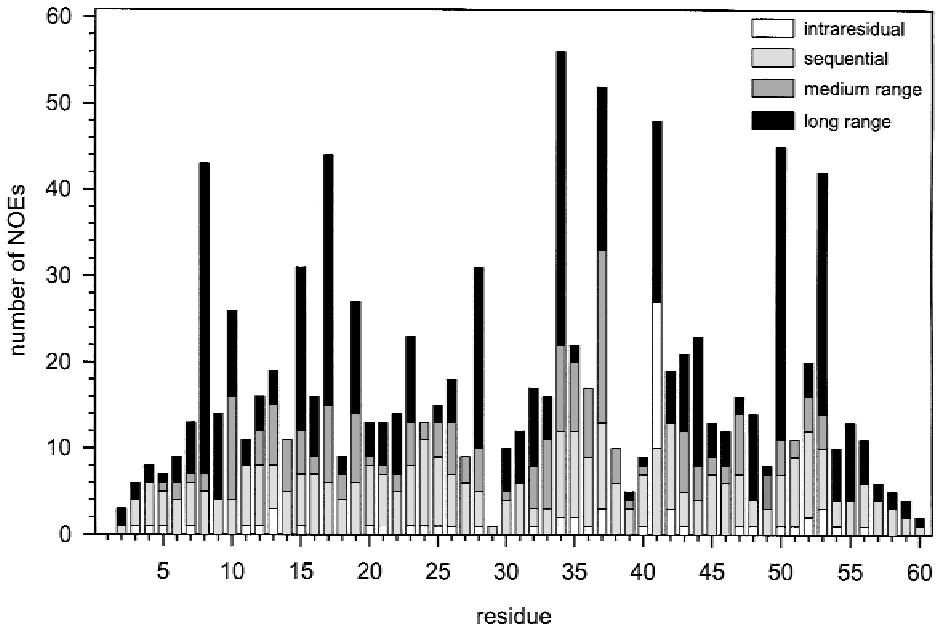

Fig. 4. Number of NOEs per residue used in the final structure calculation. NOEs are grouped into intraresidue, sequential ($i - j = 1$), medium-range ($1 < i - j < 5$), and long range ($i - j \geq 5$). Residues 61–70 have been omitted in this presentation, since only sequential NOEs have been observed for them.

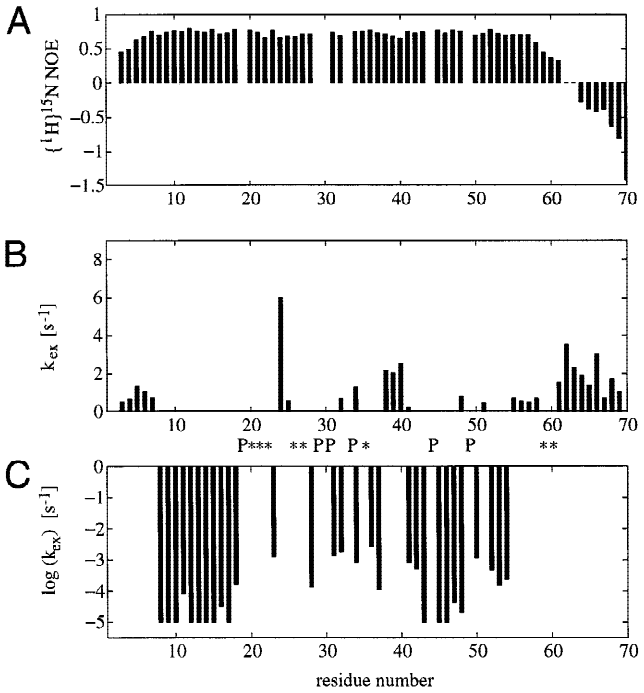

Fig. 5. $\{^1H\}$ ^{15}N NOE and amide exchange rates of *G. theta* rubredoxin. **A:** Magnitude of the $\{^1H\}$ ^{15}N NOE (I_{sat}/I_0) along the amino acid sequence. **B:** The fast amide exchange rates determined by a series of New MEXICO experiments (details of the experiment are given in Materials and methods). **C:** Logarithmic exchange rates of the slowly exchanging amide protons determined by H/D exchange. Exchange rates slower than 10^{-5} s^{-1} were arbitrary set to 10^{-5} s^{-1} (slowest quantified rate). The prolines are marked with P, and the residues with an exchange rate in the intermediate nondetectable region are labeled with an asterisk.

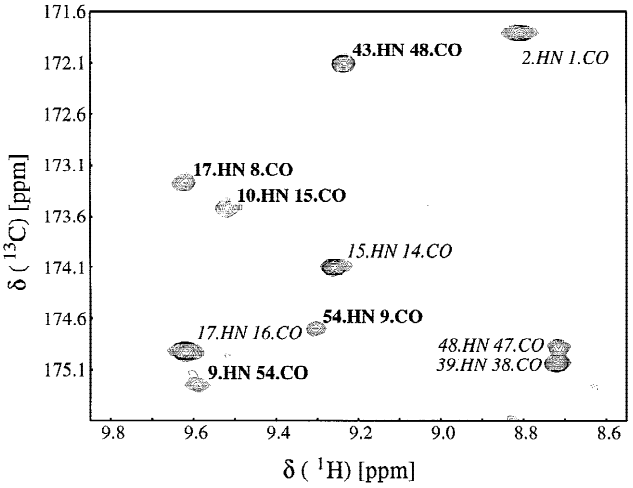

Fig. 6. Selected region of the long-range 2D H(N)CO. The H(N)CO was recorded with a dephasing delay of 133 ms for the buildup and refocusing of carbon-nitrogen antiphase coherence. The cross peaks showing correlations across the the hydrogen bonds via the $^{3h}J_{NC'}$ are marked in bold, and the residual signals caused by the $^1J_{NC'}$ are given in italics.

coupling across hydrogen bonds (Fig. 6). The corresponding acceptor for the amide protons was unambiguously identified and therefore distance restraints for these hydrogen bonds were already used in the first rounds of the structure calculation, thus facilitating the determination of the global protein fold.

Due to the selected magnetization transfer in the long-range H(N)CO experiment, only hydrogen bonds between nitrogen-bound protons and acceptors bound to carbonyl-like carbons can be detected. In contrast, H/D-exchange experiments additionally include information about hydrogen bonds formed to an acceptor different from carbonyl oxygen (e.g., a cysteine sulfur) and about protons that are buried in the interior of the protein and are thus inaccessible to water. This fact explains why the number of slowly exchanging protons observed after D_2O exchange is higher than the number of hydrogen bonds measured by ^{3h}J(N,CO) scalar coupling.

Tertiary structure of G. theta rubredoxin

The calculation of the final structures was based on 537 distance restraints (including 227 long-range NOEs), 25 dihedral angle restraints, and 18 hydrogen bonds. The resulting family of 21 converged solution structures showed no single distance violation >0.22 Å, and no torsion angle violation larger than 2°.

The carboxy-terminal residues 61–70 were excluded from further analysis because they proved to be highly unstructured, as already expected from the lack of nonsequential NOEs, fast H/D exchange, and a negative heteronuclear NOE.

According to PROCHECK analysis, residues 1–60 exhibit energetically favorable backbone conformations: 71.4% of the residues are found in the most favored regions and 28.6% in the allowed regions of the Ramachandran plot. A schematic presentation of the structure is shown in Figure 7.

