Acta Cryst. (1993). D**49**, 272–281

Structure at pH 6.5 of Ferredoxin I from *Azotobacter vinelandii* at 2.3 Å Resolution

BY E. A. MERRITT, G. H. STOUT, S. TURLEY, L. C. SIEKER AND L. H. JENSEN

Department of Biological Structure, University of Washington, Seattle, WA 98195, *USA*

AND W. H. ORME-JOHNSON

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, *USA*

(*Received* 17 *January* 1991; *accepted* 2 *July* 1992)

Abstract

Ferredoxin I from *Azotobacter vinelandii* (AvFdI) is an iron–sulfur protein composed of 106 amino acids, seven Fe atoms and eight inorganic S* atoms. A crystallographic redetermination of its structure showed the originally reported structure to be incorrect. We report here the crystal structure of AvFdI at pH 6.5. Extensive refinement has led to a final R value of 0.170 for all 6986 non-extinct reflections in the range 10–2.3 Å using a solvent model which includes 98 discrete solvent atoms with occupancies between 0.3 and 1.0 and an average B value of 22.5 Å^2. The first half of the peptide chain closely resembles that of the 55-residue ferredoxin from *Peptococcus aerogenes* (PaFd), while the remainder consists of three turns of helix and a series of loops which form a cap over part of the molecular core. Despite the similarities in structure and surroundings, the corresponding 4Fe4S* clusters in PaFd and AvFdI have strikingly different redox potentials; a possible explanation has been sought in the differing hydration models for the two molecules.

Introduction

Ferredoxin I from *Azotobacter vinelandii* (AvFdI) is an iron–sulfur protein composed of 106 amino acids with $M_r = 12456$. At one point it was thought to be a typical 8Fe ferredoxin having two low-potential 4Fe4S* clusters per molecule (S* designates inorganic sulfur), but ESR and redox studies revealed different midpoint potentials for the two clusters, −420 mV for one of them and +350 mV for the other after ferricyanide treatment (Sweeney, Rabinowitz & Yoch, 1975). The first evidence that the clusters differed geometrically came from a low-resolution (4 Å) X-ray crystallographic study in which one appeared to be a typical 4Fe4S* cluster while the other was smaller and could not be identified with certainty (Stout, 1979). Analysis of the Mössbauer spectrum implicated the −420 mV site as having three Fe atoms, assuming that the total Fe content is ≤8 atoms per mole, while ESR and Mössbauer studies established the second cluster as 4Fe4S* (Emptage *et al.*, 1980). The high redox potential reported for the 4Fe4S* cluster was later found to be an artifact of ferricyanide oxidation; the cluster in native protein instead exhibits a 2+/1+ reduction at ≤ −600 mV (Stephens *et al.*, 1991). In the initial X-ray structure determination at 2.5 Å resolution, the electron density at the smaller cluster was interpreted as a six-atom ring, three Fe atoms alternating with three S* atoms in a twisted boat conformation with an average Fe···Fe distance of 4.1 Å (Stout, Ghosh, Pattabhi & Robbins, 1980; Ghosh, Furey, O'Donnell & Stout, 1981). This model was immediately challenged by extended X-ray absorption fine-structure spectroscopy measurements on a protein containing only 3Fe clusters, identical by Mössbauer spectroscopy to the smaller cluster in AvFdI (Antonio *et al.*, 1982). This showed that the Fe···Fe distance was 2.7 Å (as in 4Fe4S* clusters) and cast doubt on the twisted boat structure proposed by Ghosh, Furey, O'Donnell & Stout (1981), which required a 1.4 Å longer Fe···Fe distance.

Other experimental studies raised serious questions concerning the validity of the proposed crystallographic model for the 3Fe3S* cluster. Thus redox studies of certain 4Fe and 8Fe ferredoxins in which the products were followed by low-temperature ESR and magnetic circular dichroism (Thomson *et al.*, 1981) found that 4Fe and 3Fe clusters are interconvertible and thus closely related. Analyses of Fe and inorganic S* in proteins having 3Fe and 4Fe clusters indicated Fe/S* ratios of 3/4 and 4/4 respectively (Beinert *et al.*, 1983). In fact, the results of the initial crystallographic refinement (Ghosh, O'Donnell, Furey, Robbins & Stout, 1982) raised serious questions regarding the validity of the crystallographic model (Jensen, 1986), and led to an independent redetermination of the AvFdI crystal structure (Stout, Turley, Sieker & Jensen, 1988). The correct model was promptly confirmed by the original investigator (Stout, 1988) who went on to refine

0907-4449/93/020272-10$06.00

© 1993 International Union of Crystallography

the corrected model against a 1.9 Å resolution X-ray data set (Stout, 1989).

The present paper reports the refinement of our model against 2.3 Å resolution data from AvFdI at pH 6.5.

Experimental

Highly purified AvFdI was used to grow crystals by the hanging-drop method at 277 K. Beginning with 1% protein in 0.15 *M* Tris maleate buffer at pH 6.5 in 35% saturated $(NH_4)_2SO_4$ solution, we observed tetragonal bipyramidal crystals growing as the solution approached 70% saturation. Visual inspection of a 10° *h0l* precession photograph identified the crystals as the tetragonal form reported by Stout (1979). The unit-cell parameters are $a = b = 55.55$ and $c = 95.65$ Å based on Cu $K\alpha = 1.5418$ Å, space group either $P4_12_12$ or $P4_32_12$. Solution of the structure established the correct space group as $P4_12_12$ (see below).

