Bau, R. et al. Crystal structure of rubredoxin from Pyrococcus furiosus at 0.95A resolution, and the structures of N-terminal methionine and formylmethionine variants of Pf Rd. Con...

ARTICLE *in* JBIC JOURNAL OF BIOLOGICAL INORGANIC CHEMISTRY · JANUARY 1998

Impact Factor: 2.54 · DOI: 10.1007/s007750050258

JBIC (1998) 3:484–493

ORIGINAL ARTICLE

Robert Bau · Douglas C. Rees · Donald M. Kurtz, Jr.
Robert A. Scott · Heshu Huang
Michael W.W. Adams · Marly K. Eidsness

Crystal structure of rubredoxin from *Pyrococcus furiosus* at 0.95 Å resolution, and the structures of N-terminal methionine and formylmethionine variants of Pf Rd. Contributions of N-terminal interactions to thermostability

Received: 27 March 1998 / Accepted: 25 June 1998

Abstract The high-resolution crystal structure of the small iron-sulfur protein rubredoxin (Rd) from the hyperthermophilic archeon *Pyrococcus furiosus* (Pf) is reported in this paper, together with those of its methionine ([_0M]Pf Rd) and formylmethionine (f[_0M]Pf Rd) variants. These studies were conducted to assess the consequences of the presence or absence of a salt bridge between the amino terminal nitrogen of Ala1 and the side chain of Glu14 to the structure and stability of this rubredoxin. The structure of wild-type Pf Rd was solved to a resolution of 0.95 Å and refined by full-matrix least-squares techniques to a crystallographic agreement factor of 12.8% [$F>2\sigma(F)$ data, 25 617 reflections], while those of the [_0M]Pf and f[_0M]Pf Rd variants were solved at slightly lower resolutions (1.1 Å, $R=11.5\%$, 17 213 reflections; 1.2 Å, $R=13.7\%$, 12 478 reflections, respectively). The quality of the data was such that about half of the hydrogen atoms of the protein were clearly visible. All three structures were ultimately refined using the program SHELXL-93 with anisotropic atomic displacement parameters for all non-hydrogen protein atoms, and calculated hydrogen positions included but not refined. In this paper we also report thermostability data for all three forms of Pf Rd, and show that they follow the sequence wild-type >[_0M]Pf>formyl[_0M]Pf. Comparison of the three Pf Rd structures in the N-terminal region show that the structures of wild-type Pf Rd and f[_0M]Pf are rather similar, while that of [_0M]Pf Rd shows a number of additional hydrogen bonds involving the extra methionine group. While the salt bridge between the Ala1 amino group and the Glu14 carboxylate is not the primary determinant of the thermostability of Pf Rd, alterations to the amino terminus do have a moderate influence on the thermostability of this protein.

Key words Rubredoxin · Iron-sulfur proteins · Hyperthermostability · Protein structure · Metalloproteins

R. Bau
Department of Chemistry, University of Southern California, Los Angeles, CA 90089, USA

D. C. Rees (✉)
Division of Chemistry and Chemical Engineering, 147-75CH, and the Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, USA

M. K. Eidsness (✉) · D. M. Kurtz, Jr. · R. A. Scott
Department of Chemistry and the Center for Metalloenzyme Studies, University of Georgia, Athens, GA 30602-2556, USA

H. Huang · M. W. W. Adams
Department of Biochemistry and Molecular Biology and the Center for Metalloenzyme Studies, University of Georgia, Athens, GA 30602-7229, USA

Supplementary information available Coordinates and structure factors for the structures reported in this paper will be deposited with the Brookhaven Protein Data Bank, file numbers 1BRF, R1BRFSF; 1BQ8, R1BQ8SF; and 1BQ9, R1BQ9SF for the wild-type, methionine and formyl methionine variants, respectively.

Introduction

Rubredoxins are small (~50 amino acid residue) proteins containing an iron atom coordinated by the sulfur atoms of four cysteine side chains [1]. Although the physiological roles for rubredoxins have not been definitively established, it is likely that they function as electron transfer proteins [2]. Rubredoxin from *Desulfovibrio gigas* has been shown to be the redox partner of rubredoxin-oxygen oxidoreductase, the terminal component of a soluble electron transfer chain [3, 4]. In *Pseudomonas oleovorans*, rubredoxin is part of a three-component system that carries out the oxidation of alkanes to alkanols [5]. Despite the uncertainty of its function in most species, rubredoxins from different organisms have been extensively studied by structural and spectroscopic methodologies, owing to their small size, stability, and ease of isolation.

Of particular relevance, rubredoxin was one of the first proteins to be crystallographically analyzed at high

resolution. Pioneering work by Watenpaugh, Sieker, Jensen and co-workers [6, 7] in 1972 described the structure analysis of rubredoxin from *Clostridium pasteurianum* (Cp Rd). This analysis, carried out at a resolution of 1.5 Å, featured the first report of a least-squares refinement of a protein on an atom-by-atom basis. The structure analysis of Cp Rd was eventually extended to 1.2 Å [8] and to 1.1 Å resolution [9], and other rubredoxin structures that have been reported over the years include those from *D. gigas* (Dg Rd, 1.4 Å resolution [10]), *D. desulfuricans* (Dd Rd, 1.5 Å [11, 12]), and *D. vulgaris* (Dv Rd, 1.5 Å [13, 14]). Recently the refinement of Dv Rd was extended to 1.0 Å resolution [15], and that data set was used in a paper describing the application of direct methods to protein crystallography [16].

A few years ago we undertook the structure determination of the first rubredoxin isolated from a hyperthermophile, *Pyrococcus furiosus* (Pf Rd), in an attempt to understand the extraordinary stability of proteins from high-temperature organisms [17]. In that study, which was carried out at 1.8 Å resolution, it was suggested that the stability of Pf Rd was in part due to multiple electrostatic interactions near the N-terminal end of the protein, involving the amino acids Glu14, Ala1, Trp3, and Phe29. In this paper we extend our studies of Pf Rd to a higher resolution of 0.95 Å, which represents the physical upper limit of our diffraction apparatus. We also report in this paper the isolation and high-resolution structures (1.1 Å and 1.2 Å, respectively) of the methionine ([_0M]Pf Rd) and formylmethionine (f[_0M]Pf Rd) variants of Pf Rd.[1]

The isolation, purification, and crystallization of wild-type Pf Rd have been described [17]. In this study we include [_0M]Pf and f[_0M]Pf Rds, which are products of a synthetic Pf Rd gene overexpressed in *Escherichia coli* [18]. The N-terminally modified proteins are apparently the result of incomplete processing of overexpressed Pf Rd by the deformylase and methionylaminopeptidase of *E. coli* [19–22]. Curiously, in studies of chimeric Rds composed of Pf Rd and Cp Rd amino acid sequences, the proportion of N-terminally modified proteins overexpressed seems to correlate with the thermostability of the Rds; those that are of hyperthermophilic (Pf Rd-like) stability produce a mixture of N-terminally modified forms, and those that are mesophilic (Cp Rd-like) produce predominantly the wild-type (N-terminally processed) Rd [18]. The purification of these three forms of Pf Rds entails separation of proteins that differ by a single amino acid residue or by a formyl group. Anion-exchange chromatography does not yield complete separation, but subsequent passage over a hydroxyapatite column delivers base-line resolution of these three forms of Pf Rd [18].