The RMSD of the 21 converged structures is 0.63 Å for the backbone heavy atoms and 1.04 Å for all heavy atoms. The overlay in Figure 8A reveals that the magnitude of the RMSD is strongly affected by residues 1–6 and 57–60 that flank the core fold. These

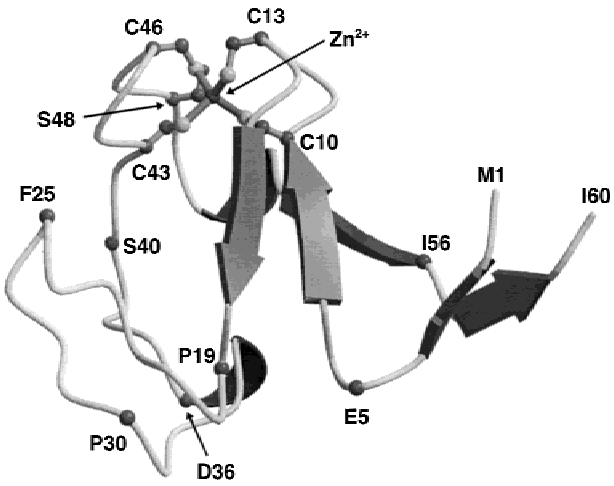

Fig. 7. Schematic presentation of the *G. theta* rubredoxin structure (residues 1–60). The four cysteines ligating the metal ion and S48, which is located in the cluster proximity, are shown in ball-and-stick presentation. For approximately each fifth residue, the location of the C^α atom is indicated in the structure and elements of secondary structure are represented schematically.

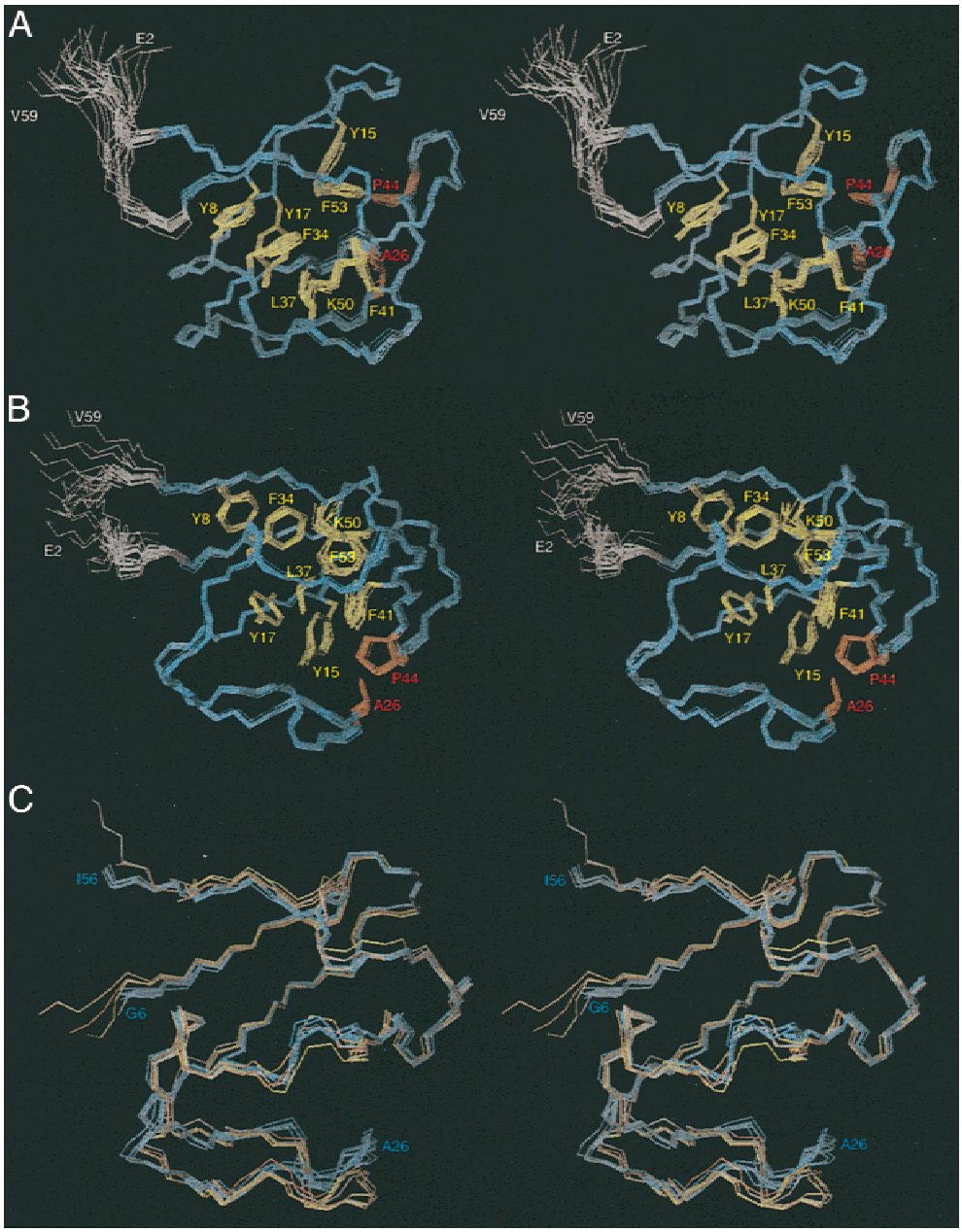

Fig. 8. **A:** Stereoview of the backbone overlay of a family of 15 structures for residues 2–59. The protein backbone is shown in blue for the rubredoxin core fold and in white for for the flanking residues. The side chains of the residues forming the hydrophobic core are shown in yellow, and A26 and P44 in red. **B:** Same presentation as in **A**, but rotated by 90° around the horizontal axis. **C:** Stereoview of the backbone overlay of the six lowest energy structures of *G. theta* rubredoxin (blue) with the high resolution crystal structures of four prokaryotic rubredoxins: *C. pasteurianum* rubredoxin (red), *D. vulgaris* rubredoxin (yellow), *D. gigas* rubredoxin (orange), and *P. furiosus* rubredoxin (white). The average pairwise RMSD of the backbone atoms of the solution structures (residues 8–56) to the crystal structures is in the range from 1.08–1.29 Å (same orientation as in **B**).

residues adopt an extended conformation and an antiparallel orientation, but exhibit an increased flexibility compared to the rubredoxin core fold.

Omission of residues 1–6 and 57–60 from the calculation of the RMSD results in significant lower values of 0.21 Å for the backbone heavy atoms and 0.67 Å for all heavy atoms (Table 1). The high precision of the structure becomes apparent from an overlay showing the protein backbone and the side chains of the amino acids that form the hydrophobic core (Fig. 8A,B). This hydrophobic core includes the aromatic amino acids Y8, Y15, Y17, F34, F41, F53, and the aliphatic side chain of L37. Stacking is observed for the rings of Y8 and F34.