Intensity data for the native crystals were collected on a four-circle diffractometer in the $\omega/2\theta$ step-scan mode with a sealed Cu target X-ray tube operated at a power in the range 1.2–1.4 kW (40–44 kV, 30–32 mA). The three crystals used in collecting the native data were truncated bipyramids having minimum and maximum dimensions of 0.225 and 0.875 mm respectively, with volumes ranging from 0.019–0.07 mm^3, and covering the resolution ranges 30–3, 3.5–2.5 and 2.8–2.3 Å. Data for the three crystals were scaled together by means of the common reflections, providing a data set of 7099 reflections in the range 30–2.3 Å. Reflections with $I/\sigma(I) < 3.0$ were corrected by the method of French & Wilson (1978). We applied no cutoff on $I/\sigma(I)$ during refinement; all of the 6986 non-extinct measured reflections from 10 to 2.3 Å were used. Intensity data were also collected for two derivatives: a $PtCl_4^{2-}$ data set with data from 32 to 3.5 Å and a UO_2^{2+} data set from 35 to 3.0 Å.

Structure solution and refinement

Our determination of the pH 6.5 structure of AvFdI was initiated in $P4_32_12$, the space group originally reported (Stout, 1979), and was based on data from the native protein and two derivatives. Friedel pairs were averaged to eliminate the effects of anomalous scattering (Stout, Turley, Sieker & Jensen, 1988). The 3.5 Å resolution electron-density map from multiple isomorphous replacement phases showed two high-density regions, one somewhat more pronounced than the other. The denser of the two was readily fitted by the standard 4Fe4S* cube-like cluster (Adman, Sieker & Jensen, 1973), while the other was fitted by a similar cluster lacking one Fe corner (3Fe4S*). Phases based only on the UO_2^{2+} derivative were extended to 3.0 Å resolution using Wang's density-modification program to resolve the phase ambiguity from a single derivative (Wang, 1985). The 3 Å map was sufficiently improved that the chain could be traced with little ambiguity. One section clearly showed three turns of left-handed helix, suggesting that the proper space group was $P4_12_12$ rather than $P4_32_12$. The choice of $P4_12_12$ was then confirmed by comparing the calculated and observed Bijvoet differences based on the Fe and U anomalous scattering. All phase information was combined, including anomalous scattering from the Fe atoms of the native protein, and after density modification yielded a figure of merit of 0.76 for all 4861 reflections to 2.6 Å. A model built into the 2.6 Å map and assigned an overall $B = 18$ Å^2 gave $R = 0.504$ for all 4758 reflections in the range 10–2.6 Å. Two cycles of rigid-body refinement using *CORELS* (Sussman, 1985) reduced R to 0.409.

The present refinement of AvFdI used *X-PLOR* (Brünger, 1988), locally modified to handle discrete disorder modeled by multiple overlapping conformations with partial occupancy. The starting coordinates were those from the above model (Stout, Turley, Sieker & Jensen, 1988) with $R = 0.466$ for the 10–2.3 Å data. The course of the main part of the refinement is shown in Fig. 1. In cycles 3, 7, 12 and 16 only positional parameters were adjusted. Cycles 10 and 19 included simulated annealing following the slow cooling protocol recommended by Brünger (1988). After cycles 3, 5 and 7 the model was rebuilt section by section based on $2F_o - F_c$ omit maps; after cycles 8, 9, 11, 14, 16 and 18 adjustments were based on $F_o - F_c$ maps with no residues omitted from F_c. R decreased very rapidly at first, from 0.47 to 0.28 in two cycles; then in four cycles, including the rebuilding sessions, it decreased to 0.25. In the next nine cycles, which included one rebuilding and four editing sessions, R still had not dropped below 0.24. Accordingly, 20 fully occupied water oxygen positions were added in the editing session after cycle 16 and another 40 after cycle 18. The added water, along with subsequent thermal parameter refinement reduced R to 0.20. Simulated annealing and one further cycle of positional and thermal parameter refinement brought R to 0.182.

Although at this point R appeared to have reached a satisfactory low value, some features of the model were questionable: the Fe···Fe distances in both the 3Fe and the 4Fe clusters were as much as 0.15–0.20 Å less than the expected distance of 2.75 Å and the S—Fe—S angles in the complexes had been improperly restrained. Furthermore, in order to minimize the number of parameters only two thermal parameters (B) had been used for each residue to this point (one for main-chain atoms and one

for side-chain atoms). No restraints had been applied to the *B* parameters, however, and in several regions of the molecule their values had become excessive.

After cycle 20, therefore, the refinement strategy was changed. *B* parameters were introduced for the individual atoms and restrained to target r.m.s. deviations of 2.5 and 4.0 Å^2 for covalently bonded atoms and 1–3 paired (angle-related) atoms, respectively. Restraints of 2.75 Å were imposed on the nonbonded Fe···Fe distances in both the 3Fe and 4Fe clusters, and the internal S*—Fe—S* and external S$^\gamma$—Fe—S* angles at the Fe atoms were restrained to 104 and 114° respectively, based on values tabulated for several Fe/S clusters by Kissinger, Adman, Sieker & Jensen (1988). Force constants for all distance constraints in the clusters were set to 80 kcal mol^{-1} (335 kJ mol^{-1}) and the weighting assigned to the X-ray residual relative to the energy terms was reduced by a factor of 2. *R* for this more tightly constrained model increased initially to 0.203 and then decreased during refinement to 0.187. At this point the entire model was rebuilt into sequential overlapping omit maps

