Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Ab initio solution and refinement of two high-potential iron protein structures at atomic resolution

Emilio Parisini,[a]† Francesco Capozzi,[b] Paolo Lubini,[a] Victor Lamzin,[c] Claudio Luchinat[d] and George M. Sheldrick[a]*

[a]Institut für Anorganische Chemie, University of Göttingen, Tammannstrasse 4, D-37077, Göttingen, Germany, [b]Institut of Agricultural Chemistry, University of Bologna, Viale Berti Pichat 10, I-40127, Bologna, Italy, [c]European Molecular Biology Laboratory (EMBL) c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, and [d]Department of Soil Science and Plant Nutrition, University of Firenze, P. le delle Cascine 28, I-50144 Firenze, Italy

† Present address: Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, England.

Correspondence e-mail: gsheldr@shelx.uni-ac.gwdg.de

Received 29 April 1999
Accepted 2 July 1999

PDB References: rc-WT, 1cku; H42Q-1, 1b0y.

The crystal structure of the reduced high-potential iron protein (HiPIP) from *Chromatium vinosum* has been redetermined in a new orthorhombic crystal modification, and the structure of its H42Q mutant has been determined in orthorhombic (H42Q-1) and cubic (H42Q-2) modifications. The first two were solved by *ab initio* direct methods using data collected to atomic resolution (1.20 and 0.93 Å, respectively). The recombinant wild type (rc-WT) with two HiPIP molecules in the asymmetric unit has 1264 protein atoms and 335 solvent sites, and is the second largest structure reported so far that has been solved by pure direct methods. The solutions were obtained in a fully automated way and included more than 80% of the protein atoms. Restrained anisotropic refinement for rc-WT and H42Q-1 converged to $R_1 = \sum \big||F_o| - |F_c|\big| / \sum |F_o|$ of 12.0 and 13.6%, respectively [data with $I > 2\sigma(I)$], and 12.8 and 15.5% (all data). H42Q-2 contains two molecules in the asymmetric unit and diffracted only to 2.6 Å. In both molecules of rc-WT and in the single unique molecule of H42Q-1 the $[Fe_4S_4]^{2+}$ cluster dimensions are very similar and show a characteristic tetragonal distortion with four short Fe—S bonds along four approximately parallel cube edges, and eight long Fe—S bonds. The unique protein molecules in H42Q-2 and rc-WT are also very similar in other respects, except for the hydrogen bonding around the mutated residue that is at the surface of the protein, supporting the hypothesis that the difference in redox potentials at lower pH values is caused primarily by differences in the charge distribution near the surface of the protein rather than by structural differences in the cluster region.

1. Introduction

$[Fe_4S_4]$ clusters in proteins (Beinert, 1990; Howard & Rees, 1991; Cammack, 1992) and model compounds (Holm, 1992) have been the subject of extensive crystallographic and spectroscopic investigations. $[Fe_4S_4]$ and other iron–sulfur clusters are found in a wide variety of proteins, usually but not always involved in electron-transfer reactions. Crystallographic studies of $[Fe_4S_4]$-cluster proteins began in the early 1970's with structures of a ferredoxin (Sieker *et al.*, 1972) and a HiPIP (Carter *et al.*, 1972; Carter, Kraut, Freer, Xuong *et al.*, 1974; Carter, 1977). Recently a very high resolution $[Fe_4S_4]$-ferredoxin structure was reported by Dauter *et al.* (1997); the highest resolution previously reported for a HiPIP is 1.5 Å (Rayment *et al.*, 1992).

Both ferredoxins and HiPIPs have a stable $[Fe_4S_4]^{2+}$ state, but whereas ferredoxins can be reduced to a $[Fe_4S_4]^{+}$ state in a biological environment, HiPIPs may be oxidized to $[Fe_4S_4]^{3+}$; in both cases only a one electron transfer is observed. To a first

Table 1
Crystal data and refinement statistics for the two forms of the wild type and the H42Q mutant.

	Wild type-I	Wild type-II	H42Q-1	H42Q-2
Molecular weight (Da)	9257	2 × 9257	9248	2 × 9248
Temperature (K)	277	100	100	277
Wavelength (Å)	1.5418	0.9091	0.9116	0.9160
Crystal system	Orthorhombic	Orthorhombic	Orthorhombic	Cubic
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$I2_13$
Unit-cell dimensions (Å)	a = 42.70	a = 35.07 (4)	a = 37.68 (4)	a = 123.0 (2)
	b = 41.86	b = 51.39 (6)	b = 41.67 (5)	b = 123.0 (2)
	c = 38.08	c = 92.39 (9)	c = 41.91 (5)	c = 123.0 (2)
Volume (Å^3)	68065	166492	65818	1860867
Solvent content (%)	30	45	30	70
Crystal size (mm)	0.8 × 0.4 × 0.4	0.6 × 0.3 × 0.2	0.8 × 0.4 × 0.4	0.15 × 0.15 × 0.08
Resolution range (Å)	30–2.0	20–1.2	20–0.93	30–2.6
Reflections collected		636238	402616	35422
Independent reflections		50551	39854	8911
Completeness (%)		97.4	93.0	100
I/σ (mean)		9.0	11.5	15.9
I/σ (outer shell)		5.4	5.3	4.5
R_{symm} (all data) (%)		8.8	7.3	6.6
R_{symm} (outer shell) (%)		21.0	29.7	27.1
Restraints/parameters		17553/14394	8801/7536	8839/5237
Final R factor [$I > 2\sigma(I)$] (%)		12.0	13.6	22.2
Final R factor (all data) (%)		12.8	15.5	23.2
R_{free} [$I > 2\sigma(I)$] (%)		17.2	18.8	28.6
R_{free} (all data) (%)		17.8	20.5	30.0

approximation, the different redox potential ranges for these two systems are commonly attributed to the stabilization of the clusters by N—H···S hydrogen bonds to the S atoms, including the cysteine S atoms that are bonded to each Fe atom; there are usually eight such hydrogen bonds in ferredoxins and five in HiPIPs (Backes *et al.*, 1991). Most model compounds do not possess this stabilization and their redox behaviour only approximately emulates that of the [Fe_4S_4] proteins. The factors affecting the redox properties of these proteins have been analysed in detail in recent reviews (Bertini, Ciurli *et al.*, 1995; Stephens *et al.*, 1996; Capozzi *et al.*, 1998).

The project described in this paper, which was started before the report of the very high resolution ferredoxin structure by Dauter *et al.* (1997), was intended to provide the first atomic resolution determination of the structure of an [Fe_4S_4] protein, and to investigate whether the change in redox potential at acidic pH on replacing His42 by Gln (Babini *et al.*, 1998) is associated with significant structural changes. It is still the first atomic resolution structure of a HiPIP, and provides an opportunity to compare precise HiPIP, ferredoxin and model compound geometries for the same total cluster oxidation state.

2. Methods and materials

2.1. Expression, purification and crystallization

The recombinant wild type (rc-WT) from *C. vinosum* HiPIP has been expressed in *E. coli* strain XL1Blue (Strategene) harboring the pLEHP20 plasmid derivative in which the synthetic gene has been introduced into the poly-linker region as described by Eltis *et al.* (1994). The HiPIPs were expressed as fusion proteins with a leading peptide of poly-histidine that allows purification by affinity chromatography (HiTrap Chelating, Pharmacia). The leading peptide was subsequently cut using the protease Factor Xa (Boehringer) and the resulting HiPIPs purified by anion-exchange chromatography (DEAE–Cellulose, Whatman). The H42Q mutant was prepared using the oligomer AATTAGCGCACTGTTGCTCTT possessing the CATTG mismatch by site-directed mutagenesis (Kunkel *et al.*, 1987). The cleavage of the N-terminal polyhistidine tag would be expected to leave an additional amino acid (Ala0) on the N-terminus relative to the natural wild type; however no evidence could be found for this amino acid in any of the electron-density maps, so it was assumed to be absent in the crystallographic analyses. It is possible that it is present but highly disordered in the crystal.

The proteins were prepared in the reduced form, which is stable to air oxidation; in the successful crystallization attempts no special precautions were taken to prevent oxidation. Attempts to grow crystals in the presence of sodium dithionite to ensure that no oxidation could take place were unsuccessful, and soaking the crystals with solutions of sodium dithionite led to cracking of the crystals. Although Carter *et al.* (1972) report that oxidation might occur at room temperature in the X-ray beam, we believe that this should not happen during a low-temperature (100 K) data collection.