Interestingly, K50 is located directly adjacent to the hydrophobic core forming numerous contacts by its methylene groups to the side chains of F34, L37, and F41, suggesting that the conservation of K50 (Fig. 1) is due to steric requirements rather than to elec-

Table 1. *Summary of structure calculation*

Experimental restraints for the final structure calculation		
Interresidual NOEs	495	
Sequential ($\|i-j\|=1$)	161	
Medium range ($\|i-j\|<5$)	107	
Long range ($\|i-j\|\geq 5$)	227	
Intraresidual NOEs	42	
Dihedral restraints		
$^3J(H^N,H^\alpha)$	25	
Hydrogen bonds	18	
Molecular dynamics statistics		
Average energy (kcal/mol)		
E_{tot}	130.8	(± 0.3)
E_{bond}	4.93	(± 0.04)
E_{angles}	93.8	(± 0.2)
$E_{improper}$	9.9	(± 0.1)
E_{repel}	12.8	(± 0.2)
E_{NOE}	9.4	(± 0.3)
E_{cdih}	0.009	(± 0.001)
RMSD from ideal distances (Å)		
NOE	0.018 (± 0.0003)	
Bonds	0.002 (± 0.000008)	
RMSD from ideal angles		
Bond angles	0.560 (± 0.005)	
Improper angles	0.372 (± 0.004)	

Atomic RMSD of 21 calculated structures (Å)

	Backbone	Heavy atoms
Residues		
Overall (residue 1–60)	0.63	1.04
RBX-fold (residue 7–56)	0.21	0.67

trostatic interactions of its amino group. Further evidence for this conclusion comes from the fact that no salt bridges were detected for the K50 side chain in *G. theta* rubredoxin and that this lysine is replaced by glutamine and alanine in *Arabidopsis* and *Alcaligenes eutrophus* rubredoxin, respectively.

Elements of regular secondary structure in *G. theta* rubredoxin include two 3_{10}-helical turns (P33–L37, K50–F53) and a triple-stranded antiparallel β-sheet formed by residues K7–E11, Y15–Y17, and Q52–I56. An additional extended region is formed by the amino- and carboxy-terminal residues E2–D4 and K57–V59 (Fig. 7).

Residues C10–Y15 connecting the first and second strand of the β-sheet constitute the first of two conserved C-x-y-C-G-z sequence motifs each containing two of the four metal-center coordinating cysteines. The second C-x-y-C-G-z sequence motif is formed by residues 43–48 of *G. theta* rubredoxin (Fig. 1).

Based on the analysis of the initial structures, 13 additional hydrogen bonds were identified on the basis of slow exchange. Six of these hydrogen bonds are formed to the four cysteines ligating the zinc ion: $A12H^N$–$C10S^\gamma$, $C13H^N$–$C10S^\gamma$, $Y15H^N$–$C13S^\gamma$, $A45H^N$–$C43S^\gamma$, $C46H^N$–$C43S^\gamma$, and $S48H^N$–$C46S^\gamma$. The fact that the acceptor is a sulfur instead of a carbonyl oxygen explains why these hydrogen bonds did not result in any signals in the long-range H(N)CO experiment.

There are, however, also seven hydrogen bonds formed to carbonyl oxygens that were not detected in the long-range H(N)CO spectrum. These hydrogen bonds are present in the two short 3_{10}-helical turns and at the ends of the β-sheet, for which an increased "fraying" is expected. Analysis of a set of structures calculated without hydrogen bond restraints revealed that the donor acceptor distance for these hydrogen bonds was in the range of 3.0–3.3 Å for the heavy atoms, which is at the upper limit for $N-H\cdots O=C$ hydrogen bonds. Assuming the reported (Cornilescu et al., 1999a, 1999b) exponential distance dependence of ^{3h}J(N,CO), couplings become very small (<0.36 Hz) for distances larger than 3.0 Å and, therefore, most likely were beyond the limit of detection in the ^{3h}J(N,CO) experiment.

Hydrogen bonds formed by the slowly exchanging Y17 H^η and T32 $H^{\gamma 1}$ were identified from the final structures (but never restrained during the calculation). For the Y17 H^η, the carbonyl oxygen of T32 could be unambiguously identified as hydrogen bond acceptor, while the T32 $H^{\gamma 1}$ forms a hydrogen bond either to the carbonyl oxygen of P29 or P30, thus stabilizing the type IV β-turn formed by the highly conserved P-P-G-T sequence (residues 29–32 in *G. theta* rubredoxin).

An unusual backbone conformation is present in the cluster vicinity: the carbonyl oxygen of Y15 forms a regular hydrogen bond to the amide proton of C10, which is typical for a β-sheet. The reverse hydrogen bond, however, does not exist and the carbonyl oxygen of C10 is hydrogen bonded to $G14H^N$ instead, while $Y15H^N$ forms a hydrogen bond to $C13S^\gamma$. This unusual backbone conformation is also reflected by the extremely small $^3J(H^N, H^\alpha)$ coupling constant (2.3 Hz) for Y15. The crystal structures of prokaryotic rubredoxins show an identical hydrogen bonding pattern, underlining the high degree of conservation of this part of the rubredoxin fold.

The flap region (E20–I28), which is one of the lesser conserved parts of rubredoxins, adopts a well-defined backbone conformation in *G. theta* rubredoxin (Fig. 8B). Tight contacts that define the orientation of this region toward the rest of the molecule are formed between the hydrophobic side chain of A26 and the ring of P44, which is located in proximity to the metal center. The spatial proximity of these residues was unambiguously confirmed by the observation of NOE cross-resonances between $A26H^\beta$ and the $H^{\delta 2}$, H^γ*, $H^{\beta 1}$, and H^α protons of P44. F25, which is not conserved among rubredoxins, does not cluster in the hydrophobic core. This residue sticks out into the solvent (Fig. 9), providing a putative recognition site for redox partners.

Structural comparison with prokaryotic rubredoxins

For comparison, the high-resolution crystal structures available for the prokaryotic rubredoxins from *C. pasteurianum* (1.1 Å; Dauter et al., 1996), *D. vulgaris* (0.92 Å; Z. Dauter, S. Butterworth, L.C. Sieker, G. Sheldrick, & K.S. Wilson, in prep.), *D. gigas* (1.4 Å; Frey et al., 1987) and *P. furiosus* (0.95 Å; Bau et al., 1998) were used. *D. desulfuricans* (Stenkamp et al., 1990) rubredoxin was excluded from the comparison because it contains a seven residue deletion corresponding to residues 24–30 of *G. theta* rubredoxin.

The comparison reveals that the overall fold of *G. theta* rubredoxin is highly similar to that found in prokaryotic rubredoxins (Fig. 8C). The spatial arrangement of the aromatic amino acids is also well conserved among all known rubredoxin structures. The replacement of the otherwise strictly conserved W41 by phenylalanine in *G. theta* rubredoxin (Fig. 1) neither affects the overall protein fold nor the side-chain orientation of residue 41 or other side chains of the hydrophobic core.

The pairwise RMSDs resulting from a best-fit superposition of the protein backbone (residues 8–56 of *G. theta* rubredoxin) are in the range from 1.08 to 1.29 Å. Exclusion of residues 21–31 (flap