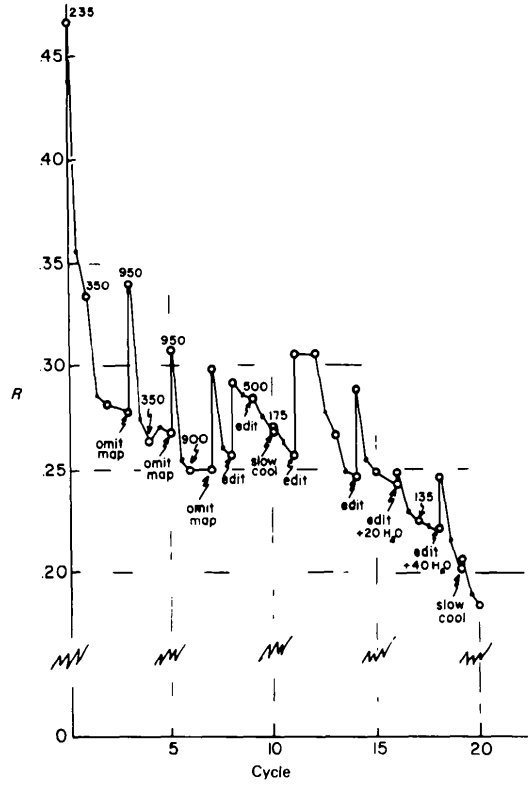

Fig. 1. *R* as a function of cycle number for refinement cycles 1–20. Numbers above the plot are the relative weights assigned to the X-ray residual, and apply until a new weight appears. Open circles (○) represent the beginning or ending of a cycle and the closed circles (●) mark the end of adjusting positional parameters in the cycle and the start of adjusting thermal parameters.

Table 1. *Refinement statistics*

Space group	$P4_12_12$
Unit-cell dimensions (Å)	$a = b = 55.55$, $c = 95.65$
No. of observations used in refinement	All 6986 (10 to 2.3 Å)
Final *R* value	0.170*
	0.201†
	0.224‡
R.m.s. variation in *B* (Å)	
Across bonds	2.8
Across angles	4.7
R.m.s. variation from ideal values	
Bond lengths (Å)	0.022
Bond angles (°)	3.1

* For 868 protein atoms and 98 discrete solvent sites.
† For 868 protein atoms and 21 solvent atoms with $Q \geq 0.7$.
‡ For protein atoms alone (no solvent model).

(prepared by omitting 11 residues from F_c but not from energy terms, refining to convergence and using the resulting phases to calculate a $2F_o - F_c$ map). 101 discrete water O atoms were similarly built into omit maps without regard to the placement of solvents in the earlier model. Finally, anomalous-scattering contributions f' were added for all atoms in the structure. No f'' terms were included, as Friedel measurements had been averaged during data scaling at an earlier stage. This model refined to $R = 0.177$ with improved geometry. At this point restraints on the Fe···Fe distances were removed and three solvent atoms deleted from the model. The cluster geometries relaxed only slightly with the removal of the restraints and *R* dropped to 0.174. The remaining effort in refinement was spent in analyzing residual density appearing in the $F_o - F_c$ map. Improper restraints on the conformation of arginyl side-chain torsion angles were corrected, the side chains of Asp 58 and Glu 66 were each modeled as exhibiting disorder in the form of two conformations with partial occupancies of 0.6 and 0.4, and 65 solvent atoms with $B > 60$ Å^2 were re-assigned occupancies of 0.5. A final refinement with 868 protein atoms, 98 discrete solvent atoms and no constraint on the Fe···Fe distances converged to $R = 0.170$ for all 6986 non-extinct reflections in the range 10.0–2.3 Å. Some further modifications to the solvent model were explored (see below) but this produced no further change in *R*. Table 1 summarizes the crystallographic and stereochemical statistics for the final model; coordinates have been deposited with the Protein Data Bank.*

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory, and are available in machine-readable form from the Protein Data Bank at Brookhaven. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37065 (as microfiche). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Fig. 2 is a plot of R *versus* $\sin\theta/\lambda$ for the AvFdI model. The sharp increase of R for $\sin\theta/\lambda < 0.1$ Å^{-1} is primarily due to the neglect of the solvent continuum in the model. Comparison of the linear part of the plot with the Luzzati (1952) lines for errors of 0.2 and 0.3 Å indicates an expected average error of approximately 0.24 Å in the determined atomic coordinates.

The φ, ψ angles in a Ramachandran plot provide a validity check of the chain folding (Weaver, Tronrud, Nicholson & Matthews, 1990). Fig. 3 shows that most values for the present structural model fall in acceptable regions of the plot. Of the five outliers with $+\varphi$, all but Gly 96 lie close to the allowed left-helical region of the plot.

The ultimate validity check is to compare two independently determined and refined models of the same structure. This is possible for AvFdI using the 1.9 Å pH 8.0 structure (Stout, 1989) and the present refinement at 2.3 Å of the pH 6.5 structure. After a rigid-body fit to correct for the slight difference in cell parameters, the overall r.m.s. difference in the pH 6.5 and 8.0 models is 0.17 Å for backbone atoms only (C, N, C^{α}), and 0.96 Å for all other protein atoms (all comparisons with the pH 8.0 model used coordinate set 4FD1 deposited with the Protein Data Bank). Assuming the same average error in both determinations, then from the Luzzati estimate above we would expect the difference in two independent determinations of the same structure to be $2^{1/2}(0.24 \text{ Å}) = 0.34$ Å.

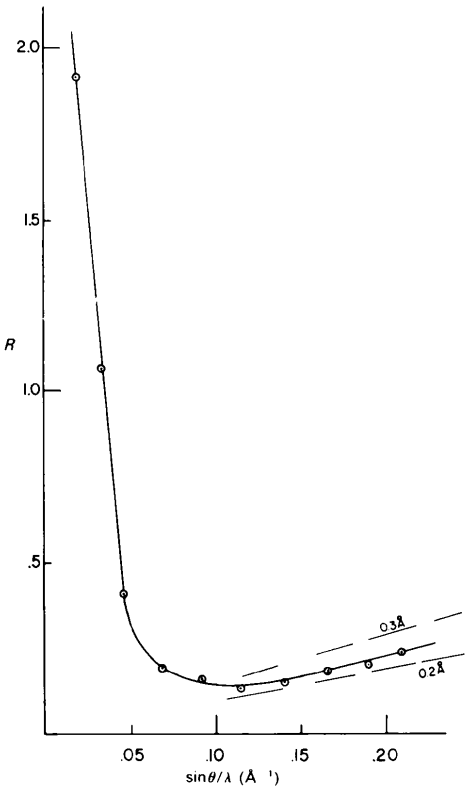

Fig. 2. R as a function of $\sin\theta/\lambda$. Luzzati lines are shown for coordinate errors of 0.2 and 0.3 Å

Comparisons and results

Most of the structural features of the AvFdI molecule have been covered in considerable detail (Stout, 1989) and need not be discussed further here. We begin by comparing AvFdI to ferredoxin from *Peptococcus aerogenes* (PaFd; Adman, Sieker & Jensen, 1973, 1976).