The H42Q mutant crystallizes in two different forms, orthorhombic (H42Q-1) in space group $P2_12_12_1$ and cubic (H42Q-2) in space group $I2_13$. Crystals were obtained by vapour diffusion using the hanging-drop technique. 2 µl of a solution containing 10 mg ml^{-1} of protein in 40 mM Tris buffer and 180 mM KCl were mixed with an equivalent volume of a 2 M solution of ammonium sulfate and equilibrated against 500 µl of the same solution in the reservoir. The plates were set aside in the dark at 277 K and relatively large crystals of the cubic form grew in about 48 h. Under the same conditions a crystal in the shape of a rectangular prism was once observed and used as seed for growing crystals of the orthorhombic form, which could not easily be obtained otherwise. Using the macroseeding technique, relatively large crystals (0.8 × 0.4 × 0.4 mm) could be obtained within about 24 h. The cryogenic data collection on H42Q-1 was performed after freezing the crystal inside a loop (Teng, 1990) in a solution containing the same concentration of ammonium sulfate and 20%(v/v) glycerol. Crystallographic data of H42Q-2 were collected at 277 K as preliminary tests of cryogenic conditions were unsuccessful.

Relatively large crystals (0.6 × 0.3 × 0.2 mm) of the recombinant wild type (rc-WT) were grown in a 1.3 *M* solution of ammonium sulfate using the vapour-diffusion method. A cryogenic solution containing the same concentration of ammonium sulfate and 25%(*v*/*v*) glycerol proved to be effective for flash freezing the crystal in a loop prior to cryogenic data collection. Although the natural wild type crystallizes in the same space group ($P2_12_12_1$) there is only one protein molecule in the asymmetric unit compared with two in rc-WT, and the cell dimensions are not related.

2.2. Data collection

Diffraction data of all the structures reported here were collected using synchrotron radiation on beamline X11 at the EMBL outstation at DESY in Hamburg; details are given in Table 1, which also includes the crystal data (Carter *et al.*, 1972) of the natural wild type for comparison. The cubic H42Q-2 diffracted to a resolution of 2.6 Å, whereas the diffraction pattern of the best crystal of the orthorhombic H42Q-1 extended to better than 1.0 Å.

For H42Q-1 data were collected to 0.93 Å at 100 K using a 180 mm MAR Research imaging-plate scanner. Three different crystal-to-detector distances (82, 180 and 380 mm for high-, medium- and low-resolution data, respectively) were employed with $\Delta\varphi$ step scans of 1.3, 2.0 and 3.0°, respectively. The overall completeness of 93.0% was achieved by combining data from two different crystals because the first data set was partially incomplete as a result of ice formation during data collection. Although the second crystal did not diffract to quite such a high resolution as the first, data from the two crystals were merged, refining a relative *B* value. 100% complete X-ray data of the cubic H42Q-2 were collected to 2.6 Å at 277 K using a 180 mm MAR Research image plate at a distance of 240 mm from the crystal.

X-ray data from a single crystal of rc-WT were collected to 1.2 Å at 100 K using a 300 mm MAR Research imaging-plate scanner. The overall completeness of 97.3% was achieved by collecting reflections at three different crystal-to-detector distances (90, 150 and 390 mm for high-, medium- and low-resolution data, respectively) with $\Delta\varphi$ step scans of 0.4, 0.8 and 2.2°, respectively. All data were processed using the *HKL* suite of programs (Otwinowski & Minor, 1997).

2.3. Structure solution

The two atomic resolution structures were solved *ab initio*, *i.e.* from the single wavelength native data alone without the use of known fragments or heavy-atom derivatives, using direct methods based on real/reciprocal space iteration (Sheldrick, 1997, 1998). This procedure involves a search for possible solutions using only the highest normalized structure factors *E* (the inner loop), followed by an expansion of the solutions with good figures of merit using all reflections (the outer loop). The outer loop, which is entered very rarely and so does not contribute much to the total computer time required, consists of iterative *E*-Fourier calculation, peak search, and elimination of potential atoms to optimize the correlation coefficient (CC) between E_o^2 and E_c^2 (Sheldrick, 1982, 1990; Sheldrick & Gould, 1995). The inner loop, which was inspired by the *Shake & Bake* philosophy (Miller *et al.*, 1993, 1994) but differs appreciably in implementation, starts with random atoms, uses only the largest *E* values and iterates tangent formula expansion, *E*-Fourier, peak search and elimination of potential atoms to maximize the function $S = \sum E_o^2(E_o^2 - 1)$.

For these two structures, it was possible to take advantage of the presence of heavier atoms by using the inner loop to search only for the Fe atoms. For rc-WT, 20 random atoms (in the first iteration) or 20 *E*-Fourier peaks (in subsequent iterations) were reduced to eight potential Fe atoms by maximizing *S*. These eight atoms provided the initial phases for two cycles of tangent expansion. The phases from the tangent expansion were used to calculate a Fourier map with amplitudes E_o, and the 20 highest peaks that were at least 2.4 Å from one another were used as potential Fe atoms in the next iteration. After five iterations using the 3140 largest $[E - \sigma(E)/2]$ and 366 865 unique triple phase relations (TPR), CC was calculated for all 51 872 data. For H42Q-1, ten potential Fe atoms were reduced to four by maximizing *S*; 2415 large *E* values and 292 560 unique TPR were employed in the inner loop, and CC calculated for all 41 951 data.

The structure solutions were repeated after full anisotropic refinement so that the *E*-weighted mean phase errors (MPE) could be calculated, taking allowed origin and enantiomorph changes into account. For rc-WT, about 0.7% of the trials led to correct solutions, corresponding to one correct solution every 9 h of 200 MHz Pentium processor time. For H42Q-2 the success rate was 1.5% or one solution every 3 h. The program has subsequently been optimized, which would reduce these times somewhat. The CC values based on the Fe atoms alone were in both cases about 6% better for the correct solutions than for the best incorrect solutions. For rc-WT, the CC for correct solutions increased from about 18% (MPE *ca* 58°) to about 75% (MPE *ca* 27°) on outer-loop expansion from the eight Fe atoms to *ca* 1365 atoms. The corresponding figures for H42Q-2 were 21% (MPE 54°) and 73% (MPE 22°) on expansion from four Fe to *ca* 615 atoms. The better final MPE for H42Q-2 can be attributed to the higher resolution of the data. Fig. 1 compares the *ab initio* maps with the electron density from the final refinement for a well defined region (low *B* values) in the vicinity of the cluster and one of the worst defined regions, a loop with high *B* values. As would be anticipated from the MPE values, the interpretation of these maps was straightforward and only a few side chains with high *B* values needed to be added with the help of difference maps. It should be emphasized that the *ab initio* solution of these structures was facilitated by the presence of the Fe atoms, and that atomic resolution was also a necessary requirement.

The structure of the low-resolution cubic H42Q-2 was solved by molecular replacement using the program *AMoRe* (Navaza, 1994). The natural wild type (Carter *et al.*, 1972) was used as search fragment, as the only difference in the sequence of the two structures consists of the mutation of His42 to

glutamine. The two space groups $I2_13$ and $I23$ are both consistent with the systematic absences; the former was tried first and proved to be correct. The unit-cell volume of 1 860 867 Å^3 suggested the presence of two independent molecules in the asymmetric unit, with a solvent content of about 70% (high but not unusual for a cubic space group). A $P1$ cell with equal axes of 80 Å was used to calculate the Fourier transform of the starting model. Data in the resolution range 15–3.0 Å were used for calculating the rotation function and to find an initial orientation of the model. No clear differentiation was observed in the correlation coefficients of the first 12 solutions (range 15.4–14.7%) but the 13th was