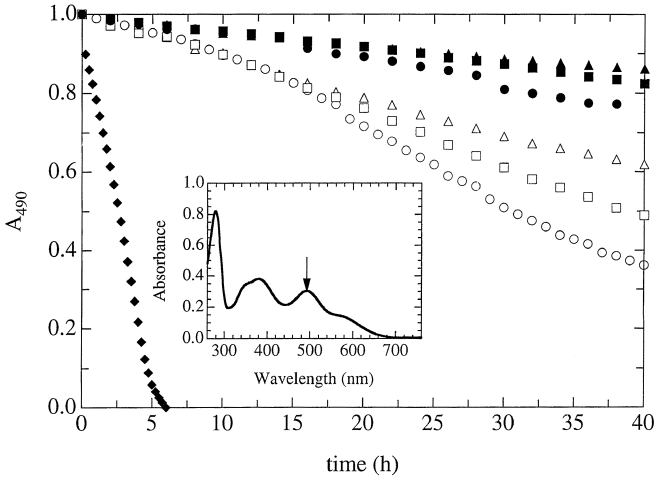

Fig. 1 Thermal stability of Pf Rd. Rd 490 nm absorbance decays vs time at 92 °C: Cp (*filled diamond*), Pf Rd (*filled triangle*), [_0M]Pf Rd (*filled square*), and f[_0M]Pf Rd (*filled circle*) in 50 mM HEPES, pH 8.2 (at room temperature); Pf Rd (*open triangle*), [_0M]Pf Rd (*open square*), and f[_0M]Pf Rd (*open circle*) in 50 mM phosphate, pH 6.3. Rd concentrations were 35 μM. *Inset*: Absorption spectrum of Pf Rd

Oxidized Rd has a distinctive UV-visible absorption spectrum (see Fig. 1, inset) with an intense band at 280 nm and bands at 385 and 490 nm arising from charge transfer transitions from cysteinyl thiolate to Fe(III) [23, 24]. The UV-visible spectral features of the N-terminally modified Pf Rds are indistinguishable from those of the wild-type Pf Rd. In addition, the reduction potentials of these Rds are identical within experimental error (E. T. Smith, M. K. Eidsness, R. A. Scott, D. M. Kurtz, Jr., unpublished results). They differ, however, in their thermostabilities. The thermostabilities of Pf Rds were assessed by monitoring the destruction of the native iron site by measuring the fractional loss of 490 nm absorbance upon heating at 92 °C (Fig. 1) [18]. Although this protocol does not directly measure the thermodynamic stability, it does provide an indication of the relative rates with which the native structure denatures at elevated temperatures. Thermostability data from the mesophilic Rd from *C. pasteurianum* (Cp Rd) are included to illustrate how extraordinarily stable the Pf Rds are by comparison. Among the Pf Rds, the order of thermostability is consistently wild-type >[_0M]Pf>f[_0M] Pf Rds when heated at 92 °C in a number of different buffers (vide infra). We have determined the high-resolution structures of the wild-type, [_0M]Pf, and f[_0M] Pf Rds and have analyzed the hydrogen bonding and salt bridge patterns to provide insights into the origins of the observed differences in thermostability.

[1] We use the terms [_0M]Pf Rd and f[_0M]Pf Rd to denote the variants of Pf Rd in which methionine and formylmethionine respectively have been added to the N-terminus. These correspond to '[_1M]Pf Rd' and 'f[_1M]Pf Rd' respectively in an earlier manuscript of ours [18]. We label the Met and fMet residues as position '0' instead of position –1 in order to maintain consistency with the numbering system of wild-type Pf Rd (e.g., so that Ala1 in wild-type Pf Rd can still be referred to as Ala1 in the Met and fMet variants)

Methods

Overexpression and purification of Pf Rds

The construction of the Pf Rd gene, its overexpression in *E. coli*, and the isolation of crude Pf Rds are described elsewhere [18]. Overexpression of Rd genes in *E. coli* results in Fe (red) and Zn (colorless) forms [25–27]. For a preparative scale purification of the Fe forms of wild-type, [_0M]Pf Rd, and formyl[_0M]Pf Rds, a three-step chromatographic procedure was followed.

1. Mono-Q HR10/10 column (Pharmacia Biotech); buffer 1 = 25 mM Tris-HCl, pH 8.5, and buffer 2 = buffer 1 + 1.0 M NaCl; elution gradient = 0–10% buffer 2 in 10 mL followed by 10–30% buffer 2 in 75 mL; flow rate = 2 mL/min.

2. Identical to step 1 except the elution gradient was 0–20% buffer 2 in 10 mL followed by 20–30% buffer 2 in 75 mL.

3. Hydroxyapatite CHT1 (10 mL) column (Bio-Rad); buffer 1 = 5 mM sodium phosphate, pH 6.0, buffer 2 = 0.5 M sodium phosphate, pH 6.0; elution gradient = 0–15% buffer 2 in 50 mL followed by 15–25% buffer 2 in 150 mL; flow rate = 4 mL/min.

At each chromatographic step, 1 min fractions of each peak were collected. Appropriate fractions were pooled, desalted, and buffer exchanged by ultrafiltration with a YM3 membrane (Amicon).

Characterization of Pf Rds

UV-visible absorption spectra were recorded on a Shimadzu UV-2101PC scanning spectrophotometer, using 1 cm path length quartz cuvettes. The Fe-containing Pf Rds were characterized by their UV-visible absorption spectra and their concentrations determined by 490 nm absorbance ($\varepsilon 490 = 9220\ M^{-1}\ cm^{-1}$) [28]. Molecular weights of the Pf Rds were determined by electrospray ionization mass spectrometry (at the Chemical and Biological Sciences Mass Spectrometry facility of the University of Georgia by D. Phillips). The molecular weight for wild-type Pf Rd is 5949, [_0M]Pf Rd is 6083, and f[_0M]Pf Rd is 6109.[2]

Thermostability measurements

Pf Rds were buffer-exchanged into water and their concentrations determined by 490 nm absorbance. The thermostability apparatus and analysis of the UV-visible absorption time courses at 92 °C have been described [18].