The first half of the 106-residue AvFdI sequence is similar to that of the 55-residue PaFd molecule. Structural comparison to PaFd used coordinates from a recent re-refinement which included an additional Cys residue at position 22 (E. T. Adman, personal communication) missed in the original protein sequence determination (Tsunoda, Yasunobu & Whiteley, 1968). Based on this revised sequence, the structural homology between the two ferredoxins is such that residues 1–8, 9–23 and 29–54 in PaFd correspond to residues 1–8, 11–25 and 32–57 in AvFdI. For these 49 residues the r.m.s. difference between the 147 corresponding backbone atoms is 0.78 Å. This corresponds to a sequence alignment in which two residues are inserted after Ile 9 of the PaFd sequence and another insertion follows Ala 27 (Fig. 4). Beyond Asp 58 the AvFdI chain bends into a three-turn helix from residue 64 to 75, then loops around one side of the body of the molecule, crossing over Cys 11 and terminating with Arg 106

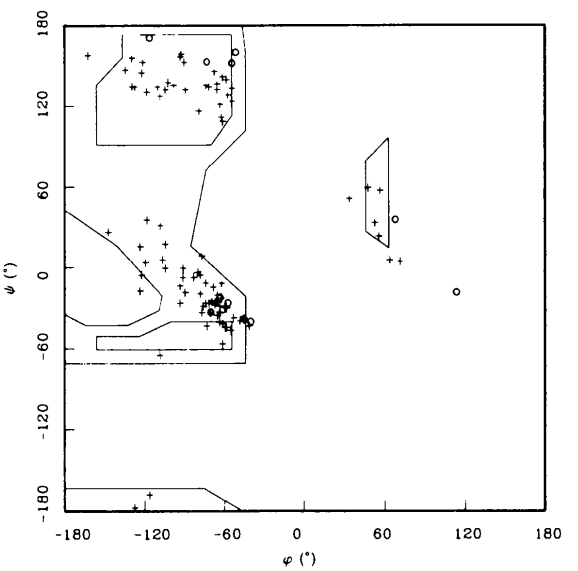

Fig. 3. Ramachandran plot for the AvFdI structure at the end of refinement. Open circles represent Gly and Pro residues.

near the end of the three-turn helix. The superposition of the AvFdI and the PaFd chains is shown in Fig. 5.

Difference in the pH 6.5 and 8.0 structures

As is clearly visible in Fig. 6, the 2.3 Å pH 6.5 model and the 1.9 Å pH 8.0 model are essentially identical except for a relatively small number of side chains which assume different conformations in the two models. Of these Gln 52, Asp 90, Glu 92 and Lys 98 in particular are surface residues with high B values in both models, so the observed differences are not likely to be significant.

```
                    10        20        30        40        50      55
PaFd:    AYVINDSC -- IACGACKPECPVNCI(-EQGAI)YAIDADSCIDCGSCASVCPVGAPNPE(D)
AvFdI:   AFVVTDNC(IK)CKYTDCVEVCPVDCF(YEGPNF)LVIHPDECIDCALCEPECPAQAIFSE(D)EV
                    10        20        30        40        50        60

PaFd:
AvFdI:   PEDMQEFIQLNAELAEVWPNITEKKDPLPDAEDWDGVKGKLQHLER
                   70        80        90       100     106
```

Fig. 4. Alignment of PaFd and AvFdI sequences based on structural homology. Residues in parentheses were not used in calculating the r.m.s. difference in the two backbones.

Fig. 5. Superposition of the backbone atoms from AvFdI (lighter shading) and PaFd (darker shading). Except for the small loop above and somewhat to the right of the 3Fe cluster (AvFdI residues 9–10) and the strand at the right of the figure (AvFdI residues 26–31) in which AvFdI contains an additional residue, the chains are very nearly identical over the length of the shorter PaFd chain.

The differences most clearly attributable to pH involve the local environment of Glu 18 and Glu 83 (Fig. 7). The present model exhibits a salt bridge between these two residues; in the pH 8.0 structure these side chains are instead directed away from each other, and a solvent molecule lies roughly in the position occupied by the Glu 18 carboxylate in the pH 6.5 structure. This solvent is, in fact, the only one of the 21 solvents in the pH 8.0 model which is not also observed in the present refinement. In this same region of the pH 6.5 model, Gln 69 has shifted to

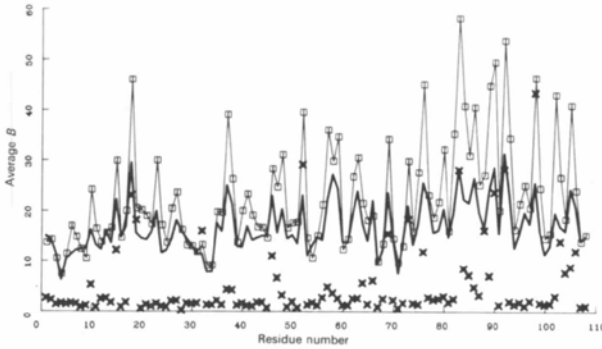

Fig. 6. Average thermal parameter B for side-chain atoms at each residue in the pH 6.5 (□; this paper) and pH 8.0 (—; Stout, 1989) structures of AvFdI. Also shown are the r.m.s. differences between the two sets of coordinates for the side chain of each residue (×; Å × 10). The overall r.m.s. difference in coordinates is 0.30 Å for backbone atoms and 0.99 Å for side-chain atoms.