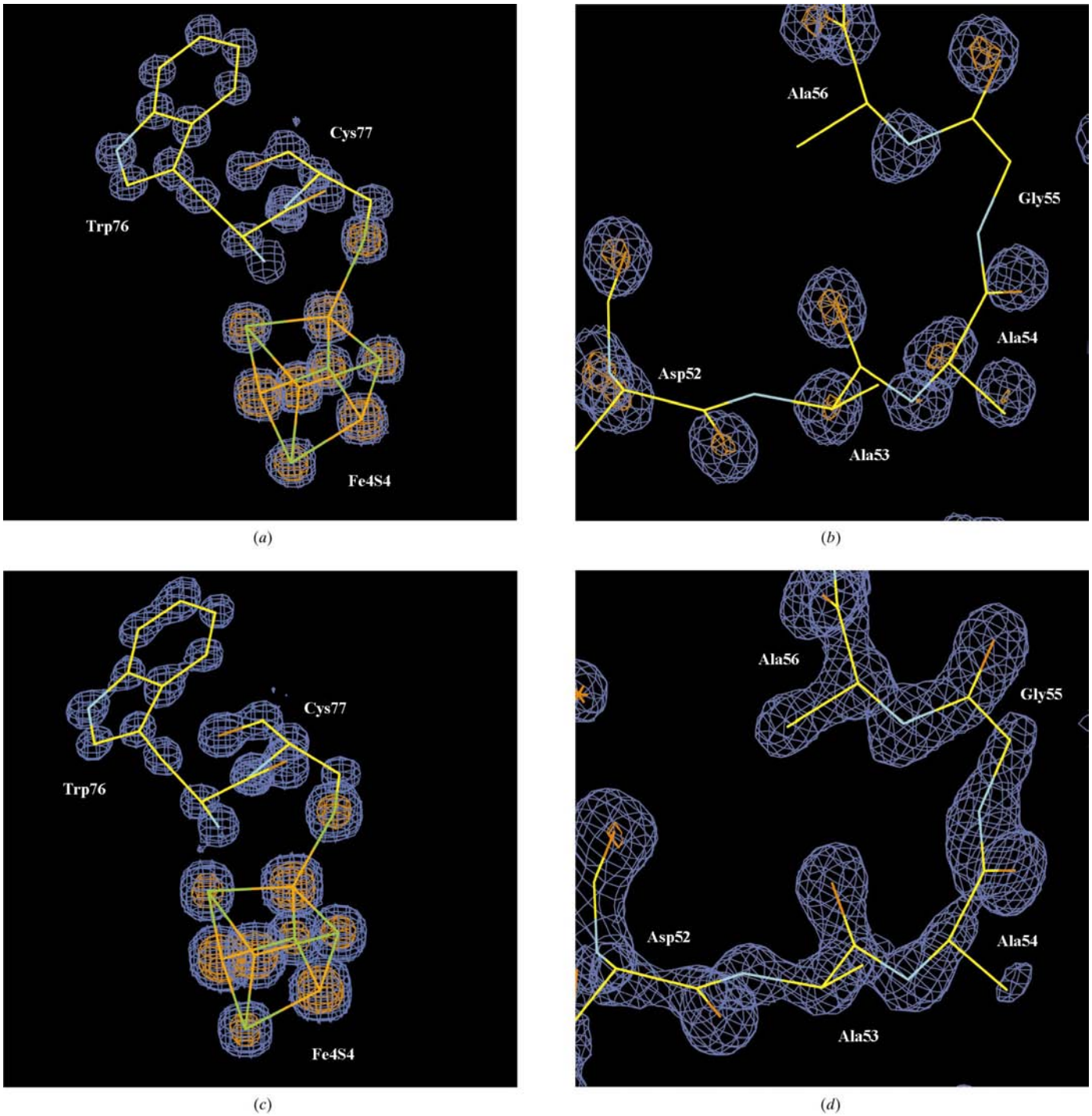

Figure 1
Comparison of the electron-density map resulting from the *ab initio* solution [(*a*) and (*b*)] and after structure refinement [(*c*) and (*d*)] for two different regions of H42Q-1. All the atoms in the rigid region around the cluster were found in their correct positions. Only some of the atoms in the side chains of residues on the surface of the protein could not be located in the initial *ab initio* map. The corresponding rc-WT maps were also of high quality. The electron-density maps (*a*) and (*c*) are contoured at the 3σ and 12σ levels; (*b*) and (*d*) are contoured at the 2σ and 6σ levels.

11.4%. The translation function was therefore calculated for the best 12 rotation solutions, using data in the same resolution range. The best translation of the first rotation solution gave the highest correlation coefficient and the lowest R factor, 38.1 and 46.3%, respectively. This solution was then used to define the orientation and position of the first molecule and the translation function was calculated for each of the remaining best rotation solutions to find the position of the expected second monomer in the asymmetric unit. The rotation solution number 8 gave the largest rise in correlation coefficient (to 61.8%) and largest drop in the R factor (to 38.2%).

2.4. Refinement

The refinement of all three structures reported here was performed using the program *SHELXL*97 (Sheldrick & Schneider, 1997). No restraints were applied to the 1,2 and 1,3 distances involving the atoms of the $[Fe_4S_4]$ cluster in the two atomic resolution structures, whereas the chemically equivalent 1,2 distances within the clusters of the low-resolution H42Q-2 were restrained to be equal with a standard deviation of 0.02 Å. In all cases the Engh & Huber (1991) restraints were applied to all 1,2 and 1,3 distances not involving the clusters with standard deviations of 0.02 and 0.04 Å, respectively. Chiral volume restraints were applied to the Cα atoms (except glycine) and to the Cβ atoms of threonine and isoleucine. Planarity restraints were applied to the atoms defining the peptide bond, to the aromatic side chains and to the appropriate parts of the amide, carboxylate and arginine side chains. The H atoms were included in calculated positions in all cases with idealized geometry and refined using a riding model. All non-H atoms in both high-resolution structures were refined anisotropically, whereas H42Q-2 was refined isotropically with the U values of spatially close atoms restrained to be equal (e.s.d.s = 0.03–0.10 Å^2). Rigid-bond and similar-ADP restraints were applied to the anisotropic atoms. The solvent water atoms were restrained to be approximately isotropic, with an e.s.d. of 0.1 Å. Two diffuse solvent parameters (Moews & Kretsinger, 1975) were refined throughout, and anti-bumping restraints were employed (Sheldrick & Schneider, 1997). Local non-crystallographic symmetry restraints (Usón *et al.*, 1999) were used in the refinement of the two independent molecules in the asymmetric unit of rc-WT and H42Q-2. These restrained corresponding 1,4 distances (derived with the help of a connectivity table) to be equal in the NCS-related molecules. No NCS restraints were applied to the anisotropic displacement parameters, but the isotropic displacement parameters of NCS-related atoms in H42Q-2 were restrained to be approximately equal (e.s.d. = 0.03 Å^2).

The refinement was performed using 95% of the data (50 551, 39 854 and 8911 reflections for rc-WT, H42Q-1 and H42Q-2, respectively). The remaining 5% (1321, 2097 and 469 reflections) were used only to monitor the refinement with the $R_{\rm free}$ test (Brünger, 1992) and not for refinement or calculation of maps. Only in the final refinement run were the working set and the reference set merged and used for calculating maps. The final free R factors (17–20%, see Table 1) probably reflect the influence of the large number of weak high-resolution data included in the calculations.

The program *XtalView* (McRee, 1992) was used for manual adjustment of side chains *etc.*, and as an independent check on the structure quality the programs *PROCHECK* (Laskowski *et al.*, 1993) and *WHAT-IF* (Vriend, 1990) were employed. The Ramachandran plots show that for both rc-WT and H42Q-1, 100% of the residues (excluding Gly and Pro) are in the core region (expected 98%) and 78% are in the inner core region (expected 80%; Kleywegt & Jones, 1996); for the low-resolution H42Q-2 the figures were 94 and 65%, respectively.

About half of the water molecules in the two high-resolution structures were placed using the automatic water divining procedure implemented in the program *SHELXWAT* (Sheldrick & Schneider, 1997). The rest of the solvent peaks were assigned by close inspection of the electron-density maps using *XtalView*. Of the 335 water sites found in the asymmetric unit of rc-WT, 264 were assigned to fully occupied water molecules and 71 to water molecules with half occupancy. For H42Q-1 the 206 water sites consisted of 132 full and 74 half occupancy. Only 50 (full) water molecules could be assigned in the case of the low-resolution structure, mostly by inspection of the electron-density maps, since the limited resolution made the automated water divining procedure less reliable.

3. Results and discussion

3.1. General comments

After completion of the structure determinations, it became apparent that H42Q-1 is essentially isostructural with the natural wild type reported by Carter *et al.* (1972). The deposited coordinates for single unique protein molecule in this structure may be converted to those found in this work by the transformation $\frac{1}{2} - z$, $1 - x$, y, which involves a cyclic permutation of the unit-cell vectors. However, the rc-WT and H42Q-2 structures, both with two molecules in the asymmetric unit, represent new crystalline modifications with unrelated packing of the protein molecules in the unit cell.

Since the resolution of the H42Q-2 data was so much worse than that of the other two structures, it will not be discussed further. The two high-resolution structures, including anisotropic displacement parameters and reflection data, have been deposited with the PDB for immediate release.

3.2. Assessment of standard uncertainties

For the smallest structure, H42Q-1, it proved possible to invert the full least-squares matrix for the unrestrained coordinates only (U^{ij} held fixed) to obtain reliable estimates of the standard uncertainties (s.u.) of the atom positions. For rc-WT the corresponding matrix inversion was unstable. The (radial) positional standard uncertainties $\sigma(r)$ are shown in Fig. 2(a) as a function of the equivalent B value ($B = 8\pi^2 U$) for H42Q-1. The lines show a simultaneous fit to a quadratic function in B for C, N and O atoms, which takes the form,

$$\sigma(r) = (0.00676 + 0.007437B + 0.0001323B^2)/Z_i^{\#}, \quad (1)$$

where $Z_i^{\#}$ is the scattering factor of the three elements at $\sin(\theta)/\lambda = 0.3$ Å^{-1}. It can be deduced that the mean positional s.u. for a C atom with the average B value (14.1) of the non-solvent atoms is 0.057 Å. Cruickshank (1999) has suggested two formulae for estimating this value, intended for use in structures not amenable to full-matrix inversion,

$$\sigma(r, B_{av}) = [3\sum(Z^{\#})^2/(n_{obs} - n_{par})]^{1/2} C^{-1/3} R\, d_{min}/Z_i^{\#} \quad (2)$$

$$\sigma(r, B_{av}) = [3\sum(Z^{\#})^2/n_{obs}]^{1/2} C^{-1/3} R_{free}\, d_{min}/Z_i^{\#}, \quad (3)$$

where n_{obs} is the number of observed reflections, n_{par} is the total number of parameters, d_{min} is the highest resolution of the data, C is the fractional completeness to this resolution, R is the conventional R index (on all F) and R_{free} is the corresponding free R value. Equations (2) and (3) lead to estimates of the mean positional s.u. of a C atom (with an average B) of 0.056 and 0.067 Å, respectively, in good agreement with the experimental value. It should be noted that the standard uncertainties $\sigma(x)$ in the atomic coordinates will be a factor of $3^{1/2}$ smaller. Tickle *et al.* (1998) suggested an alternative relation between R_{free} and R to that assumed by Cruickshank, namely,

$$R_{free} = R(2n_{obs} + n_{par})/(2n_{obs} - n_{par}). \quad (4)$$

Combining equations (2) and (4) gives an estimate of 0.62 Å for the s.u., in slightly better agreement with the experimental value of 0.57 Å than the estimate from (3).