Crystallization

Rubredoxin from *P. furiosus* was supplied as concentrated solutions containing 40 mg/mL protein (ca. 6 mM Rd), dissolved in 50 mM Tris/Tris-HCl buffer (pH 8.0) and 0.3 M NaCl. Vapor diffusion of this solution against 3.6 M NaK phosphate (equimolar $NaH_2PO_4 + K_2HPO_4$) as precipitant at room temperature yielded small prismatic crystals in about 3 days, which were crushed and used as seeds. Crystals suitable for diffraction work were obtained by macroseeding [29] of hanging drops, which had been pre-equilibrated at 4 °C against either 3.0 M NaK phosphate or 2.8 M ammonium sulfate in 0.1 M Tris buffer (pH 8.5). After the introduction of seeds, brick-like crystals suitable for diffraction work appeared overnight and were allowed to grow for several days, some of which reached dimensions as large as $0.6 \times 0.7 \times 0.8$ mm.

The [_0M]Pf and f[_0M]Pf Rd variants of Pf Rd were also supplied as concentrated solutions (30 mg/mL and 20 mg/mL, respectively, in phosphate buffer), and suitable crystals were grown at 4 °C with essentially the same procedure, using tiny crystals of wild-type Pf Rd as seeds.

Data collection

X-ray diffraction data were collected at –150 °C on a Siemens rotating-anode generator equipped with an X-1000 area detector (Siemens Analytical Instruments, Madison, Wis., USA). Crystals were harvested without the use of any added cryo-solvent, rapidly covered with a layer of Paratone-N oil, and then flash-frozen [30, 31] in a cold gaseous nitrogen stream. Data were collected with Cu $K\alpha$ radiation up to maximum 2θ (detector) settings of 80°, 60°, and 50° for crystals of wild-type Pf Rd, [_0M]Pf, and f[_0 M]Pf Rd variants, respectively. In all three cases, only one crystal was used for the entire data set. Data were processed using the XDS/XSCALE data processing package [32], and final statistics are given in Table 1.

Structure analysis and refinement

Wild-type Pf Rd

The atomic positions derived from the previous 1.8 Å resolution analysis [17] of Pf Rd were used as the starting point of the refinement, which was carried out using the program SHELXL-93 [33]. In this procedure, geometrical restraints are imposed during least-squares refinement whereby the 1,2 and 1,3 distances between protein atoms (corresponding to interatomic bond lengths and bond angles, respectively) are kept within 0.03 Å of ideal values, and distances involving water molecules are kept within 0.1 Å of minimum allowable contact values. A water molecule is considered present if it has electron density greater than 1.0 e/Å^3 and if it is no closer than 3.2 Å, 2.7 Å, and 2.6 Å from existing C, N, and O atoms, respectively. 'Anti-bumping' restraints were imposed during least-squares refinement to keep the water molecules separated by at least these minimum distances. Other constraints involving the planarity of peptide linkages and of aromatic rings were also enforced.

Refinement with isotropic atomic displacement factors for all atoms (including 119 water molecules) resulted in an agreement factor of $R = 17.4\%$. At this stage, much of the remaining electron density occurred in the solvent regions, but these peaks were too close to existing water molecules. It was thus concluded that some solvent regions were disordered, especially those remote from the protein, and so these regions were modeled on the basis of two overlapping sets of partially occupied water molecules.[3] Further refinement cycles then involved the following features: (1) all protein atoms were allowed to refine anisotropically; (2) two sets of

[2] The sequence-derived molecular weights of the apo (iron-free) forms of wild-type and [_0M]Pf Rds (5896 and 6027Da, respectively) were calculated by the program peptidesort (Program Manual for the Wisconsin Package, version 8, September 1994; Genetics Computer Group, 575 Science Drive, Madison WI 53711, USA). The molecular weight of apo formyl [_0M]Pf Rd (6055) was determined by adjusting the [_0M]Pf Rd molecular weight by subtracting one proton and adding a formyl group (HCO). The molecular weights of the holo Rds were then determined by adding the mass of Fe (56) and subtracting 4 protons (1.0) from the cysteines to the apo Rd molecular weights. The experimental and calculated Rd molecular weights agree to within 0.1% of the Rd mass, which is the error typically associated with electrospray ionization mass spectrometry

[3] Each set of partially occupied water molecules was internally self-consistent, in the sense that the water molecules were kept at least 2.6 Å apart from each other using 'anti-bumping' restraints. Of course, no contraints were imposed on water molecules *between* the two sets since they are not expected to exist simultaneously in the same region of space

Table 1 Summary of data collection statistics

	Wild-type Pf Rd	[_0M] Pf Rd	f[_0M] Pf Rd
Unit cell parameters			
a (Å)	34.123(1)	33.823(1)	34.029(1)
b (Å)	34.874(1)	34.705(1)	34.474(1)
c (Å)	43.683(1)	43.204(1)	43.470(1)
Detector settings (2θ)	20°, 50°, 80°	20°, 60°	20°, 50°
Max. resolution	0.95 Å	1.1 Å	1.2 Å
No. of measurements	76664	45654	87169
Independent reflections	32303	20698	15400
Overall completeness	98.8%	96.2%	97.5%
R(merge) (for *I*)	0.037	0.056	0.060

High resolution limit (Å)	Number of observed reflections	Number of unique reflections	*R*(merge)	Completeness of data
Statistics for wild-type Pf Rd				
4.00	1037	497	2.3%	96.5%
3.00	2359	663	2.9%	99.7%
2.50	2902	799	3.7%	99.8%
2.25	2468	705	3.7%	99.9%
2.05	1689	810	3.4%	98.4%
1.85	2025	1190	4.0%	97.3%
1.65	4690	1865	4.2%	99.3%
1.45	10079	3004	4.8%	99.8%
1.25	17233	5224	6.2%	99.7%
1.15	8995	4122	5.8%	99.8%
1.05	10962	5796	6.0%	99.7%
0.95	12225	7628	9.7%	89.6%
Statistics for [_0M] Pf Rd (methionine variant)				
4.00	952	502	3.1%	97.5%
3.00	1441	648	3.5%	97.4%
2.50	2908	796	5.0%	99.4%
2.20	3381	879	5.2%	99.8%
2.00	3558	911	5.4%	99.9%
1.80	2774	1305	5.7%	96.9%
1.60	4358	2052	6.8%	97.4%
1.40	7250	3407	7.6%	99.0%
1.20	11935	5952	9.1%	97.7%
1.10	7097	4246	12.5%	87.2%
Statistics for f[_0M] Pf Rd (formylmethionine variant)				
4.00	5199	512	5.0%	99.4%
3.00	8193	665	5.1%	100.0%
2.50	10508	801	5.4%	100.0%
2.20	10938	881	6.3%	100.0%
2.00	10204	912	7.8%	100.0%
1.80	5712	1307	7.9%	97.0%
1.60	8554	2014	9.0%	95.6%
1.40	12563	3206	10.9%	93.1%
1.20	15289	5102	16.6%	83.8%