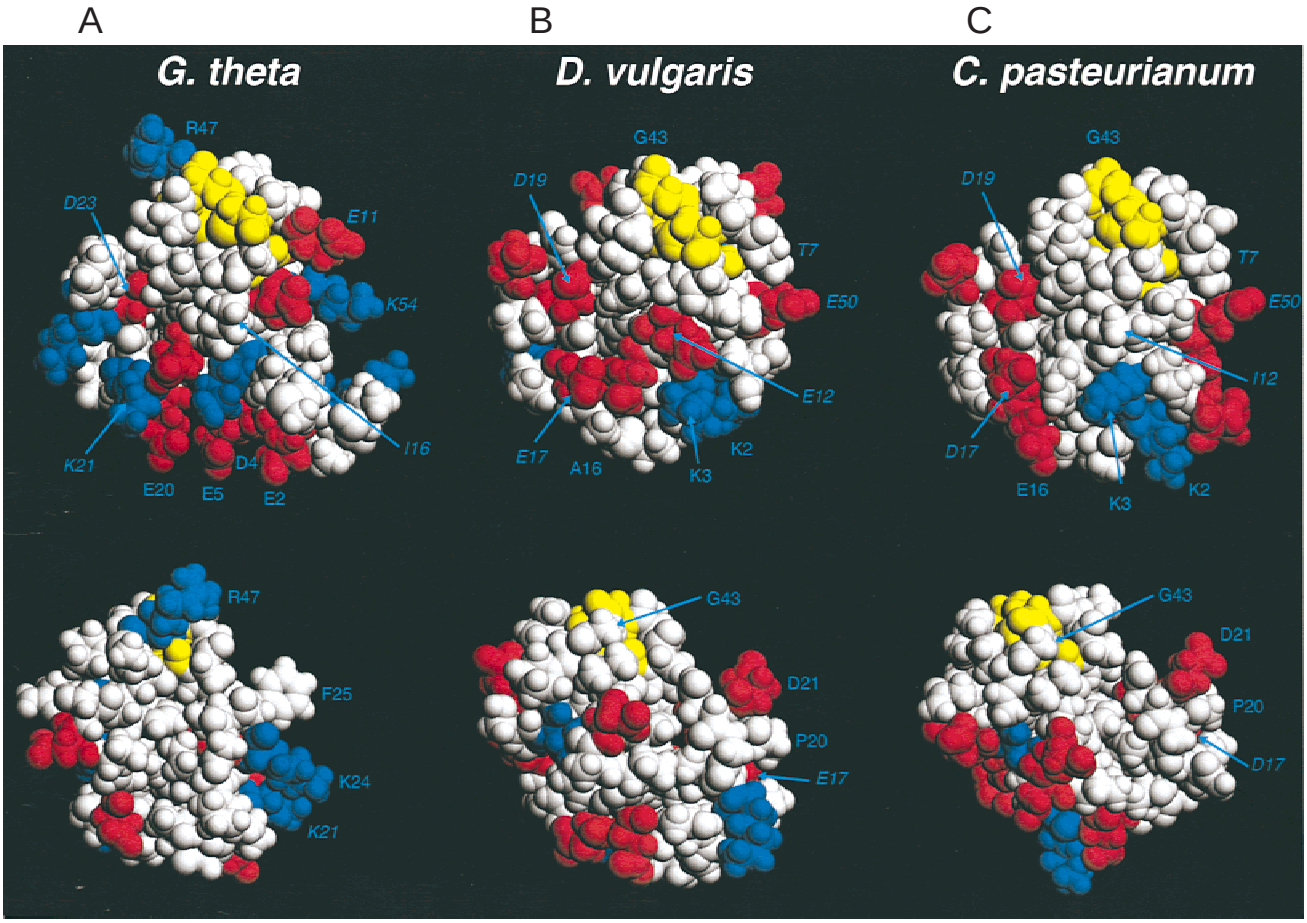

Fig. 9. Electrostatic surface properties of the rubredoxins from (**A**) *G. theta*, (**B**) *D. vulgaris*, and (**C**) *C. pasteurianum*. The upper and lower row show two different views of the proteins. Acidic residues are shown in red, basic residues in blue, and the metal-coordinating cysteines in yellow. Residues discussed in the text are labeled. Those polar residues that were proposed to be important for the rubredoxin–cytochrome c_3 interaction in *D. vulgaris* (Stewart et al., 1989) are emphasized in italics.

region), which are only moderately conserved among rubredoxins, results in a lower pairwise backbone RMSD of 0.71–0.88 Å. As evident from Figure 8C, this region is oriented closer toward the rest of the molecule in the set of *G. theta* rubredoxin structures than in the crystal structures of the other rubredoxins.

These differences most likely result from different types of interactions made by the residue at the tip of the flap region. In *G. theta* rubredoxin, an alanine (A26) is present at this position forming hydrophobic contacts to the strictly conserved P44 located in the cluster proximity (Fig. 8B). In contrast, in the structurally characterized prokaryotic rubredoxins an asparagine or serine is present that forms completely different interactions due to its polar character: the asparagine found in the rubredoxins from *C. pasteurianum*, *D. vulgaris*, and *P. furiosus* forms a highly conserved hydrogen bond to an aspartate side chain, which is three residues away in sequence, while the serine in *D. gigas* rubredoxin is solvent exposed and most probably forms a hydrogen bond to a water molecule 2.8 Å away.

The presence of an overall similar fold (including the side-chain orientation of the amino acids forming the hydrophobic core) alone, however, is insufficient to draw direct conclusions on the physiological function of *G. theta* rubredoxin because interactions with redox partners are generally mediated by the solvent-exposed side chains (Beißinger et al., 1998; Sticht & Rösch, 1998). For that reason we compared the surface properties of several rubredoxins to find clues of whether *G. theta* rubredoxin is expected to function in similar metabolic pathway as its prokaryotic counterparts or not.

This surface analysis reveals several pronounced differences between the two prokaryotic rubredoxins from *C. pasterianum* and *D. vulgaris* on one hand and *G. theta* rubredoxin on the other hand (Fig. 9). One example is found in the flap region, where lysine (K24) and phenylalanine (F25) are present in the *G. theta* protein instead of proline and aspartate, respectively (Fig. 9; lower row). Because both residues are highly solvent exposed, these substitutions have drastic effects on the shape and electrostatic properties of the protein surface. These two substitutions, together with a replacement of glycine by arginine at position 47, results in considerably altered surface properties of *G. theta* rubredoxin.

Additional differences include a large negatively charged surface patch formed by E2, D4, E5, and E20 in *G. theta* rubredoxin, while two positively charged lysines are present at the amino-terminus of *C. pasteurianum* and *D. vulgaris* rubredoxin (Fig. 9; upper row). There are also three positively charged residues (K21, K24, R47) in *G. theta* rubredoxin that have no counterparts in most other prokaryotic rubredoxins (including those from *C. pasteurianum* and *D. vulgaris*).

In addition to R47, S48 represents a second uncommon substitution observed in the immediate cluster vicinity of *G. theta* rubredoxin. Site-specific mutagenesis studies on *C. pasteurianum* and *P. furiosus* rubredoxins proved that the identity of this residue, which is generally a nonpolar amino acid, has strong effects on the redox potential (Eidsness et al., 1999). An alanine favors a redox potential close to 0 mV, while a valine results in a more negative potential of −50 mV. The side-chain orientation of S48 is identical to that found for alanine and valine in other rubredoxins, placing the O^{γ}-atom of S48 in a distance of ~4.5 Å from the metal center (Fig. 7). The fact that the side chain of S48 does not protrude into the solvent but is oriented toward the metal center suggests that this substitution might have a significant effect on the redox potential.

For *D. vulgaris* rubredoxin, a model of the complex with the tetraheme cytochrome c_3 has been proposed using computer graphic modeling and NMR spectroscopy (Stewart et al., 1989). This model suggested the negatively charged side chains of E12, E17, D19, and E50 of *D. vulgaris* rubredoxin to play a key role for the interaction with the lysines surrounding the heme crevices of cytochrome c_3. Comparison to the structure of *G. theta* rubredoxin reveals that the corresponding sequence positions are occupied by I16, K21, D23, and K54. This finding renders an interaction with cytochromes of the c-type very unlikely because this protein family generally contains a large excess of positive charges in the vicinity of the heme crevice (Stewart et al., 1989; Beißinger et al., 1998).

In summary, surface analysis revealed that *G. theta* rubredoxin differs considerably from prokaryotic rubredoxins with respect to shape and charge distribution, rendering a common physiological function of these rubredoxins rather unlikely. Biochemical assays aimed to identify the electron transfer partners of *G. theta* rubredoxin are currently carried out in our laboratory.