The bond-length standard uncertainties from the unrestrained H42Q-1 refinement (Fig. 2*b*) show an almost linear monotonic increase with the mean equivalent B value of the two atoms concerned; C—C bonds tend to have higher standard uncertainties than C—O bonds with the same mean B values, with intermediate standard uncertainties for the C—N bonds. Quite a different picture (Fig. 2*c*) emerges from the corresponding restrained refinement; all the standard uncertainties in bond lengths asymptote to the value of 0.02 Å that

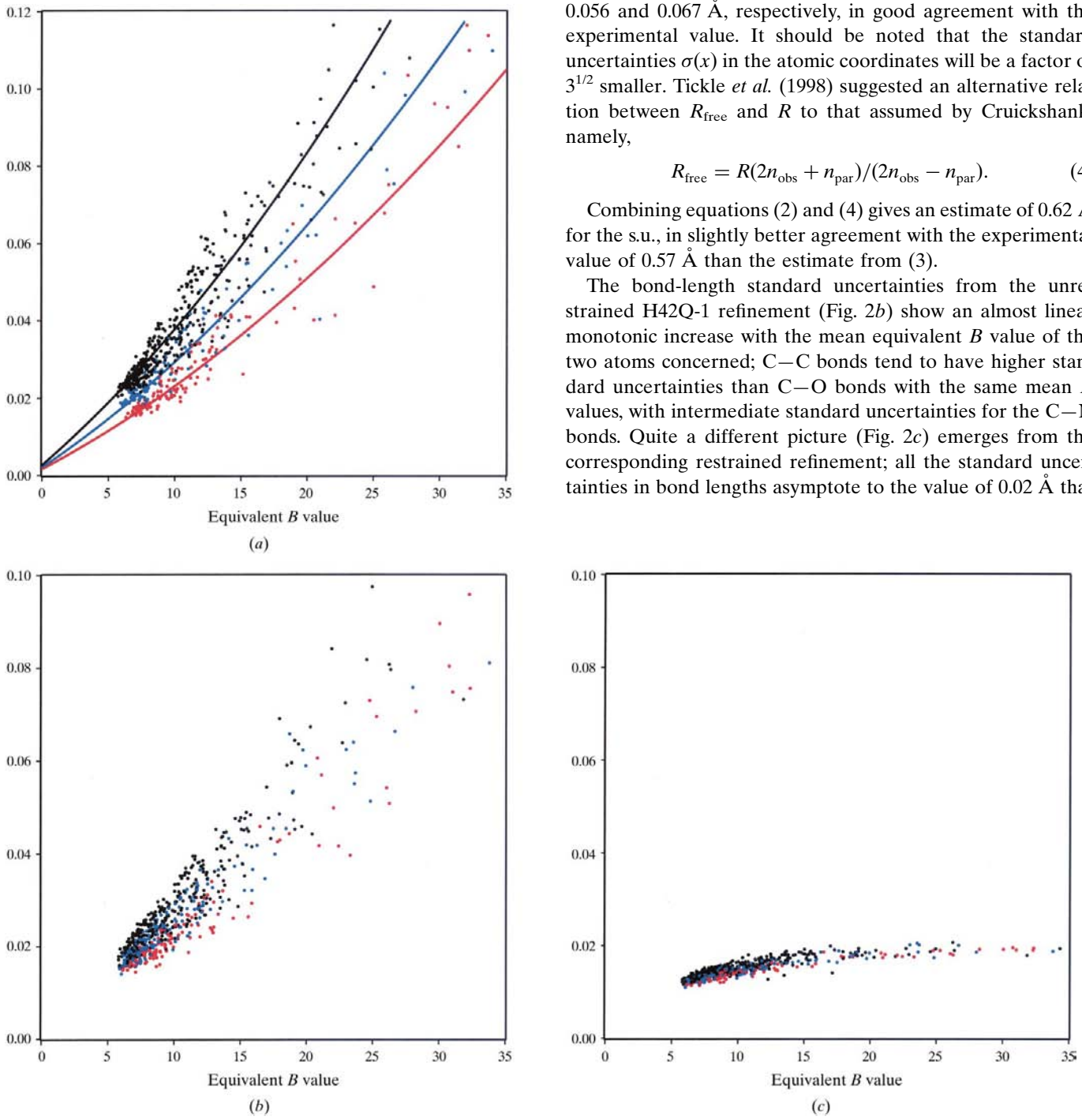

Figure 2
Standard uncertainties from the full-matrix refinement of H42Q-1 in (*a*) atomic positions, colour coded black (C), blue (N) and red (O); (*b*) bond distances from the unrestrained refinement, colour coded black (C—C), blue (C—N) and red (C—O), (*c*) bond distances from the restrained refinement (with the same colour scheme).

Table 2
Bond distances in the Fe_4S_4 clusters of the two atomic resolution structures.

	Wild type-II Mol *A*	Wild type-II Mol *B*	H42Q-2	Mean
Short Fe—S				
Fe(1)—S(2)	2.238 (6)	2.234 (5)	2.222 (4)	
Fe(2)—S(1)	2.233 (6)	2.212 (5)	2.228 (4)	
Fe(3)—S(4)	2.245 (6)	2.261 (6)	2.242 (4)	
Fe(4)—S(3)	2.281 (6)	2.273 (5)	2.266 (4)	2.244
Long Fe—S				
Fe(1)—S(3)	2.327 (5)	2.336 (6)	2.321 (4)	
Fe(1)—S(4)	2.310 (6)	2.313 (6)	2.292 (5)	
Fe(2)—S(3)	2.299 (6)	2.304 (6)	2.295 (5)	
Fe(2)—S(4)	2.296 (6)	2.323 (6)	2.310 (4)	
Fe(3)—S(1)	2.313 (6)	2.320 (6)	2.307 (4)	
Fe(3)—S(2)	2.325 (6)	2.317 (6)	2.306 (4)	
Fe(4)—S(1)	2.285 (6)	2.282 (6)	2.284 (5)	
Fe(4)—S(2)	2.297 (6)	2.308 (6)	2.295 (4)	2.307
Short Fe···Fe				
Fe(1)···Fe(3)	2.705 (4)	2.724 (4)	2.709 (3)	
Fe(1)···Fe(4)	2.711 (5)	2.695 (4)	2.702 (3)	
Fe(2)···Fe(3)	2.704 (4)	2.729 (4)	2.707 (3)	
Fe(2)···Fe(4)	2.711 (5)	2.693 (5)	2.699 (3)	2.708
Long Fe···Fe				
Fe(1)···Fe(2)	2.731 (5)	2.734 (4)	2.747 (3)	
Fe(3)···Fe(4)	2.757 (4)	2.738 (4)	2.722 (3)	2.738
Fe—SG				
Fe(1)—SG_43	2.249 (6)	2.267 (6)	2.252 (4)	
Fe(2)—SG_46	2.299 (7)	2.304 (6)	2.274 (4)	
Fe(3)—SG_63	2.265 (6)	2.272 (6)	2.258 (4)	
Fe(4)—SG_77	2.273 (6)	2.269 (6)	2.253 (4)	2.270

was used as a restraint standard deviation in the refinement. Thus, except for the best defined atoms, the bond lengths are determined primarily by the restraint target values. This does not apply to the torsion angles and to the cluster geometry, which were not restrained during the refinement.

The standard uncertainties in the bond lengths within the H42Q-1 cluster (0.93 Å resolution) shown in Table 2 are around 0.004 Å, similar to those quoted for the 0.94 Å resolution structure of the ferredoxin from *Clostridium acidurici* (Dauter *et al.*, 1997) and the model cluster $[Fe_4S_4(SBu^t_4)_4]^{2-}$ (Mascharak *et al.*, 1983), which in both cases were about 0.003 Å. As would be expected for the lower resolution (1.20 Å), the standard uncertainties for the rc-WT structure (also shown in Table 2) are a little larger, however if the equivalent distances in the two molecules in the asymmetric unit are averaged, the resulting standard uncertainties are also about 0.004 Å. These standard uncertainties are an order of magnitude less than the precision associated with earlier studies, and justify a re-examination of the cluster geometry to identify possible distortions and differences between these structures.