disordered, overlapping water molecules with partial occupancies were introduced; and (3) calculated hydrogen positions in the protein were added and kept in a riding mode during the subsequent refinement. Exhaustive refinement under these conditions, with many cycles of conjugate-gradient least-squares refinement [34] followed by several cycles of full-matrix least-squares refinement, reduced the *R* factor to a final value of 12.8% for 25 617 reflections. The final model included 88 fully occupied and 98 partially occupied solvent molecules, giving a net total of $88+98/2=137$ water molecules. For reference, it is estimated that $\sim$185 water molecules per rubredoxin molecule are present in the crystal, based on the unit cell volume, the protein volume, and the volume/water molecule (30 $Å^3$). Various refinement statistics are given in Table 2.

The final rubredoxin structure is extremely well-ordered, especially the hydrophobic region in the center of the molecule. As reported earlier [17], a few of the surface residues (such as Glu30, Glu31, Glu49, and Glu52) had high atomic displacement factors and showed evidence of disorder, but the structure could be refined successfully just using the major conformations of the disordered groups.[4]

[4] In some cases, difference-Fourier maps suggested the presence of a major conformation and a minor conformation, but the minor conformation could not be refined successfully because of low occupancy factors

Table 2 Summary of refinement statistics

	Wild-type Pf Rd	[_0M] Pf Rd	f[_0M] Pf Rd
Resolution limit (Å)[a]	0.95	1.1	1.2
R factor [$I>4\sigma(I)$ data]	12.8%	11.5%	13.3%
Number of reflections [$I>4\sigma(I)$ data]	25617	17213	12478
R factor[b] (all data)	13.2%	11.9%	13.7%
Number of reflections (all data)	28498	18377	14030
R(free)	16.2%	14.9%	16.6%
Number of reflections [for *R*(free)]	3165	2041	1559
Number of parameters	4433	4541	4664
Number of restraints	4783	4888	5076
No. of fully occupied water molecules	88	130	81
No. of partially occupied water molecules	98	60	78
Net total number of water molecules	137	160	120
Average ESD in bond distances	0.014 Å	0.016 Å	0.022 Å
Averages ESD in bond angles	0.9°	1.2°	1.7°

[a] Data with resolution lower than 10 Å were not used in the refinement
[b] All *R* factors in this table are based on *F*

Met variant ([_0M]Pf Rd)

The structure analysis and refinement of the N-terminal methionine variant was carried out in essentially the same manner as that of the wild-type protein. Initial isotropic refinement using SHELXL-93 [33], followed by a difference Fourier calculation with XPLOR [35] and map visualization with the program O [36], readily revealed the position of the extra Met residue at the N terminus. It was also noticed that one of the side chains, Lys2, had moved away from its position in the wild-type protein. Inclusion of these atoms and exhaustive least-squares refinement, using the same stepwise procedure used for wild-type Pf Rd, yielded an excellent fit [$R(F)=11.5\%$ for 17 213 reflections].

The quality of the maps for this crystal (the methionine variant) was in fact the best among the three structures described here, revealing the positions of many of the hydrogen atoms in the structure (vide infra). The high quality of the maps also enabled many more solvent molecules to be located, and, in the final model, 130 fully occupied and 60 partially occupied water molecules were refined, giving a net total of $130+60/2=160$ solvent molecules.

The methionine residue is situated at a site normally occupied by solvent molecules in the wild-type Pf Rd structure. Additional hydrogen bonds were found involving the backbone CO and NH groups of this Met group, as described later. As in the case of the wild-type Pf Rd structure, disorder problems involving hydrophilic surface residues were not particularly severe, and the structure could be refined successfully without taking into account multiple conformations of side chains.

fMet variant (f[_0M]Pf Rd)

The structure analysis of the formylmethionine variant was carried out in exactly the same manner as that of the methionine variant. Location of the f[_0M]Pf Rd atoms, followed by least-squares refinement, yielded a final agreement factor of $R(F)=13.3\%$ for 12 478 reflections. The final model included 81 fully occupied and 78 partially occupied water molecules, yielding a net total of $81+78/2=120$ water molecules. Table 2 provides a summary of refinement details of all three structure analyses.

In contrast to those of wild-type Pf Rd and the Met variant, the structure analysis of the fMet variant was complicated by pronounced packing disorder involving several side chains. Portions of Lys6, Glu31, Asp35, and Glu49, as well as the terminal atoms of residues Ile7, Ile40, and Lys50, were disordered and had to be refined in two orientations with partial occupancies.

Results

Descriptions of the structures

The structure of wild-type Pf Rd at high resolution is, as expected, essentially the same as that reported earlier at 1.8 Å resolution [17]. The structure consists of a three-stranded antiparallel β-sheet (Fig. 2), with a hydrophobic core containing six aromatic residues. The molecule is approximately ellipsoidal in overall shape, with a hydrophilic 'tail' (consisting of the residues Lys51, Leu52, and Glu53) that protrudes into the solvent at the C-terminus region. At the present resolution (0.95 Å), almost every non-H atom can be resolved, with the exception of a few poorly defined surface residues.