Role of the sequences flanking the core fold

The structure determination of *G. theta* rubredoxin reveals that residues 1–6 and 57–70 do not exhibit pronounced interactions with the core fold and thus have to be considered as a distinct structural elements. This finding is in agreement with the observation that the dynamic behavior of these flanking sequences differs significantly from that of the core fold.

Despite their increased flexibility, the amino-terminal residues containing the putative recognition/cleavage site of the signal peptidase are not unstructured, but adopt a defined 3D conformation in which residues 5 and 6 are bulged out while residues 2–4 are aligned in an antiparallel fashion with residues 57–59 as evidenced from the NOE data.

These findings might give a first clue to the yet unanswered question about the structural framework that is recognized by the signal peptidase: our present results suggest that in the case of *G. theta* rubredoxin, the peptidase might recognize a defined 3D conformation rather than a simple sequence motif.

It is particularly remarkable that this conformation predominantly results from nonsequential contacts to residues of the carboxy-terminal region. The charge complementarity observed between residues E2, D4, E5, E9, and K54, K57, and K58 suggests an important role of electrostatic interactions for stabilization.

One major difference compared to prokaryotic rubredoxins is a charge inversion placing negatively charged residues in the amino-terminal and positively charged residues in the carboxy-terminal part of the sequence. This feature might be a consequence of a specific requirement of the peptidase for negative residues in the vicinity of the cleavage site, while the flexibility observed for this sequence region might be important for access of the peptidase or for allowing a dissociation of the signal sequence after cleavage.

Further conclusions on this subject, however, are hampered by the fact that the signal peptidase responsible for rubredoxin processing in *G. theta* as well as the exact cleavage site are not known up to date. Information from sequence and structure analysis of additional proteins targeted to the chloroplast might be helpful in proving whether the observations made for *G. theta* rubredoxin represent a general principle.

Materials and methods

Cloning, expression, and purification of G. theta rubredoxin

G. theta rubredoxin, spanning amino acids 57–126 of the rubredoxin precursor (TrEMBL accession number Q9XG40), was cloned into pET28a (Novagen, Madison, Wisconsin) and overexpressed in *Escherichia coli* BL21(DE3) (Grodberg & Dunn, 1989). The numbering scheme used throughout this paper will refer to the expressed protein, starting from M1 instead of M57.

^{15}N-/^{13}C-labeled protein samples were prepared from cells grown in M9-medium (Sambrook et al., 1989) supplemented with 2 g/L ^{13}C-glucose, 0.5 g/L $^{15}NH_4Cl$, 2 mL/L TS2 trace element solution (Meyer & Schlegel, 1983), 0.1 mM $CaCl_2$, and 2 mM $MgSO_4$, lacking any further Fe-salts except for possible impurities of the used chemicals. $ZnCl_2$ was added to a final concentration of 50 μM to yield Zn-substituted *G. theta* rubredoxin. The protein was purified on a DEAE-sepharose column, followed by gelfiltration and a Mono-Q-sepharose column.

NMR spectroscopy

All spectra were recorded at 25 °C on a Bruker DRX600 spectrometer with pulsed field gradient capabilities. Quadrature detection in the indirect dimensions was obtained by the States-TPPI method (Marion et al., 1989) or by the echo/antiecho method (Kay et al., 1992; Schleucher et al., 1993), if coherence selection with gradients was employed. All experiments were recorded on samples containing 2.6 mM protein, 10 mM potassium phosphate, pH 6.5, in H_2O/D_2O (9:1).

For assignment of the backbone and H^{β}/C^{β} chemical shifts, the following set of experiments were recorded: 2D ^1H, ^{15}N FHSQC (Mori et al., 1995), 3D CT-HNCO, 3D CT-HNCA (Grzesiek & Bax, 1992a), 3D HNCACB (Wittekind & Mueller, 1993), 3D CBCA(CO)NH (Grzesiek & Bax, 1992b), 3D HNHA (Vuister & Bax, 1993; Zhang et al., 1997), and 3D HBHA(CO)NH (Grzesiek & Bax, 1993a). The binomial 3-9-19 watergate element was used for water suppression in these experiments (Sklenar et al., 1993). All experiments with a starting $^1H \rightarrow {}^{15}N$ INEPT sequence included a water flipback scheme for minimizing saturation transfer (Grzesiek & Bax, 1993b). All carbon pulses were applied on one radiofrequency channel, using appropriate combinations of on/off resonant band-selective pulses [G3 and G4 Gaussian cascades (Emsley & Bodenhausen, 1990)] as well as proper calibrated rectangular pulses. For further aliphatic $^{13}C/^1H$ assignments 3D C(CO)NH and 3D H(CCO)NH (Grzesiek et al., 1993) with Watergate for water suppression and ^1H-^{13}C CT-HSQC (Vuister & Bax, 1992),

and 3D H(C)CH-COSY with gradient coherence selection (Ikura et al., 1991; Gehring & Ekiel, 1998) were recorded. A ^{15}N-NOESY-HSQC with a water flipback scheme (Talluri & Wagner, 1996) and a ^{13}C-NOESYHSQC using a sensitivity enhanced HSQC step with gradient coherence selection were recorded for deriving distance restraints for structure calculation. A low-power GARP-I sequence (Shaka et al., 1985) was applied for ^{15}N or ^{13}C broadband decoupling during the data acquisition. Proton broadband decoupling during the heteronuclear transfer steps in the triple resonance and H(C)CH-COSY experiments was achieved by the WALTZ-16 sequence (Shaka et al., 1983). Aromatic proton resonances were assigned from a 2D NOESY experiment (Jeener et al., 1979) with 120 ms mixing time and a 2D TOCSY (Bax & Davis, 1985) spectra with 80 ms mixing time recorded with an unlabeled sample. The DIPSI-2rc sequence was applied for ^{1}H TOCSY mixing (Cavanagh & Rance, 1992). For both experiments excitation sculpting for water suppression was employed (Hwang & Shaka, 1995). A 2D long-range H(N)CO was recorded for detecting hydrogen bonds involving amide protons and carbonyl groups (Cordier & Grzesiek, 1999; Cornilescu et al., 1999a, 1999b). A series of ^{15}N/^{1}H HSQCs were recorded after dissolving lyophilized protein in D_2O for identifying slowly exchanging amide protons, the experiments with a total duration of 5 min were acquired 10, 15, 21, 27, 35, 65, 95, 155, 216, 1,169, and 1,540 min after dissolving the sample in D_2O. A series of New Mexico experiments (mixing times 5, 10, 15, 25, 50, 100, and 200 ms) were recorded for measuring the exchange rates of the fast exchanging amide protons (Koide et al., 1995). The original pulse sequence was modified with a final FHSQC detection scheme (Mori et al., 1995) instead of the published HSQC with coherence selection by gradients. The $\{^{1}H\}^{15}N$ NOE experiments were recorded using the pulse sequences of Dayie and Wagner (1994). The relaxation delay was 4 s, and the proton saturation was performed by 120° high-power pulses with an interpulse delay of 5 ms for the final 3 s of the relaxation delay of the saturation experiment. Experimental details of each experiment including sweep width and time domain matrix sizes are given in Table 2.

NMR data processing and analysis

The NMR datasets were processed using in house written software and analyzed with the program packages NMRView (Johnson & Blevins, 1994) and NDEE (SpinUp Inc., Dortmund, Germany).

Data processing consists typically of SVD-Linear Prediction (Barkhuijsen et al., 1985) with root reflection (Press et al., 1992) in one heteronuclear dimension (normally the ^{15}N dimension of triple resonance experiments or the X-dimension in the X-edited spectra), apodization with 60–90° shifted squared sinebells, one zero filling in all dimensions, and Fourier transformation. For constant time evolution periods, mirror image linear prediction (Zhu & Bax, 1990) was employed. Finally baseline correction in the acquisition dimension was performed using a model free algorithm (Friedrich, 1995).