3.3. Comparison of the three independent molecules

A least-squares fit of the cluster atoms and the Cα atoms of residues 5–81 inclusive gave a root-mean-square deviation (r.m.s.d.) of 0.44 Å between the two independent molecules of rc-WT but only 0.31 and 0.32 Å between H42Q-1 and each of these two molecules. Fig. 3 shows the resulting superposition of the main-chain and cluster atoms. A possible explanation of this surprising result is that the H42Q-1 structure is more accurate because of the higher resolution, and so is closer to the average rc-WT structure than the two independent rc-WT molecules are to one another. The Kleywegt (1996) plot (Fig. 4) confirms the good agreement of the φ and ψ angles in the two molecules of rc-WT. All three molecules are also consistent with the NMR structures reported by Banci *et al.* (1995) and Bertini, Dikiy *et al.* (1995) for the oxidized and reduced forms of rc-WT in solution, within the much larger standard uncertainties of the NMR structures.

Fig. 5 shows the differences in the side-chain χ_1 angles of the three possible pairs of molecules. Eight of the side chains switch between the two possible *gauche* conformations when the two independent rc-WT molecules are compared; in each of these eight cases the H42Q-1 conformation corresponds closely to one or other rc-WT molecule. There are also differences in the ring conformations of Pro3 and Pro67, but in both cases the mutant is closer to one of the rc-WT molecules than they are to each other. That leaves only four side chains (Ser1, Asn11, Val29 and Val73) where the mutant conformation is different to both rc-WT molecules, and of these the OH group of Ser1 makes a strong hydrogen bond to the carbonyl O atom of a symmetry-related Lys25 in the mutant only, the side-chain NH2 of Asn11 makes a hydrogen bond with a symmetry-related Asp58 side chain in the mutant only, and Val73 is close to the site of the mutation (see Fig. 6).

The conformational differences introduced by the mutation are thus small and, on the whole, even smaller than the differences between the two independent rc-WT molecules in the same crystal. It is likely that most of these small differences arise from intermolecular interactions.

At the pH of 5.3, at which the crystals were grown, the only histidine (His42) in the wild type is predominantly doubly

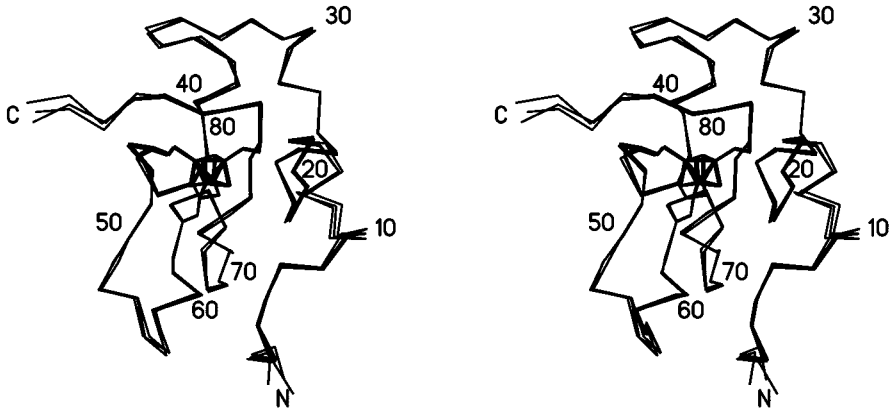

Figure 3
Stereo plot of the best fit of the three independent molecules in the high-resolution structures, which are seen to have very similar backbone conformations.

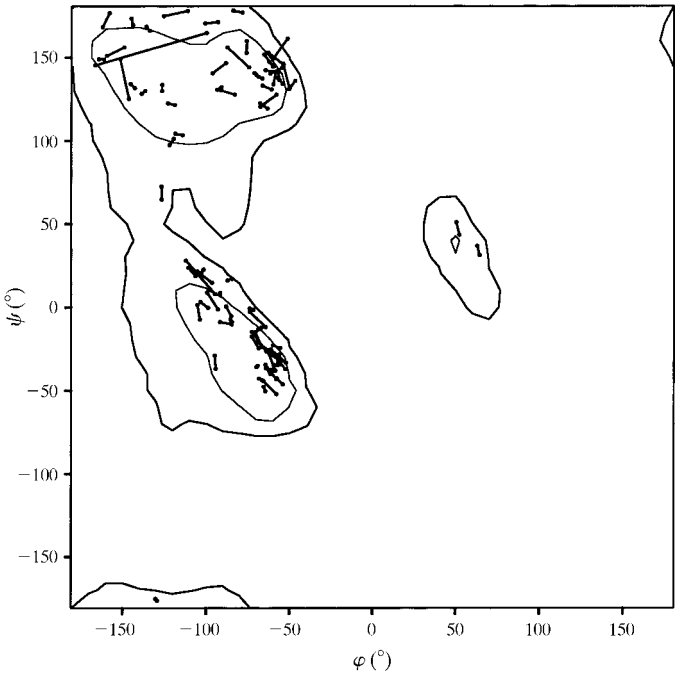

Figure 4
Kleywegt (1996) NCS plot for rc-WT. The φ/ψ points for the same amino acids of the two independent molecules are joined by lines; this shows how well the NCS is obeyed by the main-chain atoms. The plot was drawn using *SHELXPRO* (Sheldrick & Schneider, 1997).

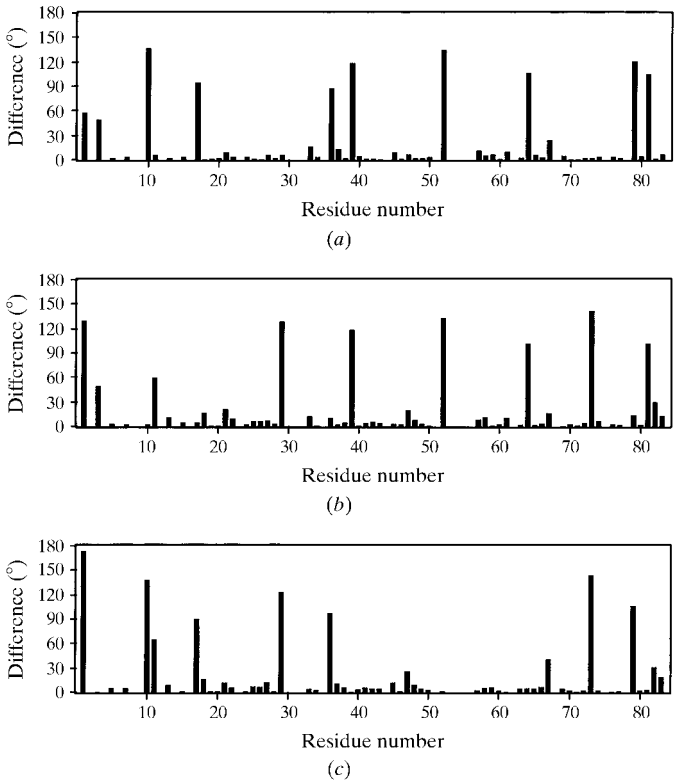

Figure 5
Absolute differences in χ_1 between (*a*) the two independent molecules *A* and *B* of rc-WT, (*b*) H42Q-1 and molecule *A* of rc-WT and (*c*) H42Q-1 and molecule *B* of rc-WT.

protonated and, therefore, positively charged. As shown in Fig. 6, the His42 residue in monomer *A* takes part in hydrogen-bonding contacts with the backbone carbonyl O atom of Val73 of the same molecule [2.77 (2) Å] and with that of Ala32 of a symmetry-related molecule [2.65 (2) Å], whereas the His42 residue in monomer *B* (not shown) makes only weak interactions; it is more exposed to the solvent and may be a mixture of two conformations. In the mutant, residue 42 is