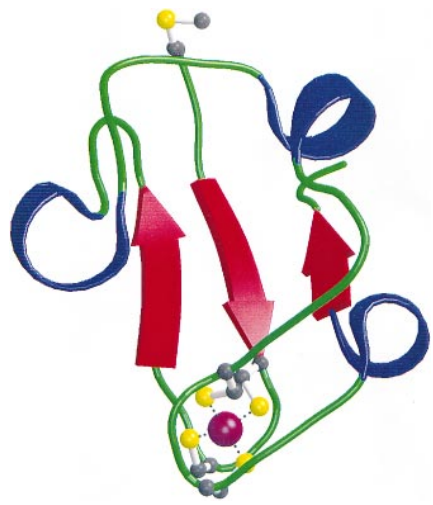

Fig. 2 Schematic diagram of the structure of the methionine variant ([_0M]Pf Rd) showing, among other things, the three-stranded antiparallel β-sheet. The FeS_4 core is shown at the *bottom* of this plot, and the methionine residue in the N-terminus appears at the *top*

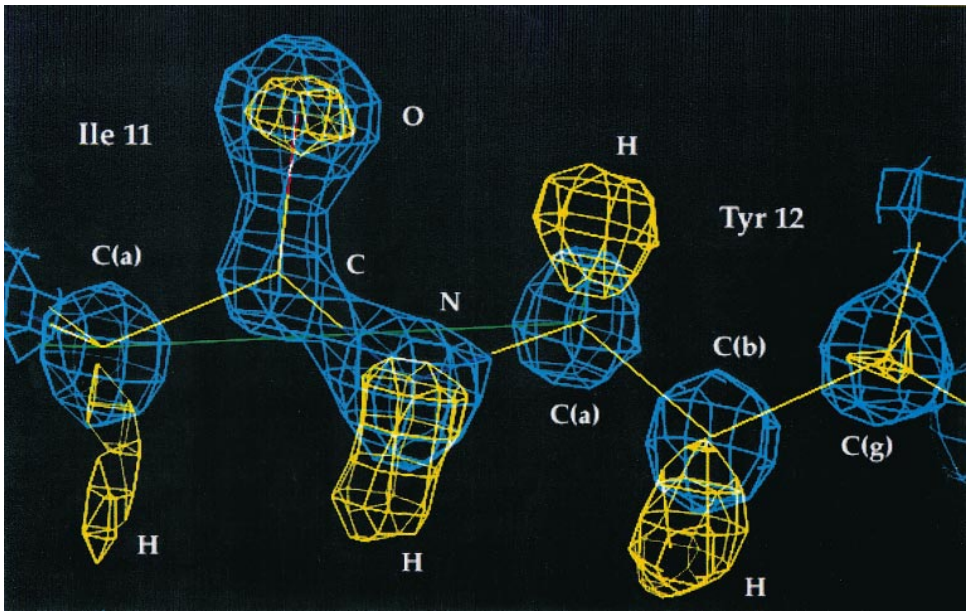

Fig. 3 Difference map, calculated without input hydrogen positions, showing some H atoms in the region near residues Ile11 and Tyr12. *Blue contours* (3σ level) correspond to a ($2F_o$–F_c) map and *yellow contours* (2σ level) represent a (F_o–F_c) map

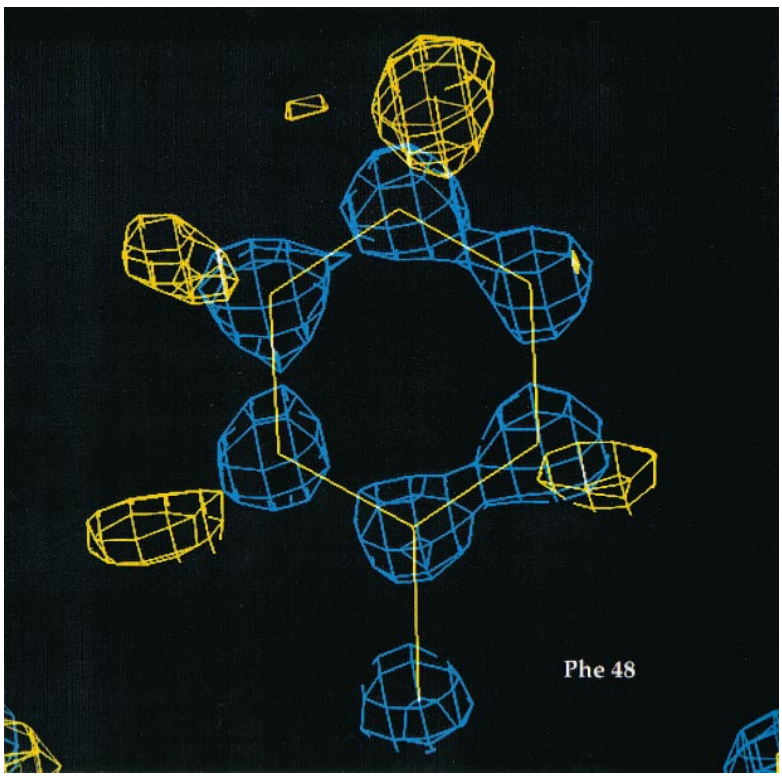

Fig. 4 Same as Fig. 3, showing four of the five aromatic hydrogen atoms of residue Phe48

Full-matrix least-squares refinement of the structure resulted in average estimated standard deviations [33] of 0.014 Å for bond distances and 0.9° for bond angles, among the lowest reported for a protein structure.[5] As expected, the water structure in the immediate vicinity of the protein is well ordered, but it becomes more disordered as one moves further from the surface of the protein molecule. In certain regions, five-membered [38] and six-membered rings of water molecules can be clearly discerned, but the presence of these rings is by no means a pervasive phenomenon.

About half of the hydrogen atoms could be clearly seen in the difference maps. Figure 3 shows a typical region near a section of the main backbone of the Met variant, while Fig. 4 shows one of the phenylalanine re-

[5] For comparison, we note that average ESDs of σ(bond) = 0.0126 Å and σ(angle) = 0.89° were reported by Teeter and co-workers [37] in a high-resolution study (0.83 Å) and full-matrix least-squares refinement of the structure of crambin

sidues in the molecule. Interestingly, the maps were actually cleaner for the 1.1 Å data (Met variant, [_0M]Pf Rd) than for the 0.95 Å wild-type data, which brings to mind a known technique in small-molecule crystallography whereby high-angle data are actually thrown out to improve visualization of H atom positions [39].

Iron-sulfur core

In Table 3 are listed the distances and angles found for the FeS_4 cores of the three structures, which are very similar to each other. Figure 5 shows this FeS_4 center. In all three cases, two of the six S-Fe-S angles are significantly smaller than the other four, a feature that has been characteristic of virtually all other high-resolution rubredoxin structures that have been reported thus far [6–17]. The reason for this two-fold distortion remains unexplained, although it is curious to note that the axis of this distortion is not along the iron-surface direction. [The two cysteine residues near the surface of Pf Rd are Cys8 and Cys41, while the two small angles are (Cys5)-Fe-(Cys41) and (Cys8)-Fe-(Cys38)]. We also find pairs of slightly longer (2.28–2.30 Å) and slightly shorter (2.25–2.26 Å) Fe-S distances in all three structures. The shorter Fe-S distances are associated with the two residues near the surface of the protein (Cys8 and Cys41). Finally, we find ten N-H···S hydrogen bonds, two each to Cys8 and Cys41 Sγs and three each to Cys5 and Cys38 Sγs, in all three structures (Table 4). These distortions from T_d symmetry are characteristic of the MS4 center in all known rubredoxin structures [6–17]. For example, these same distortions (including the pattern of N-H···S hydrogen bonds) are found in both the Cp Fe(III) and Zn(II) Rds [9]. Thus the protein, not the specific metal ion, appears to dictate these characteristic structural features, and there is nothing unique about the nature of these features in Pf Rd.