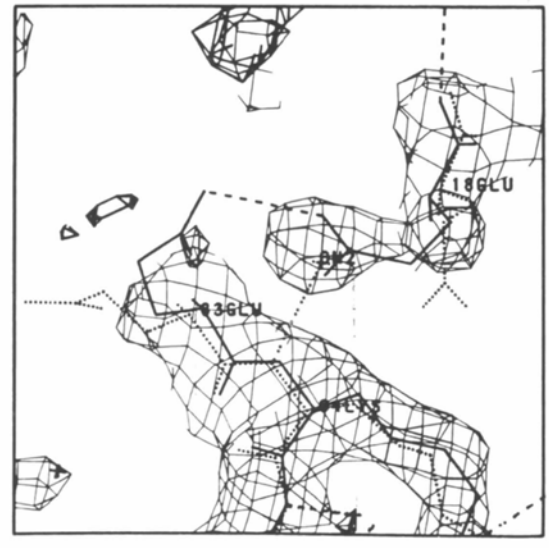

Fig. 7. Electron density in the region of the salt bridge between Glu 18 and Glu 83. The solid bonds indicate the positions observed in the present pH 6.5 structure; the dotted bonds are the positions reported for the pH 8.0 structure. The electron density is from an $|F_o - F_c|$ map calculated after omitting residues 17–19, 82–84, and nearby solvent molecules from the F_c.

form hydrogen-bonding interactions with the Glu 18 carbonyl oxygen, the secondary conformation assigned to the Glu 66 side chain, and to a solvent molecule. In order to test whether the pH 8.0 conformation for Glu 18 and its local environment would also be consistent with our pH 6.5 diffraction data, we replaced the refined coordinates for the Glu 18 side chain with those from the pH 8.0 structure, added a solvent molecule corresponding to that in the pH 8.0 structure, and subjected the entire model to further positional and B refinement. This test supported the original placement, since refinement of the trial alternative conformation for Glu 18 caused its B values to rise to the highest in the entire structure ($B > 70$ Å^2), the overall R increased slightly, and an $F_o - F_c$ map showed negative density at the new location of Glu 18 and residual positive density for the original placement despite the presence of the additional solvent molecule at that site.

The side chains of Asp 58 and Glu 66 are fitted as a 60/40 mixture of two alternate conformations in the present refinement. For both residues the primary conformation is the same as in the pH 8.0 model.

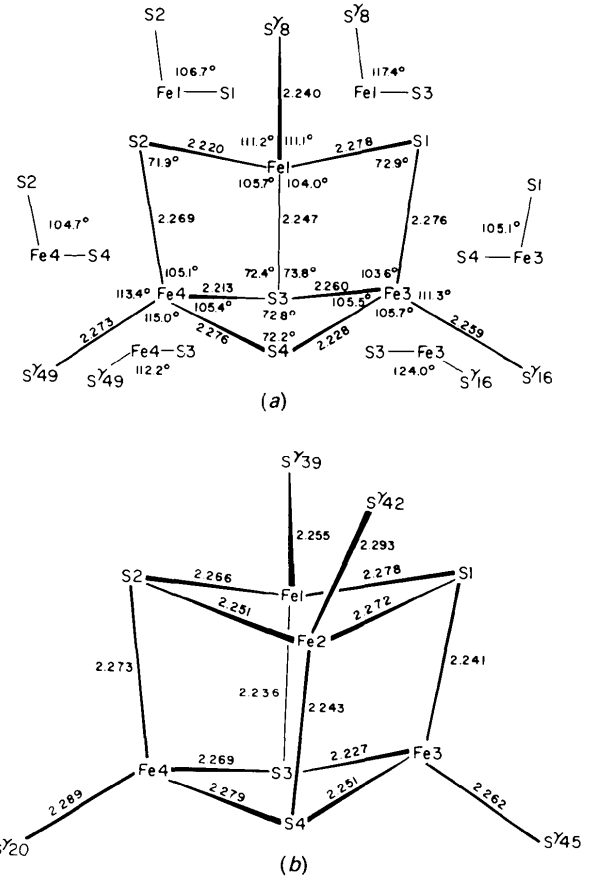

Fig. 8. (*a*) Bond lengths (Å) and angles ( ) in the 3Fe complex of the pH 6.5 AvFdI model. (*b*) Bond lengths (Å) in the 4Fe complex of the pH 6.5 AvFdI model.

Table 2. *Bond angles* (°) *in the* 4Fe *complex*

S1—Fe1—S2	101.8	Sγ39—Fe1—S1	105.2	Fe1—S1—Fe2	73.9
S1—Fe1—S3	107.8	Sγ39—Fe1—S2	112.4	Fe1—S1—Fe3	69.9
S2—Fe1—S3	105.3	Sγ39—Fe1—S3	122.3	Fe2—S1—Fe3	74.2
S1—Fe2—S2	102.5	Sγ42—Fe2—S1	115.8	Fe1—S2—Fe2	74.3
S1—Fe2—S4	102.5	Sγ42—Fe2—S2	124.1	Fe1—S2—Fe4	73.0
S2—Fe2—S4	108.0	Sγ42—Fe2—S4	101.9	Fe2—S2—Fe4	72.1
S1—Fe3—S3	109.5	Sγ45—Fe3—S1	115.5	Fe1—S3—Fe3	70.8
S1—Fe3—S4	103.2	Sγ45—Fe3—S3	119.0	Fe1—S3—Fe4	73.6
S3—Fe3—S4	100.8	Sγ45—Fe3—S4	106.4	Fe3—S3—Fe4	77.0
S2—Fe4—S3	104.0	Sγ20—Fe4—S2	115.3	Fe2—S4—Fe3	74.6
S2—Fe4—S4	106.1	Sγ20—Fe4—S3	119.8	Fe2—S4—Fe4	72.2
S3—Fe4—S4	98.7	Sγ20—Fe4—S4	111.0	Fe3—S4—Fe4	76.3