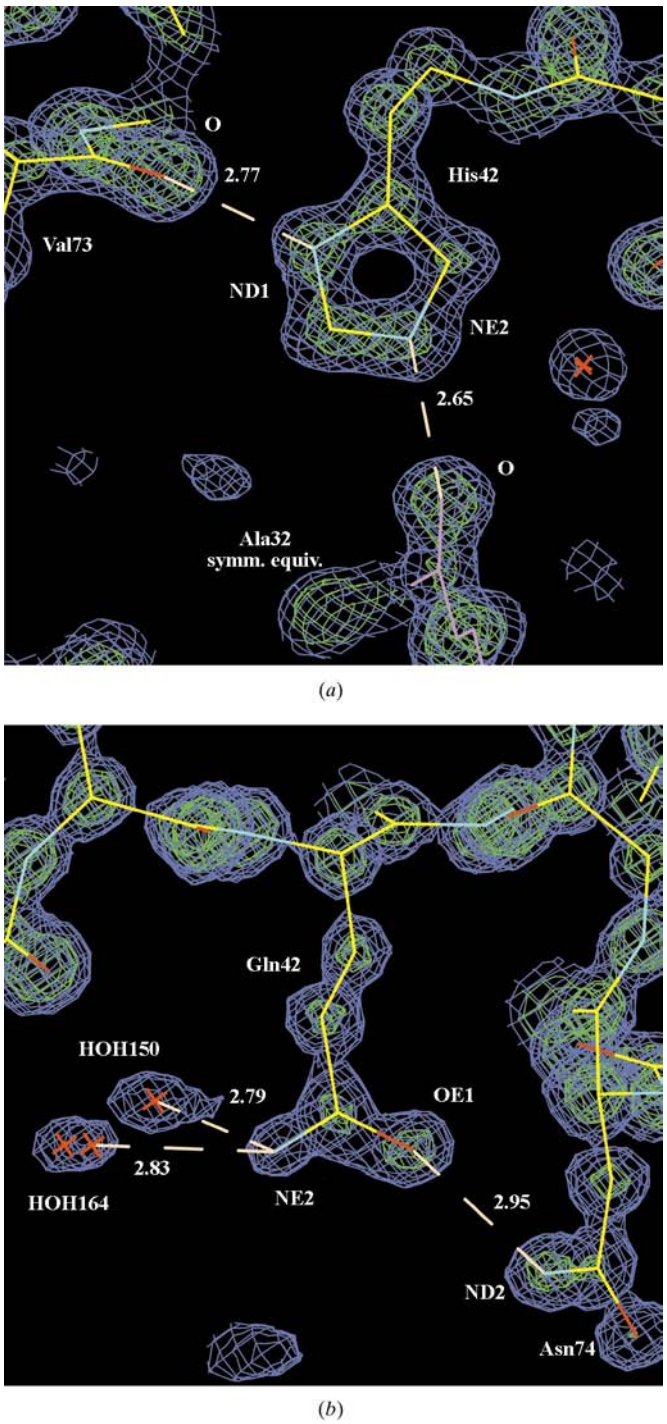

Figure 6
Electron-density map of (*a*) molecule *A* of rc-WT and (*b*) H42Q-1 in the region of the mutation. These maps are contoured at 2σ and 4σ levels.

Table 3
Hydrogen-bonding interactions involving the protein matrix and the S atoms in the clusters.

Donor	Acceptor	$D\cdots A$ distance wild-type II (Å)		$D\cdots A$ distance H42Q-1 (Å)
		Mol *A*	Mol *B*	
N_48	SG_46	3.460 (14)	3.444 (15)	3.400 (15)
N_65	SG_63	3.298 (14)	3.298 (14)	3.351 (11)
N_77	S3_87	3.411 (15)	3.411 (16)	3.396 (12)
N_79	SG_77	3.423 (15)	3.309 (15)	3.400 (10)
N_81	SG_46	3.781 (15)	3.563 (15)	3.554 (11)

replaced by Gln and so is uncharged; it accepts an hydrogen bond from the side chain of Asn74 and also interacts with two water molecules.

3.4. Cluster geometry

Studies of model compounds (Averill *et al.*, 1973, Berg & Holm, 1982; Mascharak *et al.*, 1983) and of the *C. vinosum* HiPIP at 2 Å resolution (Carter *et al.*, 1972) shown the presence of tetragonally distorted clusters with approximate $D_2d(-42m)$ symmetry and four short and eight long Fe—S bonds. As can be seen from Table 2, this is also true to a good first approximation for all three independent molecules reported here. Similarly there are four short and two long Fe···Fe distances. Fig. 7 shows a thermal ellipsoid plot of the cluster with the four short bonds roughly from top to bottom, and the eight long bonds at right angles to them. Two of the 'short' Fe—S bonds (the first two rows of the Table) are systematically shorter than the other two in all three molecules, reducing the ideal symmetry to $C_{2v}(2mm)$.

The mean of the four short Fe—S bonds is 2.244 Å, which should be compared with 2.25 Å in the lower resolution study (Carter, Kraut, Freer & Alden, 1974) and 2.25 Å in the 1.5 Å structure of *R. tenuis* HiPIP (Rayment *et al.*, 1992) which shows a similar $C_{2v}(2mm)$ distortion. The corresponding mean values for the 'long' Fe—S bonds are 2.307 Å (for the three molecules reported here), 2.36 Å (*C. vinosum* at 2 Å resolution) and 2.30 Å (*R. tenius*). The 0.94 Å *C. acidurici* ferredoxin structure reported by Dauter *et al.* (1997) contains two $[Fe_4S_4]^{2+}$ clusters; one has the typical D_2 distortion with mean Fe—S distances of 2.251 (four short) and 2.302 Å (eight long bonds), the other is less symmetrical, the best D_2d approximation gives corresponding mean Fe—S distances of 2.264 and 2.291 Å. A recent 1.35 Å structure of a 7-Fe ferredoxin (Stout *et al.*, 1998) has mean distances of 2.32, 2.29 and 2.28 Å for the three sets of four parallel edges of the $[Fe_4S_4]^{2+}$ unit. Allowing for the lower precision of the original *C. vinosum* structure, the consistency of these mean bond lengths is impressive, and also indicates that there are no significant differences in the cluster geometry of the $[Fe_4S_4]^{2+}$ clusters in reduced HiPIPs and oxidized ferrodoxins, despite the different redox potentials and different number of N—H···S hydrogen bonds stabilizing the clusters.

Despite the above result, the Fe—S bonds in the three independent molecules reported here are on average slightly longer when the S atom accepts an N—H···S hydrogen bond (Table 3), as would have been anticipated by the bond-valence model. For example the longest Fe—S(Cys) distances in all three molecules involve SG of Cys46, the only sulfur that accepts two such hydrogen bonds.

3.5. Solvent structure and conserved waters

The solvent content of the crystals is approximately 30% in H42Q-1, 45% in rc-WT, and about 70% in H42Q-2, showing the usual correlation between solvent content and maximum resolution of the observed diffraction data. Most of the solvent molecules are in the first hydration shell and make hydrogen bonds with the protein molecules. The occupancy sum of this first hydration shell is 112 and 121.5 for the two unique HiPIP molecules in rc-WT and 95 in H42Q-1. The occupancy sum of solvent molecules that are simultaneously in the first solvation shells of two different HiPIP molecules is 35 in rc-WT and 26.5 in H42Q-1. The smaller numbers for H42Q-1 imply larger areas of direct contact between the protein molecules, consistent with the low solvent content, but the higher resolution data also enabled more alternative sites to be modelled for this structure.

Comparison of the hydrogen-bonding patterns of the waters in the first hydration shell revealed six strongly conserved sites, of which three are sufficiently close to the clusters to be relevant to the question of solvent accessibility of the clusters, which is believed to be a factor affecting redox processes. The environment of these three sites is shown in stereo in Fig. 8. The site shown in Fig. 8(*a*) involves a disordered solvent molecule in H42Q-1, and both (*a*) and (*b*) would appear to

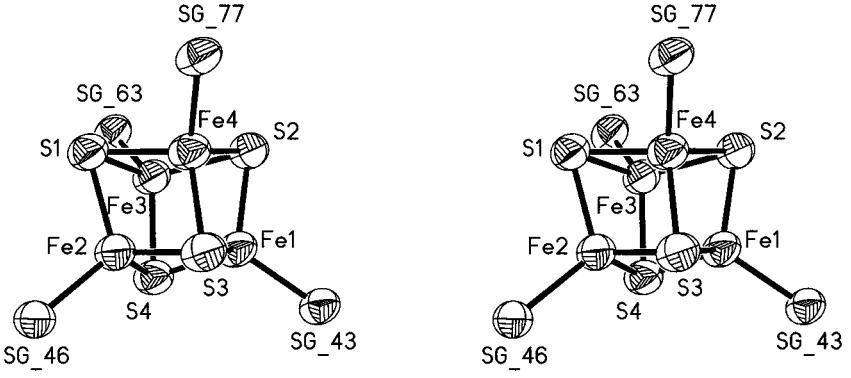

Figure 7
Stereo *ORTEP*-style plot of the $[Fe_4S_4]$ cluster and the four S atoms of the cysteines bound to the Fe atoms showing 50% probability thermal ellipsoids. The bonds that are shown vertically in this diagram are the shortest in the cluster (see Table 2).

be in a position for rapid exchange with the bulk solvent. The site shown in Fig. 8(*c*) is of interest because the residue Tyr19 (the only completely conserved residue in all HiPIPs apart from those that bind the cluster) may be implicated in electron transfer (Carter, Kraut, Freer & Alden, 1974) though recent work indicates that its primary role is to improve the hydrolytic stability of the cluster (Agarwal *et al.*, 1995; Iwagami *et al.*, 1995). The conserved water molecule is hydrogen bonded to its aromatic OH group.