Table 3 Distances and angles around Fe-S_4 Core[a]

	Wild-type Pf Rd	[_0M] Pf Rd	f[_0M] Pf Rd
Fe-(Cys5)	2.283(3)	2.285(4)	2.278(6)
Fe-(Cys8)	2.248(3)	2.261(5)	2.256(6)
Fe-(Cys38)	2.297(3)	2.304(4)	2.297(6)
Fe-(Cys41)	2.259(3)	2.266(4)	2.255(6)
(Cys5)-Fe-(Cys8)	113.4(1)	113.6(2)	113.0(2)
(Cys5)-Fe-(Cys38)	110.7(1)	111.2(2)	111.3(2)
(Cys5)-Fe-(Cys41)	103.8(1)	103.8(2)	104.1(2)
(Cys8)-Fe-(Cys38)	102.4(1)	102.7(2)	102.7(2)
(Cys8)-Fe-(Cys41)	114.3(1)	113.3(2)	113.5(3)
(Cys38)-Fe-(Cys41)	112.6(1)	112.6(2)	112.5(2)

[a] Cys8 and Cys41 are the two cysteine residues close to the surface of the molecule

Table 4 Hydrogen bonds near the N-terminus of Rd Pf

Donor	Acceptor	Wild-type Pf Rd	[_0M] Pf Rd	f[_0M] Pf Rd
N Met0	OD2 Asp15[a]	–	2.75	–
N Ala1	OE1 Glu14[a]	2.78	3.00	2.78
NZ Lys2	O Met0	–	3.00	–
NZ Lys2	OD1 Asp13[a]	–	2.83	–
NZ Lys2	OD2 Asp15[a]	–	2.91	–
N Glu14	O Ala1	3.02	2.94	2.93
N Glu14	OD1 Asp13[a]	3.10	3.10	3.02
N Asp15	OD1 Asp13[a]	2.97	3.00	2.97
N Asp15	O Asp13	3.10	3.06	3.12
N Ala16	O Asp13	2.88	2.84	2.81
N Glu-17	O Asp13	2.98	2.97	2.96
N-H···S hydrogen bonds				
N Lys6	S Cys5	3.71	3.74	3.73
N Ile7	S Cys5	3.60	3.59	3.60
N Cys8	S Cys5	3.69	3.66	3.66
N Gly9	S Cys8	3.59	3.59	3.58
N Tyr10	S Cys8	3.51	3.54	3.51
N Pro39	S Cys38	3.56	3.55	3.58
N Ile40	S Cys38	3.53	3.54	3.52
N Cys41	S Cys38	3.62	3.61	3.59
N Gly42	S Cys41	3.62	3.61	3.62
N Ala43	S Cys41	3.53	3.53	3.53

[a] Salt bridge

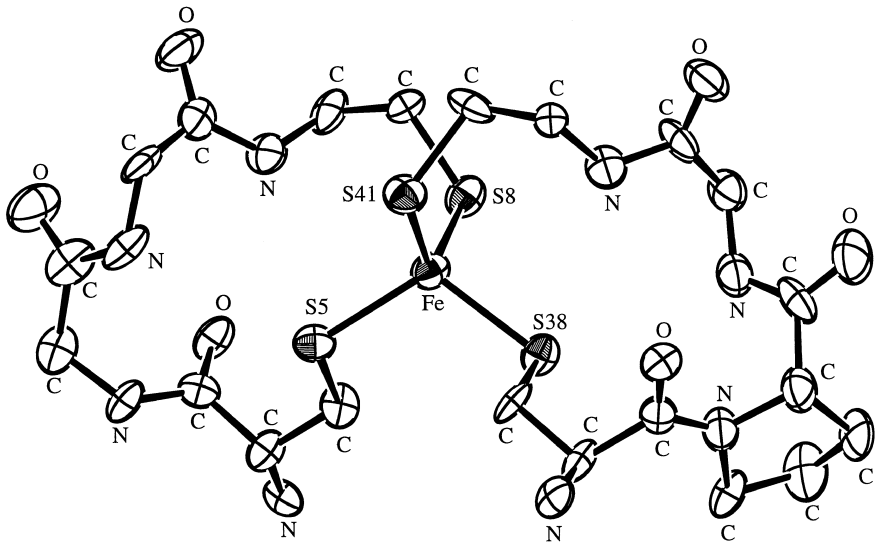

Fig. 5 An ORTEP plot showing the FeS_4 core and surrounding backbone atoms, including the anisotropic thermal ellipsoids. The top of this diagram, including residues Cys8 and Cys41, is near the surface of the protein

Differences between wild-type Pf Rd and the N-terminal variants

Met variant ([_0M]Pf Rd)

The additional methionine residue (labelled ′Met 0′) at the N-terminus of this variant occupies a region of the unit cell normally populated by water molecules, so none of the hydrogen bonds that normally exist in the wild-type Rd structure in this region are disrupted by the introduction of the extra amino acid residue (see Table 4 for a listing of hydrogen bonds that vary in the three structures). The H bonds from N(Ala1) to OE1(Glu14) (a salt bridge in wild-type Pf Rd) and N(Glu14) to O(Ala1) are still present, but the extra Met residue has introduced two extra H-bonding sites on the main backbone, the carbonyl oxygen, O(Met0), and the new amino terminus, N(Met0). O(Met0) is found to H-bond to the side-chain amine group (NZ) of Lys2, which has 'swung around' to form a salt bridge with the side chain of Asp13 (Fig. 6). The terminal amino atom, N(Met0), is found to form a salt bridge to a carboxylate oxygen (OD2) from Asp15, which also 'swivels' slightly to make this hydrogen-bonding contact.