The proton chemical shifts were referenced to external DSS at 0.0 ppm. The chemical shifts of ^{13}C and ^{15}N resonances were referenced indirectly using the Ξ ratios of the zero-point frequencies at 298 K: 0.10132905 for ^{15}N/^{1}H and 0.25144952 for ^{13}C/^{1}H (Live et al., 1984).

The backbone resonances were automatically assigned with an in-house written search algorithm using inter- and intraresidual C^{α} and C^{β} chemical shifts for sequential linking of amide resonances and amino acid type determination. Aliphatic side-chain carbon and proton resonances were assigned by analyzing the HBHA(CO)NH, H(CCO)NH, C(CO)NH, and H(C)CH-COSY data. Aromatic proton resonance assignments were made by analysis of the homonuclear 2D NMR experiments. $^{3}J(H^{N},H^{\alpha})$ coupling con-

Table 2. *Experiments recorded with G. theta rubredoxin samples*

	F1[a]			F2			F3				
Experiment	Nuc[b]	SW[c]	TD[d]	Nuc	SW	TD	Nuc	SW	TD	NS[e]	Mixing time (ms)
HNCO	^{13}C	1,509.3	40	^{15}N	1,307.6	30	^{1}H	7,183.9	512	8	
HNCA	^{13}C	4225.7	40	^{15}N	1,307.6	30	^{1}H	7,183.9	512	16	
CBCA(CO)NH	^{13}C	8,725.7	52	^{15}N	1,307.6	30	^{1}H	7,183.9	512	16	
HNCACB	^{13}C	8,752.7	44	^{15}N	1,307.6	30	^{1}H	7,183.9	512	16	
HBHA(CO)NH	^{1}H	3,600.7	64	^{15}N	1,307.6	30	^{1}H	7,183.9	512	8	
C(CO)NH	^{13}C	9,960.2	64	^{15}N	1,307.6	30	^{1}H	7,183.9	512	16	16
H(CCO)NH	^{1}H	4,200.8	64	^{15}N	1,307.6	30	^{1}H	7,183.9	512	16	16
HCCH-COSY	^{13}C	5,281.9	32	^{1}H	4,249.9	70	^{1}H	8,389.3	512	8	
^{13}C NOESY-HSQC	^{1}H	7,501.9	128	^{13}C	5,281.9	32	^{1}H	8,389.3	512	8	120
^{15}N NOESY-HSQC	^{1}H	7,501.9	120	^{15}N	1,307.6	30	^{1}H	7,183.9	512	8	120
HNHA	^{1}H	4,249.9	64	^{15}N	1,307.6	36	^{1}H	7,183.9	512	16	
^{1}H, ^{15}N HSQC	^{15}N	1,307.6	128	^{1}H	7,183.9	1,024				8	
^{1}H, ^{13}C ct HSQC	^{13}C	9,657.2	480	^{1}H	8,389.3	1,024				16	
Long-range H(N)CO	^{13}C	1,660.2	128	^{1}H	7,183.9	512				256	
TOCSY	^{1}H	8,389.9	400	^{1}H	8,389.9	1,024				32	
NOESY	^{1}H	8,389.9	512	^{1}H	8,389.9	1,024				32	100

[a] F1,2,3 = frequency dimension.
[b] nuc = observed nucleus.
[c] SW = spectral width (Hz).
[d] TD = complex time domain data points.
[e] NS = number of scans.

stants were measured from the cross peak to diagonal peak ratios in the HNHA, corrected by a factor 1.05 (Vuister & Bax, 1993; Düx et al., 1997).

Data analysis for exchange experiments

The slow exchange rates were determined from a nonlinear two parameter fit (I_0, k_{ex}) to a single exponential decay (Equation 1) of the time-dependent peak intensities of the HSQC series after dissolving the labeled sample in D_2O.

$$I(t) = I_0 \exp\{-k_{ex}\, t\}. \quad (1)$$

The fast exchange rates were determined from a nonlinear parameter fit for the longitudinal relaxation rate of the amide proton, R_1^A, and the exchange rate k_{ex} to Equation 2 with the intensities of the New Mexico experiments (Mori et al., 1997).

$$I(t_m)/I_{ref} = [\exp\{-R_1^{H_2O} * t_m\} - \exp\{-(R_1^A + k_{ex}) * t_m\}] * k_{ex}/(k_{ex} + R_1^A - R_1^{H_2O}). \quad (2)$$

The longitudinal relaxation rate of the water $R_1^{H_2O}$ was determined with a gradient assisted inversion recovery experiment (0.42 s^{-1}) to avoid radiation damping. Peak amplitudes instead of integrals were used because the detection scheme was the same for all experiments of a time series. Therefore, the same line shape for a given amide resonance was obtained for all data points. Data fitting was performed using the Nelder–Mead algorithm implemented in the MATLAB software package (Mathworks, Inc., Natick, Massachusetts).

Experimental restraints for the structure calculation

NOE cross peaks were categorized as "strong," "medium," and "weak," and converted into upper limit distance constraints of 2.7, 3.5, and 5.0 Å, respectively. For distances involving either methylene protons without stereospecific assignments or methyl protons $\langle r^{-6}\rangle^{-1/6}$ averaged distances were applied. For the calculation of initial structures only unambiguous restraints were used. Additional restraints were included in several rounds of structure calculation after inspection of the initial structures.

$^3J_{HN\alpha}$ values < 6.0 Hz were converted to φ-angles according to the Karplus equation allowing deviations of $\pm 25°$ from the derived angle and $^3J_{HN\alpha}$-coupling constants >8.0 Hz were translated to φ angle constraints of $-120 \pm 40°$.

Hydrogen bond restraints were directly included in the first rounds of structure calculation if the hydrogen bond acceptor could be unambiguously deduced from the 2D long-range H(N)CO experiments. Hydrogen bonds deduced from slow exchange in D_2O were introduced in the final round of the calculation if three criteria were met: Slow exchange ($<10^{-2}$ Hz) of the corresponding amide proton, a N—H···O distance <2.3 Å, and an O···H—N angle >120° in at least 70% of the unrestrained structures. For each hydrogen bond, two distance restraints were introduced into the calculation: $d_{HN\text{-}O}$ = 1.7–2.3 Å, $d_{N\text{-}O}$ = 2.4–3.3 Å. For hydrogen bonds formed to a sulfur atom as acceptor, the upper distance limits were increased by 0.7 Å.

In the calculations, the metal center was represented by a Zn^{2+}-ion tetrahedrally coordinated by the four sulfurs of cysteines 10, 13, 43, and 46. The bond length (Zn-S = 2.35 Å) and angles (S-Zn-S = 109.5°; Zn-S-C^β = 100°) were chosen in analogy to the crystal structure of the Zn-substituted form of *C. pasteurianum* rubredoxin (Dauter et al., 1996).

Structure calculation and analysis

All structures were calculated using X-PLOR 3.851 (Brünger, 1993) by ab initio simulated annealing (Kharrat et al., 1995; Kemmink et al., 1996) including floating assignment of prochiral groups and a reduced presentation for nonbonded interactions for part of the calculation. Each round of the structure calculation started from templates with random backbone torsion angles. In the conformational search phase, 120 ps of MD were simulated at 2,000 K (2 fs time step) computing nonbonded interactions only between C_α atoms and one carbon of each side chain using van der Waals radii of 2.25 Å (Nilges, 1993) to increase efficiency. The subsequent procedure accomplished cooling from 2,000 to 1,000 K within 90 ps (1 fs time step) concomitantly increasing the force constants for the nonbonded interactions to their final values. In the final stage of the calculation, the system was cooled from 1,000 to 100 K within 45 ps (1 fs time step), applying the high force constants obtained at the end of the previous cooling stage, followed by 200 steps of energy minimization.