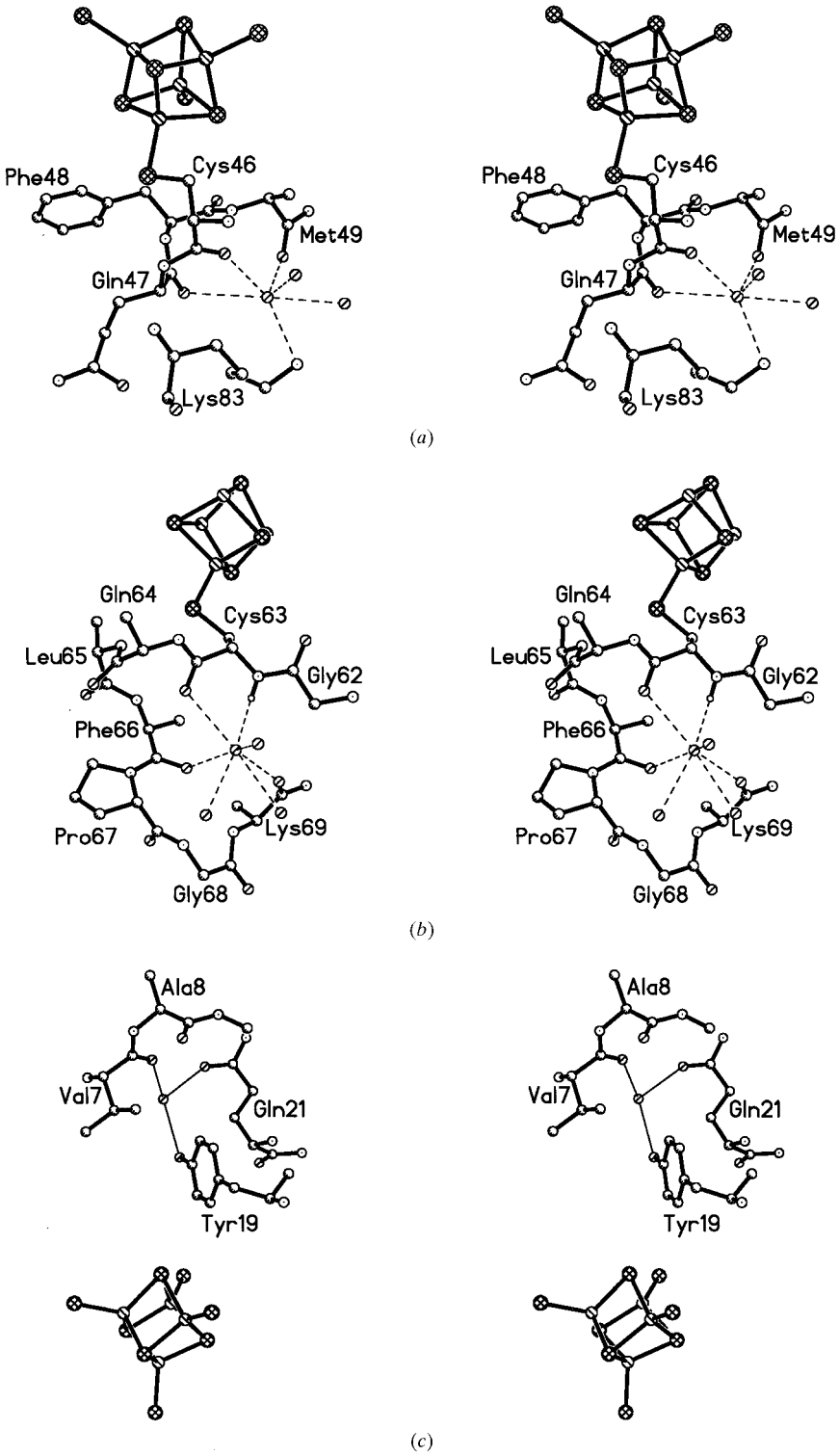

Figure 8
Stereoview of three of the water sites that are conserved for the three independent molecules in the high-resolution structures. They are close enough to the cluster to be relevant to the question of solvent accessibility.

3.6. Secondary structure

The secondary structure of the monomeric form of the wild type has been described by Carter, Kraut, Freer, Xuong *et al.* (1974) as mainly consisting of two short α-helices (spanning residues 12–16 and residues 28–31, respectively) and 17 hairpin turns, with all the remaining residues in an extended conformation. At atomic resolution this same pattern is observed for all three independent protein molecules. Of the 17 β-turns observed, 12 are of type I, two of type II and three of type III, with three pairs of linked hairpins in positions 23–25, 38–40 and 43–45. The two α-helical segments, although relatively short, are characterized by combinations of φ/ψ angles that are in close agreement with those expected for regular secondary-structure elements of this type.

Three short antiparallel β-strands extend from residues 48 to 50, 59 to 64 and 70 to 73. Although the rather long portions of polypeptide chain connecting the three strands allow the third strand to be nearly perpendicular to the plane of the first one, the pronounced right-handed twist observed in the second strand accounts for the observed hydrogen-bonding interactions with the other two segments of the chain and, therefore, the stability of their stranded conformation.

It is interesting to note that the loop between the first two antiparallel β-strands (shown in Figs. 1*b* and 1*d*) that shows larger displacement parameters (*B* values) also shows more pronounced disorder in the solution structures (Banci *et al.*, 1995; Bertini, Dikiy *et al.*, 1995), faster exchange of the peptide protons (Bertini *et al.*, 1997) and higher mobility in the unfolding intermediate obtained in the presence of guanidinium chloride (Bertini *et al.*, 1997; Bentrop *et al.*, 1999).

3.7. Intermolecular interactions

The existence of significant areas of contact between pairs of molecules in the solid state suggests that these interactions might also lead to dimer or oligomer

formation in solution, although in solution each molecule is almost certainly involved in dynamic exchange processes. There is electron spin resonance evidence for dimer formation by the *C. vinosum* HiPIP in solution (Dunham *et al.*, 1991), and dimers have also been identified in the crystal structures of the *R. tenuis* and *E. halophila* HiPIPs (Rayment *et al.*, 1992) that are similar to the dimer modelled from the spectroscopic data. The dimer structure is relevant to the rate of electron transfer between oxidized and reduced forms of the same protein molecule in solution.

The shortest approaches of the clusters in the rc-WT and H42Q-2 structures are shown in Fig. 9; the distance between the centroids of the two clusters is 16.7 Å in rc-WT (Fig. 9*a*) and 22.2 Å in H42Q-2 (Fig. 9*b*). The mutual orientation of the two independent molecules of rc-WT is quite different to that found for the two molecules forming the asymmetric unit of the cubic form of H42Q-2. Although the contact surface and the inter-cluster distance in the rc-WT dimer are similar to those found in the previously proposed dimer, the two monomers appear to have been rotated by nearly 180° about the line joining the two clusters. It appears that the dimer found in the rc-WT structure could also be consistent with the spectroscopic evidence, so the presence of both types of dimer in solution would explain the observation that the spectra indicated a mixture of two closely related but different dimers (Dunham *et al.*, 1991).

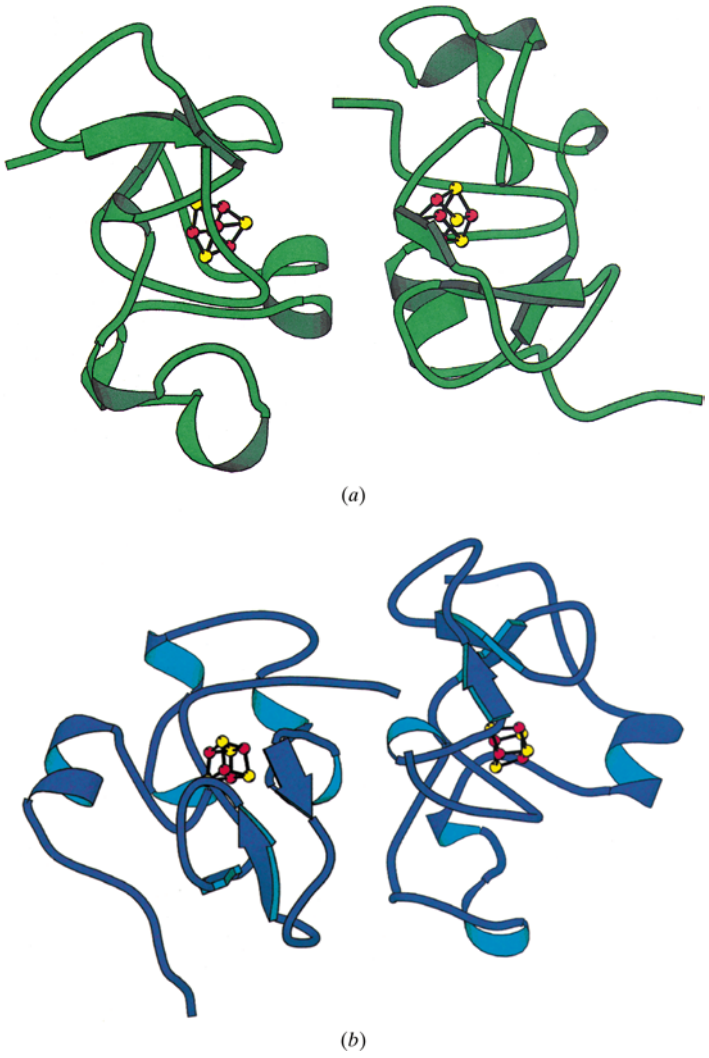

Figure 9
A *MOLSCRIPT* (Kraulis, 1991) plot of the asymmetric unit of (*a*) rc-WT and (*b*) H42Q-2. The interface between the two molecules in these crystallographic dimers and their inter-cluster distances are quite different.