Fe—S *complexes*

Fig. 8 shows the labeling of the atoms in the 3Fe and 4Fe complexes in AvFdI along with the Fe—S bond lengths in both complexes and the bond angles in the 3Fe complex. The bond angles in the 4Fe complex are listed in Table 2. Comparison of the coordinating cysteine residues in the two 4Fe complexes of the PaFd structure (complex I: 8, 11, 14, 46; complex II: 36, 39, 42, 18) with the sequence of the two proteins in Fig. 4 shows that the 3Fe and 4Fe complexes in AvFdI correspond to complexes I and II respectively in PaFd (Adman, Sieker & Jensen, 1973, 1976). It is Fe2 in complex I of PaFd that is missing in the 3Fe complex in AvFdI, resulting in the free SH group in the Cys 11 side chain.

The target values used for the stereochemical restraint of the complexes in the present refinement were based on the 1.7 Å refinement of the model for ferredoxin from *Desulfovibrio gigas* (DgFd) in which the positional parameters of all atoms in the 3Fe4S* 3S$^\gamma$ complex were unrestrained (Kissinger, Sieker, Adman & Jensen, 1991). In the final stages of the present refinement a distinction was made between the internal S*—Fe—S* angles, which were restrained to a target angle of 104 , and the external S*—Fe—S$^\gamma$ angles, which were restrained to a target angle of 114 . In the final stages no restraints were imposed on the non-bonded distances in the complexes. While subclasses of both distance and angular parameters may exist within the restraint classes chosen, we regard the restraints used as adequate, at least at the resolution of the present refinement. Fig. 9 shows the distribution of the individual bond lengths, angles and restraint targets for the clusters. Note the greater range of values for the external S$^\gamma$—Fe—S* angles compared to the internal S*—Fe—S* angles.

Tables 3 and 4 present the geometry and possible NH—S hydrogen-bonding interactions of the liganding sulfur atoms. None of these values differs significantly enough from those seen in the FeS complexes of other ferredoxins to easily explain the

observed differences in reduction potentials on this basis.

Solvent model

The ordered water in crystalline proteins is an important component of the structure, playing a key role in the hydrogen-bonding system that laces the protein molecules together. In redetermining the solvent model for AvFdI in the final refinement cycles, we included 101 fully occupied sites that were within hydrogen-bonding distance of candidate donor or acceptor atoms. Three of these were subsequently deleted when they did not refine satisfactorily, leaving 98 sites. Of these, 84 were first-shell solvents with a direct hydrogen-bonding interaction to a protein atom, while 14 were second-shell solvents with hydrogen bonds only to other solvent oxygen atoms. At this point a number of these solvent sites were clearly visible in omit maps, were well behaved during refinement and were apparently anchored by hydrogen-bonding networks, yet their B parameters had adjusted to large values during refinement. It seemed possible that a better interpretation of the density was to model solvent sites with lower occupancy, and presumably with a correspondingly lower B value. Therefore, the occupancy Q was reduced to 0.5 for 65 solvent sites whose B value had previously exceeded 60 Å^2, and positional and thermal parameters were again refined to convergence. The average refined B value was 29.7 Å^2, with an r.m.s. variation of 9.6 Å^2 and a maximum of 56 Å^2. Fig. 10 is a plot of B *versus* peak height for the 98 solvent O sites at this stage in refinement. Note that the plot falls into two well defined sections, suggesting that even in refining against data at 2.3 Å resolution one is justified in using ΔQ values smaller than 0.5 (*cf.* Kundrot & Richards, 1987; Jensen, 1990). To test this we omitted ten solvent atoms at a time from F_c in a series of omit maps and tabulated the electron-density peak height at the site of the omitted solvent oxygens. All modeled solvents reappeared at their previously assigned site with density greater than 3σ above the mean density in the corresponding omit map. New occupancies were then assigned in increments of $\Delta Q = 0.1$ relative to $Q = 1.0$ for the average peak height of the top five solvent peaks. This procedure yielded occupancies of $Q = 1.0$ (5 sites), $Q = 0.9$ (2 sites), $Q = 0.8$ (5 sites), $Q = 0.7$ (9 sites), $Q = 0.6$ (13 sites), $Q = 0.5$ (31 sites), $Q = 0.4$ (30 sites) and $Q = 0.3$ (3 sites). The thermal parameter B was reset to 21 Å^2 for all solvent sites and then re-refined to convergence with the new occupancies. This resulted in an averge solvent B of 22.5 Å^2, with an r.m.s. variation of 6.4 Å^2 and a

Table 3. *Possible* NH—S *hydrogen bonds* (Å, °)

Acceptor	Donor NH	Distance	NH—S angle	Distance for AvFdI*
Cys 8 S^γ	Leu 32 N	3.55	152.50	3.56
Cys 11 S^γ	Lys 100 N^ζ	3.56		3.46
Cys 39 S^γ	Phe 2 N	3.50	153.43	3.40
Cys 39 S^γ	Asp 41 N	3.32	152.10	3.30
Cys 42 S^γ	Leu 44 N	3.46	157.47	3.42
Cys 49 S^γ	Ala 51 N	3.27	140.29	3.37
Cys 49 S^γ	Ala 53 N	3.42	171.89	3.30
FeS3 107 S1	Thr 14 N	3.19	149.90	3.15
FeS3 107 S2	Tyr 13 N	3.44	171.89	3.59
FeS3 107 S4	Asp 15 N	3.39	139.01	3.48
FeS3 107 S4	Cys 16 N	3.48	165.02	3.43
FeS4 108 S1	Ile 40 N	3.35	155.74	3.46
FeS4 108 S4	Ala 43 N	3.24	142.94	3.22

* Values for AvFdI at pH 8.0 calculated from the Protein Data Bank coordinate set 4FD1 (Stout, 1989).