Of the 60 structures resulting from the final round of structure calculation, those 21 structures that showed the lowest energy and the least number of violations of the experimental data were selected for further characterization. Geometry of the structures, structural parameters, and elements of secondary structure were analyzed using the programs DSSP (Kabsch & Sander, 1983), PROCHECK (Laskowski et al., 1993), and PROMOTIF (Hutchinson & Thornton, 1996). For the graphical presentation of the structures RasMol (Sayle, 1995), SYBYL (Tripos, St. Louis, Missouri), MOLSCRIPT (Kraulis, 1991), and Raster3D (Merritt & Murphy, 1994) were used. The coordinates have been deposited in the Protein Data Bank, Brookhaven National Laboratory (Upton, New York) with accession code 1DX8 and chemical shifts have been deposited at the BioMagResBank, University of Wisconsin, with accession code 4382.

References

Adman ET, Sieker LC, Jensen LH. 1991. Structure of rubredoxin from *Desulfovibrio vulgaris* at 1.5 Å resolution. *J Mol Biol 217*:337–352.

Barkhuijsen H, de Beer R, Bovee WMMJ, van Ormondt D. 1985. Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure. *J Magn Res 61*:465–481.

Barton GJ. 1993. ALSCRIPT: A tool to format multiple sequence alignments. *Protein Eng 6*:37–40.

Bau R, Rees DC, Kurtz DM, Scott RA, Huang H, Adams MWW, Eidsness MK. 1998. Crystal structure of rubredoxin from *Pyrococcus furiosus* at 0.95 Å resolution, and the structures of the N-terminal methionine and formylmethionine variants of Pf Rd. Contributions of N-terminal interactions to thermostability. *J Biol Inorg Chem 3*:484–493.

Bax A, Davis DG. 1985. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J Magn Res 65*:355–360.

Beißinger M, Sticht H, Sutter M, Ejchart A, Haehnel W, Rösch P. 1998. Solution structure of cytochrome c6 from the thermophilic cyanobacterium *Synechococcus elongatus*. *EMBO J 17*:27–36.

Bertini I, Kurtz DM Jr, Eidsness MK, Liu G, Luchinat C, Rosato A, Scott RA. 1998. Solution structure of reduced *Clostridium pasteurianum* rubredoxin. *J Biol Inorg Chem 3*:401–410.

Blake PR, Park JB, Zhou ZH, Hare DR, Adams MW, Summers MF. 1992. Solution-state structure by NMR of zinc-substituted rubredoxin from the marine hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Protein Sci 1*:1508–1521.

Brünger AT. 1993. *X-PLOR version 3.1*. New Haven, Connecticut: Howard Hughes Medical Institute & Yale University.

Cavanagh J, Rance M. 1992. Suppression of cross relaxation effects in TOCSY spectra via a modified DIPSI-2 mixing sequence. *J Magn Res 96*:670–678.

Chen JC, Mortensen LE. 1992. Two open reading frames (ORFs) identified near the hydrogenase structural genes in *Azotobacter vinelandii*, the first ORF may encode for a polypeptide similar to rubredoxins. *Biochim Biophys Acta 1131*:122–124.

Cordier F, Grzesiek S. 1999. Direct observation of hydrogen bonds in proteins by interresidue $^{3h}J(NC')$ scalar couplings. *J Am Chem Soc 121*:1601–1602.

Cornilescu G, Hu JS, Bax A. 1999a. Identification of the hydrogen bonding network in a protein by scalar couplings. *J Am Chem Soc 121*:2949–2950.

Cornilescu G, Ramirez BE, Frank MK, Clore GM, Gronenborn AM, Bax A. 1999b. Correlation between $^{3h}J(NC')$ and hydrogen bond length in proteins. *J Am Chem Soc 121*:6275–6279.

Dauter Z, Sieker LC, Wilson KS. 1992. Refinement of rubredoxin from *Desulfovibrio vulgaris* at 1.0 A with and without restraints. *Acta Crystallogr B48*:42–59.

Dauter Z, Wilson KS, Sieker LC, Moulis JM, Meyer J. 1996. Zinc- and iron-rubredoxins from *Clostridium pasteurianum* at atomic resolution: A high-precision model of a ZnS_4 coordination unit in a protein. *Proc Natl Acad Sci USA 93*:8836–8840.

Day MW, Hsu BT, Joshua-Tor L, Park JB, Zhou ZH, Adams MW, Rees DC. 1992. X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Protein Sci 1*:1494–1507.

Dayie KT, Wagner G. 1994. Relaxation-rate measurements for ^{15}N-^{1}H groups with pulsed field gradients and preservation of coherence pathways. *J Magn Res 111A*:121–126.

Düx P, Whitehead B, Boelens R, Kaptein R, Vuister GW. 1997. Measurements of ^{15}N-^{1}H couplings constants in uniformly ^{15}N-labeled proteins: Application to the photoactive yellow protein. *J Biomol NMR 10*:301–306.

Emsley L, Bodenhausen G. 1990. Gaussian pulse cascades—New analytical functions for rectangular selective inversion and in-phase excitation in NMR. *Chem Phys Lett 165*:469–476.

Eidsness MK, Burden AE, Richie KA, Kurtz DM, Scott RA, Smith ET, Ichiye T, Beard B, Min T, Kang C. 1999. Modulation of the redox potential of the $[Fe(SCys)_4]$ site in rubredoxin by the orientation of a peptide dipole. *Biochemistry 38*:14803–14809.

Frey M, Sieker L, Payan F, Haser R, Bruschi M, Pepe G, LeGall J. 1987. Rubredoxin from *Desulfovibrio gigas*. A molecular model of the oxidized form at 1.4 Å resolution. *J Mol Biol 197*:525–541.

Friedrich MS. 1995. A model-free algorithm for the removal of baseline artifacts. *J Biomol NMR 5*:147–153.

Gehring K, Ekiel I. 1998. H(C)CH-COSY and (H)CCH-COSY experiments for ^{13}C-labeled proteins in H_2O solution. *J Magn Res 135*:185–193.

Geissdorfer W, Frosch SC, Haspel H, Ehrt S, Hillen W. 1995. Two genes encoding proteins with similarities to rubredoxin and rubredoxin reductase are required for conversion of dodecane to lauric acid in *Acinetobacter calcoaceticus* ADP1. *Microbiology 141*:1425–1432.

Gilson PR, Maier U-G, McFadden GI. 1997. Size isn't everything: Lessons in genetic miniaturisation from nucleomorphs. *Curr Opin Genet Dev 7*:800–806.

Gomes CM, Silva G, Oliveira S, LeGall J, Liu MY, Xavier AV, Rodrigues-Pousada C, Teixeira M. 1997. Studies on the redox centers of the terminal oxidase from *Desulfovibrio gigas* and evidence for its interaction with rubredoxin. *J Biol Chem 272*:22502–22508.

Grzesiek S, Bax A. 1992a. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Res 96*:432–440.

Grzesiek S, Bax A. 1992b. Correlating backbone amide and sidechain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc 114*:6291–6293.

Grzesiek S, Bax A. 1993a. Amino-acid type determination in the sequential assignment procedure uf uniformly C13/N15 enriched proteins. *J Biomol NMR 3*:185–204.