Thus, the introduction of the extra residue, Met0, has caused two side chains that had been protruding into the solvent in wild-type Pf Rd, Lys2 and Asp15, to become involved in H-bonding; in doing so, four new H-bonds or salt bridges have been formed: NZ(Lys2) to O(Met0); NZ(Lys2) to OD2(Asp15); NZ(Lys2) to OD1(Asp13); and N(Met0) to OD2(Asp15) (Fig. 6). Movement of Lys2 involves a 120° pivot about the CG-CD bond and a substantial repositioning of three atoms (CD, CE, and NZ), whereas in the case of Asp15 a smaller rotation of about 60° about the CB-CG bond was necessary to bring one of the carboxylate oxygen atoms into a favorable hydrogen-bonding geometry.

fMet variant (f[_0M]Pf Rd)

Curiously, the new hydrogen bonds shown in Fig. 6 are not found in the formylmethionine variant. Perhaps the steric bulk of the formyl group on the terminal nitrogen atom is interfering with the formation of the N(fMet0) to OD2(Asp15) hydrogen bond. Lys2 and Asp15 in fMet are roughly in the same positions as they are in the wild-type structure (i.e., protruding into the solvent), and the carbonyl oxygen of the terminal fMet group [O(fMet0)] is not oriented properly to hydrogen-bond to any other residue. Thus, in the N-terminal region the structure of the fMet variant more closely resembles that of the wild-type Rd than that of the Met variant.

Conclusions

The relative stabilities of the three title proteins toward thermal degradation decrease in the order (wild-type Pf Rd)>[_0M]Pf Rd>f[_0M]Pf Rd. Nevertheless, all of

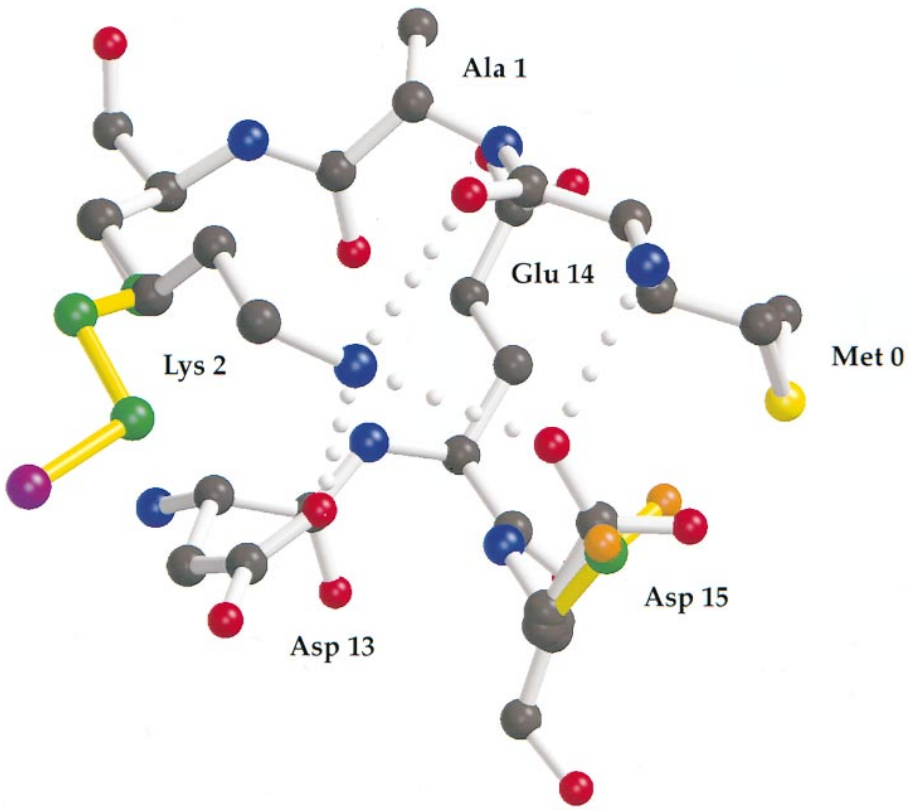

Fig. 6 A composite (superposition) diagram comparing some amino acids near the N-terminal regions of wild-type Pf Rd (*yellow bonds*) and the Met variant (*white bonds*). Four hydrogen bonds which are present in the structure of the Met variant, [_0M]Pf, but not in wild-type Pf Rd, are indicated as *dotted lines*. Note how Lys2 of wild-type Pf Rd (*green and purple atoms*) 'swings over' to form H bonds to three oxygen atoms in the Met variant, and how the carboxylate group of Asp15 (*green and orange atoms*) tilts to form two H bonds (see text)

these proteins remain much more thermostable than Cp Rd at pH 8.2 and 6.3. While these results show that the thermostability of Pf Rd is influenced by the interactions among amino terminal residues, they also demonstrate that the salt bridge between the amino nitrogen of Ala1 and the carboxylate group of Glu14 is not essential for retention of thermostability characteristic of wild-type Pf Rd. A structural comparison reveals that the wild-type and formyl-Met Pf Rds are very similar, except for the addition of the formyl-Met group to the former. The decreased thermostability associated with this modification could in principle be attributed to the replacement of one salt bridge by a hydrogen bond between the Ala1 NH and Glu14 carboxylate groups. However, this simple counting of salt bridges or hydrogen bonds does not explain why [_0M]Pf Rd is slightly less thermostable than the wild-type protein (cf. Fig. 1), despite the larger number of salt bridge/hydrogen bond interactions at the N-terminus of this variant (cf. Fig. 6).[6] As discussed elsewhere [40–42], the energetic contributions of electrostatic interactions to protein stability are quite sensitive to the details of these interactions. The high resolution of the Pf Rd structures described in this paper may permit a more quantitatively reliable analysis of the contributions of the electrostatic (and other types of) interactions to protein thermostability. Nevertheless, from our present results we conclude that neither a larger number of salt-bridge/hydrogen bonds involving N-terminal residues nor differences in the FeS_4 site structure explain Pf Rds extraordinary thermostability. As shown in Fig. 1, all three of the Pf Rds examined here are much more thermostable than Cp Rd, despite the essentially identical FeS_4 structure in all four proteins. We have found no simple correlation between redox potentials of various mutated Rds and their thermostabilities (F. Bonomi, M. K. Eidsness, D. M. Kurtz, Jr., R. A. Scott, E. T. Smith, unpublished results).

Pf Rd joins an increasing number of protein structures that have been determined at resolutions better than 1 Å [43]. At the 0.95 Å resolution of the wild-type protein, the data-to-parameter ratio achieved in this study (better than 6:1) is comparable to that of a typical small-molecule structure determination. In addition to the accurate determination of structural parameters, the ability to successfully locate many of the hydrogen atoms in these structures illustrates that X-ray data can be sufficient to provide useful information about hydrogen positions in macromolecules. It is worth noting that, as illustrated by exceptional cases such as Pf Rd, atomic resolution data on proteins can be achieved with laboratory X-ray sources. Although the coupling between protein stability and structure is complex, the combination of structural information and stability measurements on variants of proteins such as Pf Rd may ultimately yield a more quantitative understanding of the origins of protein hyperthermostability.