Table 4. *Torsion angles* (°) *for* Fe—S^γ—C^β—C^α *bonds in FdIAv* FeS *structures*

FeS 107 S^γ 8	85.61	FeS 108 S^γ 42	−64.31
FeS 107 S^γ 16	−114.56	FeS 108 S^γ 45	−108.47
FeS 107 S^γ 49	76.46	FeS 108 S^γ 20	68.68
FeS 108 S^γ 39	62.93		

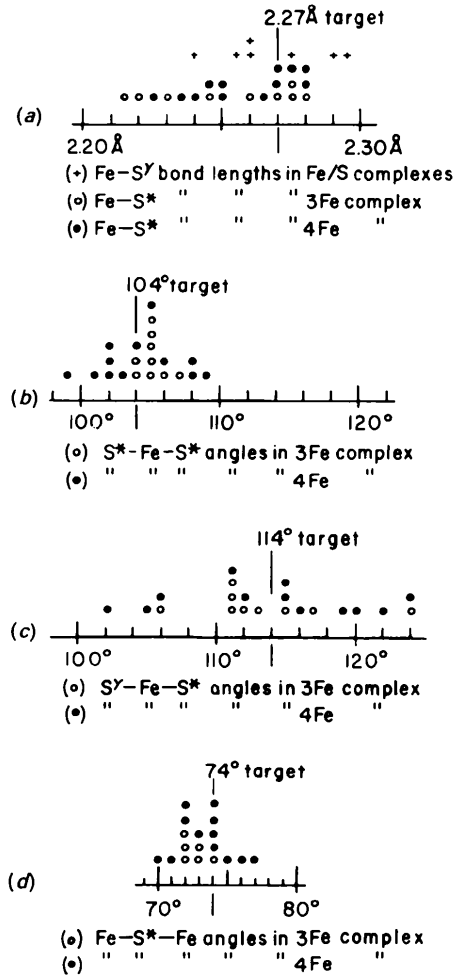

Fig. 9. (*a*)–(*d*) Distribution of bond lengths and angles in the 3Fe and 4Fe complexes of pH 6.5 AvFdI; lengths are rounded to 0.005 Å, angles are rounded to integral values.

maximum of 38 Å^2. The R factor remained at 0.170 although the revised water model comprised the equivalent of 12.2 fewer water molecules (53.3 *versus* 65.5). No constraints on the solvent B values were applied at any point during refinement.

Discussion

The rationale for such a detailed treatment of the solvent structure was to increase the interpretability and significance of minor features in the residual electron density shown by the $|F_o - F_c|$ synthesis. For example, the residual density near side chains Asp 58 and Glu 66 in early maps did not permit reliable analysis, whereas after inclusion of a detailed solvent model the density seemed clearly to indicate discrete disordered conformations for these two side chains. At the end of refinement four peaks with $\rho >$ 4 r.m.s.(ρ) remained in the $|F_o - F_c|$ residual density map. The three largest of these are plausible candidates for second-shell solvent sites. The fourth site lies adjacent to the 4Fe cluster, 1.7 Å from S$^\gamma$ of Cys 24 (the nearest atom). Considered in isolation the side-chain torsion angles of Cys 24 could be adjusted in the model to fit this peak as an alternate location for S$^\gamma$. Such a conformation for Cys 24 could not be accommodated without concomitant adjustment of the protein main-chain atoms in Glu 38 and Cys 39, but some support for this possibility is lent by the fact that the B values for these main-chain atoms are the highest in the first half of the protein backbone. This residual peak is of interest in the context of recently reported studies of site-directed mutants of AvFdI in which Cys 20 is replaced by Ala (Martín *et al.*, 1990). In that C20$^\alpha$ mutant a conformational shift of Cys 24 allows it to replace the deliberately removed Cys 20 as the fourth ligand in the 4Fe4S cluster without substantial deformation of the protein's peptide backbone. If the residual density which we observe does in fact indicate a small component of static disorder in the crystal structure at this site, then this may be evidence that the conformational flexibility of Cys 24 exists even in the native protein.

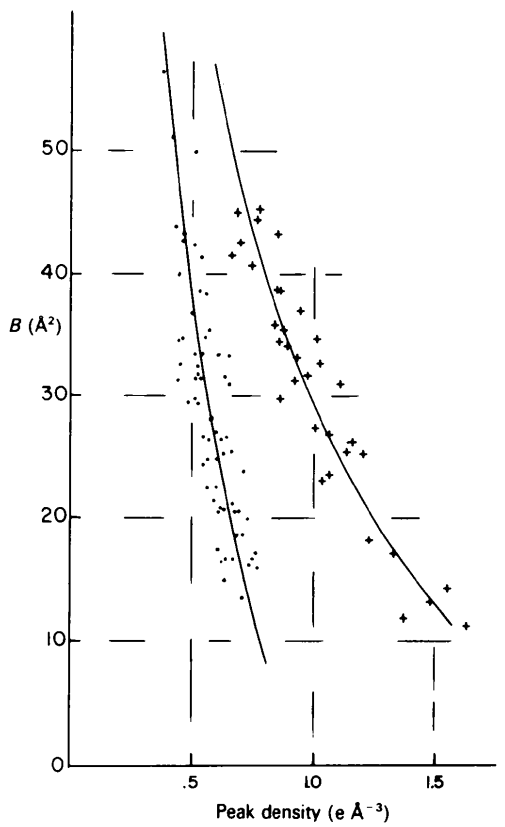

Fig. 10. B as a function of peak electron density for the 98 discrete water sites at a point during refinement at which all solvent atoms were assigned an occupancy of 1.0 (+) or 0.5 (•).