Grzesiek S, Bax A. 1993b. The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements. *J Am Chem Soc 115*:12593–12594.

Grzesiek S, Anglister J, Bax A. 1993. Correlation of backbone amide and aliphatic side-chain resonances in $^{13}C/^{15}N$-enriched proteins by isotropic mixing of ^{13}C magnetization. *J Mag Reson 101B*:114–119.

Grodberg J, Dunn JJ. 1989. Comparison of *Escherichia coli* K-12 outer membrane protease OmpT and *Salmonella typhimurium* E protein. *J Bacteriol 170*:1245–1253.

Higgins DG, Bleasby AJ, Fuchs R. 1992. CLUSTAL V: Improved software for multiple sequence alignment. *Comput Appl Biosci 8*:189–191.

Hutchinson EG, Thornton JM. 1996. PROMOTIF—A program to identify and analyze structural motifs in proteins. *Protein Sci 5*:212–220.

Hwang TL, Shaka AJ. 1995. Water suppression that works: Excitation sculpting using arbitrary waveforms and pulsed field gradients. *J Magn Res 112A*: 275–279.

Ikura M, Kay LE, Bax A. 1991. Improved three-dimensional ^{1}H-^{13}C-^{1}H correlation spectroscopy of a ^{13}C-labeled protein using constant-time evolution. *J Biomol NMR 1*:299–304.

Jeener J, Meier BH, Bachmann P, Ernst RR. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Chem Phys 71*:4546–4553.

Johnson BA, Blevins RA. 1994. NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR 4*:603–614.

Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*:2577–2637.

Kay LE, Keifer P, Saarinen T. 1992. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc 114*:10663–10665.

Kay LE, Torchia DA, Bax A. 1989. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry 28*:8972–8979.

Kemmink J, Darby NJ, Dijkstra K, Nilges M, Creighton TE. 1996. Structure determination of the N-terminal thioredoxin-like domain of protein disulfide isomerase using multidimensional heteronuclear $^{13}C/^{15}N$ NMR spectroscopy. *Biochemistry 35*:7684–7691.

Kharrat A, Macias MJ, Gibson TJ, Nilges M, Pastore A. 1995. Structure of the dsRNA binding domain of *E. coli* RNase III. *EMBO J 14*:3572–3584.

Koide S, Jahnke W, Wright PE. 1995. Measurement of intrinsic exchange rates of amide protons in a ^{15}N-labeled peptide. *J Biomol NMR 6*:306–312.

Kraulis P. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Live DH, Davis DG, Agosta WC, Cowburn D. 1984. Long-range hydrogen bond mediated effects in peptides—^{15}N NMR-study of gramicidin-S in water and organic solvents. *J Am Chem Soc 106*:1939–1941.

Marion D, Ikura M, Tschudin R, Bax A. 1989. Rapid recording of 2D NMR spectra without phase cycling. Applications to the study of hydrogen exchange in proteins. *J Magn Res 85*:393–399.

McFadden GI, Gilson PR, Douglas SE, Cavalier-Smith T, Hofmann CJB, Maier U-G. 1997. Bonsai genomics: Sequencing the smallest eukaryotic genomes. *Trends Genet 13*:46–49.

Merritt EA, Murphy MEP. 1994. Raster3D version 2.0—A programm for photorealistic molecular graphics. *Acta Crystallogr D50*:869–873.

Meyer O, Schlegel HG. 1983. Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu Rev Microbiol 37*:277–310.

Misaki S, Morimoto Y, Ogata M, Yagi T, Higuchi Y, Yasuoka N. 1999. Structure determination of rubredoxin from *Desulfovibrio vulgaris* Miyazaki F in two crystal forms. *Acta Crystallogr D55*:408–413.

Mori S, Abeygunawardana C, Berg JM, van Zijl PCM. 1997. NMR study of rapidly exchanging backbone amid protons in *Staphylococcal nuclease* and the correlation with structural and dynamic properties. *J Am Chem Soc 119*:6844–6852.

Mori S, Abeygunawardana C, O'Neill Johnson M, van Zijl PCM. 1995. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J Magn Res 108B*:94–98.

Nilges M. 1993. A calculation strategy for the structure determination of symmetric dimers by 1H NMR. *Proteins 17*:297–309.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1992. *Numerical recipes in C*, 2nd ed. New York: Cambridge University Press.

Rensing SA, Goddemaier M, Hofmann CJB, Maier U-G. 1994. The presence of a nucleomorph hsp70 gene is a common feature of *Cryptophyta* and *Chlorarachniophyta*. *Curr Genet 26*:451–455.

Rost B, Sander C. 1993. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol 232*:584–599.

Sambrook J, Fritsch EF, Maniatis M. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press.

Sayle R. 1995. *RasMol V2.6. molecular visualisation program*. Stevewnage, Hertfordshire, United Kingdom: Glaxo Wellcome Research and Development.

Schleucher J, Sattler M, Griesinger C. 1993. Coherence selection by gradients without signal attenuation: Application to the three-dimensional HNCO experiment. *Angew Chem Int Ed Engl 32*:1489–1491.

Seki S, Ikeda A, Ishimoto M. 1988. Rubredoxin as an intermediary electron carrier for nitrate reduction by NAD(P)H in *Clostridium perfringens. J Biochem 103*:583–584.

Shaka AJ, Keeler J, Frenkiel T, Freeman R. 1983. An improved sequence for broadband decoupling: Waltz-16. *J Magn Res 53*:335–338.

Shaka A, Barker PB, Freeman R. 1985. Computer-optimized decoupling scheme for wideband applications and low-level operation. *J Magn Res 64*:547–552.

Sieker LC, Stenkamp RE, Le Gall J. 1994. Rubredoxin in crystalline state. *Methods Enzymol 243*:203–216.

Sklenar V, Piotto M, Leppik R, Saudek V. 1993. Gradient-tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity. *J Magn Res 102A*:241–245.

Stenkamp RE, Sieker LC, Jensen LH. 1990. The structure of rubredoxin from *Desulfovibrio desulfuricans* strain 27774 at 1.5 Å resolution. *Proteins 8*:352–364.

Stewart DE, Legall J, Moura I, Moura JJ, Peck HD, Xavier AV, Weiner PK, Wampler JE. 1989. Electron transport in sulfate-reducing bacteria. Molecular modeling and NMR studies of the rubredoxin–tetraheme-cytochrome-c3 complex. *Eur J Biochem 185*:695–700.

Sticht H, Rösch P. 1998. The structure of iron-sulfur proteins. *Prog Biophys Mol Biol 70*:95–136.

Talluri S, Wagner G. 1996. An optimized 3D NOESY-HSQC. *J Magn Res 112B*:200–205.

Vuister GW, Bax A. 1992. Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J Mag Reson 98*:428–435.

Vuister GW, Bax A. 1993. Quantitative J correlation: A new approach for measuring homonuclear three-bound $J(H^{N}\text{-}H^{\alpha})$ coupling constants in ^{15}N-enriched proteins. *J Am Chem Soc 115*:7772–7777.

Watenpaugh KD, Sieker LC, Jensen LH. 1979. The structure of rubredoxin at 1.2 Å resolution. *J Mol Biol 217*:509–522.

Wittekind M, Mueller L. 1993. HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with alpha- and beta-carbon resonances in proteins. *J Magn Res 101B*:201–205.

Zhang W, Smithgall T, Gmeiner WH. 1997. Three-dimensional structure of the Hck SH2 domain in solution. *J Biomol NMR 10*:263–272.

Zhu G, Bax A. 1990. Improved linear prediction for truncated signals of known phase. *J Magn Res 90*:405–410.