[6] The altered conformation of the [_0M]Pf Rd N-terminal region may explain why this variant behaves differently on a hydroxyapatite column (having both hydrophobic and ion exchange sites) compared to wild-type and f[_0M]Pf Rds which nearly co-elute. The stability gained by the additional salt bridges and hydrogen bonds in [_0M]Pf Rd may be offset by a loss of entropy upon removing Lys2 from solvent. The electrostatic penalty for desolvation of a charged side chain is rarely recovered in favorable interactions in the folded protein [40]

Acknowledgements This work was supported by the National Science Foundation (Grant DMB 91-18689 to D.C.R., Grant BCS-96-32657 to M.W.W.A.), the National Institutes of Health (Grant GM-50736 to D.M.K.), and NSF Research Training Group Award DIR-90-14281 to the Center for Metalloenzyme Studies, University of Georgia. R.B. acknowledges sabbatical leave support from the University of Southern California.

References

1. Spiro TG (1982) Iron-sulfur proteins. Wiley-Interscience, New York
2. Sieker LC, Stenkamp RE, Legall J (1994) Methods Enzymol 243:203–216
3. Gomes CM, Silva G, Oliveira S, LeGall J, Liu MY, Xavier AV, Rodrigues-Pousada C, Teixeira M (1997) J Biol Chem, 272:22502–22508
4. Santos H, Fareleira P, Xavier AV, Chen L, Liu MY, LeGall J (1993) Biochem Biophys Res Commun 195:551–557
5. Eggink G, Engel H, Vriend G, Terpstra P, Witholt B (1990) J Mol Biol 212:135–142
6. Watenpaugh KD, Sieker, LC, Herriott, JR, Jensen LH (1972) Cold Spring Harbor Symp Quant Biol 36:359–367
7. Watenpaugh KD, Sieker LC, Herriott JR, Jensen LH (1973) Acta Crystallogr B29:943–956
8. Watenpaugh KD, Sieker LC and Jensen LH (1979) J Mol Biol 131:509–522
9. Dauter Z, Wilson KS, Sieker LC, Moulis JM, Meyer J, (1996) Proc Natl Acad Sci USA 93:8836–8840
10. Frey M, Sieker L, Payan F, Haser R, Bruschi M, Pepe G, LeGall J (1987) J Mol Biol 197:525–541
11. Sieker LC, Stenkamp RE, Jensen LH, Prickril B, LeGall J (1986) FEBS Lett 208:73–76
12. Stenkamp RE, Sieker LC, Jensen LH (1990) Proteins Struct Funct Genet 8:352–364
13. Adman ET, Sieker LC, Jensen LH, Bruschi M, LeGall J (1977) J Mol Biol 112:113–120
14. Adman ET, Sieker LC, Jensen LH (1991) J Mol Biol 217:337–352
15. Dauter Z, Sieker LC, Wilson KS (1992) Acta Crystallogr B48:42–59
16. Sheldrick GM, Dauter Z, Wilson KS, Hope H, Sieker LC (1993) Acta Crystallogr D49:18–23
17. Day MW, Hsu BT, Joshua-Tor L, Park JB, Zhou ZH, Adams MWW, Rees DC (1992) Protein Sci 1:1494–1507
18. Eidsness MK, Richie KA, Burden AE, Kurtz DM Jr, Scott RA (1997) Biochemistry 36:10406–10413
19. Ben-Bassat A, Bauer K, Chang SY, Myambo K, Boosman A, Chang S (1987) J Bacteriol 169:751–757
20. Dalbøge H, Bayne S, Pedersen J (1990) FEBS Lett 266:1–3
21. Hirel PH, Schmitter JM, Dessen P, Fayat G, Blanquet S (1989) Proc Natl Acad Sci USA 86:8247–8251
22. Meinnel T, Mechulam Y, Blanquet S (1993) Biochemie 75:1061–1075
23. Eaton WA, Lovenberg W (1973) In: Lovenberg W (ed) Iron-sulfur proteins. Academic Press, New York, pp 131–162
24. Lowery MD, Guckert JA, Gebhard MS, Solomon EI (1993) J Am Chem Soc 115:3012–3013
25. Richie KA, Teng Q, Elkin CJ, Kurtz DM Jr (1996) Protein Sci 5:883–894

26. Eidsness MK, O'Dell SE, Kurtz DM Jr, Robson RL, Scott RA (1992) Protein Eng 5:367–371
27. Zeng Q, Smith ET, Kurtz DM Jr, Scott RA (1996) Inorg Chim Acta 242–243:245–251
28. Blake PR, Park JB, Bryant FO, Aono S, Magnuson JK, Eccleston E, Howard JB, Summers MF, Adams MWW (1991) Biochemistry 30:10885–10895
29. Stura EA, Wilson IA (1992) In: Ducruix A, Giegé R (eds) Crystallization of nucleic acids and proteins, chap 5. Oxford University Press, Oxford
30. Hope H (1988) Acta Crystallogr B44:22–26
31. Hope H (1990) Annu Rev Biophys Biophys Chem 19:107–126
32. Kabsch W (1988) J Appl Crystallogr 21:916
33. Sheldrick GM (1993) SHELXL-93; Program for crystal structure refinement. University of Göttingen, Germany
34. Hendrickson WA, Konnert JH (1980) In: Diamond R, Ramaseshan S, Venkatesan K (eds) Computing in crystallography. IUCr and Indian Academy of Sciences, Bangalore, pp 13.01–13.25
35. Brünger AT (1990) Program X-PLOR; A system for X-ray crystallography and NMR (version 2.1, manual). Yale University Press, Yale, Conn.
36. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Acta Crystallogr A47:110–119
37. Stec B, Zhou RS, Teeter MM (1995) Acta Crystallogr D51:663–681
38. Teeter, MM (1984) Proc Natl Acad Sci USA 81:6014–6018
39. Kirtley SW, Olsen JP, Bau R (1973) J Am Chem Soc 95:4532
40. Hendsch ZS, Tidor B (1994) Protein Sci 3:211–226
41. Dao-pin S, Sauer U, Nicholson H, Matthews BW (1991) Biochemistry 30:7142–7154
42. Hendsch ZS, Jonsson T, Sauer RT, Tidor B (1996) Biochemistry 35:7621–7625
43. Dauter Z, Lamzin VS, Wilson KS (1997) Curr Opin Struct Biol 7:681–688