The residues at the interface between the two monomers forming the asymmetric unit of rc-WT are Gln47, Phe48, Met49, Gln64, Leu65 and Ser79 in the first monomer and Ala16, Leu17 and Leu65 in the second one. The interaction between the two monomers provides a rather efficient way of stabilizing this large non-polar portion of the protein surface. It is interesting that three of the residues involved in this interaction also make N—H···S hydrogen bonds to the cluster, suggesting a possible role (*e.g.* for the relatively conserved Phe48) in electron transfer.

The residues forming most of the interface between the two molecules in the asymmetric unit of the cubic form of the H42Q mutant are Leu36, Glu40, Gln42, Ala44 and Asp45, and the corresponding residues in the second monomer. The mutation of a positively charged residue (His42) at the working pH to a polar but uncharged amino acid (Gln) has had the effect of removing a positive charge on that portion of the protein surface, thus leaving a large hydrophobic patch that forms the contact surface in the dimer. This may explain why the cubic form crystallized preferentially for the mutant but has not been observed for the wild type.

4. Concluding remarks: implications for electron transfer

The two atomic resolution structures presented here show that the mutation, despite its effect on the redox potential, produced no significant change in the overall conformation of the protein or the detailed geometry of the iron–sulfur cluster. The availability of standard uncertainties from full-matrix refinement and of two crystallographically independent molecules in the rc-WT structure enabled a reliable assessment of the significance of the small differences observed. The changes introduced by the mutation appear to be localized to its immediate vicinity on the surface of the protein. The difference in redox potential between the mutant and wild type at low pH can thus be attributed to the presence of the positively charge protonated histidine in the wild type. At higher pH this residue is no longer protonated and so the redox potentials are similar, though some pH dependence is still observed for the mutant (Babini *et al.*, 1998). This confirms the view (Stephens *et al.*, 1996) that these redox potentials are determined by the overall arrangement of point charges and dipoles in the protein rather than by changes in the iron–sulfur cluster geometry or the immediate environment of the cluster. When other electrostatic factors do not change, a change in the

charges on the surface of the HiPIP will have a predictable effect on the redox potential (Banci *et al.*, 1996; Capozzi *et al.*, 1998). A comparison of the three independent molecules in the high-resolution structures reported here also indicated three conserved waters that could provide solvent accessibility to the cluster, which may also be important in the modelling of redox potentials. The relevance of these water molecules as fixed electric dipoles, as opposed to the rapidly reorientating dipoles of bulk water that are responsible for its high dielectric constant, will have to be assessed by NMR measurements of their lifetimes in the binding sites.

The dimer formed between the two crystallographically independent molecules in the rc-WT structure involves a rotation about the inter-cluster vector relative to the previously reported dimer. This observation may require a further analysis of earlier EPR data that were interpreted as suggesting the presence of two different but closely related dimers in solution (Dunham *et al.*, 1991).

EP and PL thank the European Community and the Swiss Bundesamt für Bildung und Wissenschaft for post-doctoral grants (ERB CHBG CT 940731 and BBW 94.0162).

References

Agarwal, A., Li, D. & Cowan, J. A. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 9440–9444.

Averill, B. A., Herskovitz, T., Holm, R. H. & Ibers, J. A. (1973). *J. Am. Chem. Soc.* **95**, 3523–3534.

Babini, E., Borsari, M. & Capozzi, F. (1998). *Inorg. Chim. Acta*, **275**, 230–233.

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

Banci, L., Bertini, I., Dikiy, A., Kastrau, D. H. W., Luchinat, C. & Sompornpisut, P. (1995). *Biochemistry*, **34**, 206–219.

Banci, L., Bertini, I., Savellini, G. G. & Luchinat, C. (1996). *Inorg. Chem.* **35**, 4248–4253.

Beinert, H. (1990). *FASEB J.* **4**, 2483–2491.

Bentrop, D., Bertini, I., Luchinat, C., Piccioli, M. & Rosato, A. (1999). In preparation.

Berg, J. M. & Holm, R. H. (1982). *Metal Ions in Biology*, Vol. 4, *Iron-Sulfur Proteins*, edited by T. G. Spiro, pp. 1–66. New York: Wiley-Interscience.

Bertini, I., Ciurli, S. & Luchinat, C. (1995). *Struct. Bonding*, **83**, 1–53.

Bertini, I., Cowan, J. A., Luchinat, C., Natarajan, K. & Piccioli, M. (1997). *Biochemistry*, **36**, 9332–9339.

Bertini, I., Dikiy, A., Kastrau, D. H. W., Luchinat, C. & Sompornpisut, P. (1995). *Biochemistry*, **34**, 9851–9858.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Cammack, R. (1992). *Adv. Inorg. Chem.* **38**, 281–322.

Capozzi, F., Ciurli, S. & Luchinat, C. (1998). *Struct. Bonding*, **90**, 127–160.

Carter, C. W. Jr (1977). *J. Biol. Chem.* **252**, 7892–7811.

Carter, C. W. Jr, Kraut, J., Freer, S. T. & Alden, R. A. (1974). *J. Biol. Chem.* **249**, 6339–6346.

Carter, C. W. Jr, Kraut, J., Freer, S. T., Alden, R. A., Sieker, L. C., Adman, E. & Jensen, L. H. (1972). *Proc. Natl. Acad. Sci. USA*, **69**, 3526–3529.

Carter, C. W. Jr, Kraut, J., Freer, S. T., Xuong, N., Alden, R. A. & Bartsch R. G. (1974) *J. Biol. Chem.* **249**, 4212–4225.

Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.

Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J. & Moulis, J.-M. (1997). *Biochemistry*, **36**, 16065–16072.

Dunham, W. R., Hagen, W. R., Fee, J. A., Sands, R. H., Dunbar, J. B. & Humblet, C. (1991). *Biochim. Biophys. Acta*, **1079**, 253–262.

Eltis, L. D., Iwagami, S. G. & Smith, M. (1994). *Protein Eng.* **7**, 1145–1150.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Holm, R. H. (1992). *Adv. Inorg. Chem.* **38**, 1–71.

Howard, J. B. & Rees, D. C. (1991). *Adv. Protein Chem.* **42**, 119–280.

Iwagami, S. G., Creagh, A. L., Haynes, C. A., Borsari, M., Felli, I. C., Piccioli, M. & Eltis, L. D. (1995). *Protein Sci.* **4**, 2562–2572.

Kleywegt, G. J. (1996). *Acta Cryst.* D**52**, 842–857.

Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). *Methods Enzymol.* **154**, 367–382.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

McRee, D. E. (1992). *J. Mol. Graph.* **10**, 44–46.

Mascharak, P. K., Hagen, K. S., Spence, J. T. & Holm, R. H. (1983). *Inorg. Chim. Acta*, **80**, 157–170.

Miller, R., DeTitta, G. T., Jones, R., Langs, D. A., Weeks, C. M. & Hauptman, H. (1993). *Science*, **259**, 1430–1433.

Miller, R., Gallo, S., Khalak, H. G. & Weeks, C. M. (1994). *J. Appl. Cryst.* **27**, 613–621.

Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–228.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

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

Sheldrick, G. M. (1982). *Computing in Crystallography*, edited by D. Sayre, pp. 506–514. Oxford University Press.

Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.

Sheldrick, G. M. (1997). *Recent Advances in Phasing: Proceedings of the CCP4 Study Weekend*, edited by K. S. Wilson, G. Davies, A. S. Ashton & S. Bailey, pp. 147–157. Warrington: Daresbury Laboratory.

Sheldrick, G. M. (1998). *SHELX: Applications to Macromolecules.* In *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 401–411. Dordrecht: Kluwer Academic Publishers.

Sheldrick, G. M. & Gould, R. G. (1995). *Acta Cryst.* B**51**, 423–431.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Sieker, L. C., Adman, E. T. & Jensen, L. H. (1972). *Nature (London)*, **235**, 40–42.

Stephens, P. J., Jollie, D. R. & Warshel, A. (1996). *Chem. Rev.* **96**, 2491–2513.

Stout, C. D., Stura, E. A. & McRee, D. E. (1998). *J. Mol. Biol.* **278**, 629–629.

Teng, T. Y. (1990). *J. Appl. Cryst.* **23**, 387–391.

Tickle, I. J., Laskowski, R. A. & Moss, D. S. (1998). *Acta Cryst*. D**54**, 547–557.

Usón, I., Pohl, E., Schneider, T. R., Dauter, Z., Schmidt, A., Fritz, H.-J. & Sheldrick, G. M. (1999). *Acta Cryst*. D**55**, 1150–1167.

Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.