Despite the missing Fe atom, the 3Fe complex in AvFdI has essentially the same redox potential (−420 mV) as both 4Fe complexes in PaFd. By contrast the 4Fe complex in AvFdI (corresponding to complex II in PaFd) undergoes reduction at ≤ -600 mV (Morgan *et al.*, 1984; Stephens *et al.*, 1991) even though the backbone and cluster geometries in the two proteins are nearly superimposable. Backes *et al.* (1991) speculate that the anomalously low potential for AvFdI is due to differences in the microenvironment which limit solvent accessibility to the cluster. With this in mind, we compared the apparent solvent accessibility in AvFdI and PaFd. The substantially longer peptide chain in AvFdI might be expected to limit solvent access to the 4Fe cluster. However, inspection of the solvent accessible surface for AvFdI reveals two notable invaginations pointing toward the base of the 4Fe cluster (Fig. 11). In both cases the closest approach to the 4Fe4S cluster is approximately 7 Å. In the present crystal structure the first of these invaginations is occupied by water 109 ($Q = 1.0$, $B =$ 11.4 Å^2) and water 110 ($Q = 0.8$, $B = 13.9$ Å^2). Its inner surface follows very closely the exterior suface of the smaller PaFd. The solvent molecules in this depression are separated from the 4FeS cluster by the hydrophobic side chains of Ile 34 and Phe 2, which correspond almost exactly in placement to Ile 31 and Tyr 2 of PaFd. There would, therefore, seem to be no difference in the solvent accessibility of the two proteins from this direction. The second invagination, however, is harder to analyze. Its tip is occupied by water 111 ($Q = 1.0$, $B = 14.2$ Å^2) and water 123 ($Q = 0.7$, $B = 12.4$ Å^2). Water 111 is hydrogen bonded to the amide N of Phe 25, to the carbonyl O of Ile 81 (for which there is no homologue in PaFd), and to water 123. The volume between

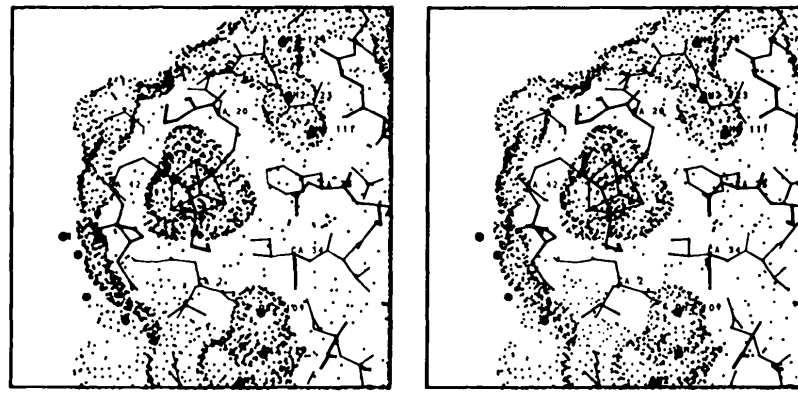

Fig. 11. Solvent accessible surface generated using a probe radius of 1.4 Å and considering only the atoms belonging to the polypeptide chain (*i.e.* the Fe and S* atoms were excluded). In the vicinity of the 4Fe cluster there are two substantial invaginations of the surface occupied by discretely ordered water molecules. Either or both may conceivably affect the electronic environment of the 4Fe cluster. Both OH2 111 and OH2 123 (in the upper invagination) lie within 5 Å of C^{β} and S^{γ} of the liganding residue Cys 20. However, OH2 109 (near the bottom of the figure) is largely shielded from the cluster by the hydrophobic side chain of Ile 34.

water 111 and the cluster is largely occupied by the side chain of Cys 20. Since S^{γ} 20 is a cluster ligand, it is possible that the solvent proximity may be influencing the electronic behavior of the 4Fe4S cluster. While PaFd does have a water molecule (*OW* 108) at roughly the position of water 111 in the present structure, in PaFd the side chain of Asn 12 is interposed between it and the bulk solvent region on the protein's exterior; *OW* 108 is apparently hydrogen bonded to the amide N of Ile 23, the carbonyl O of Lys 15 and Ile 23, and to O^{δ} of Asn 21. Given the likely flexibility of the Asn 21 side chain in solution, one cannot say whether the conformations seen in the two structures actually indicate a significant difference in solvent accessibility.

The virtually identical Fe/S cluster geometries observed for the pH 6.5 and the pH 8.0 AvFdI structures are consistent with spectroscopic studies. Both of these crystal structures are of the as-isolated (oxidized) form of the protein ($[4Fe4S]^{2+}$, $[3Fe4S]^{+}$). No change in the low-temperature magnetic circular dichroism (MCD) spectrum is observed for oxidized AvFdI over the pH range 6.3 to 8.5 (Johnson, Bennett, Fee & Sweeney, 1987; Morgan *et al.*, 1984). By contrast, when the 3Fe cluster is reduced there is a marked pH dependence of the MCD spectrum indicating a reversible change in the electronic structure over this same pH range (Stephens *et al.*, 1991; Johnson, Bennett, Fee & Sweeney, 1987). The site of the presumed protonation event associated with this change is not determinable from the crystal structures.

Concluding remarks

The folding of the chain in the present model for AvFdI based on refining against 2.3 Å resolution X-ray data is identical within experimental error to that of the 1.9 Å model reported by Stout (1989). Most of the side chains in the two models are essentially the same, but the difference in conformation in one region of the molecule may be attributable to the difference in pH at which the crystals were grown.

In comparing the plots of R *versus* $\sin\theta/\lambda$ for the final structure factors from the lower and higher resolution refinements [*cf.* Fig. 2 in this paper and Table 2 in Stout (1989)] we note that the precision of the two models is essentially the same, and for the main-chain atoms can be taken as $0.17\ \text{Å}/2^{1/2} = 0.12$ Å for the positional parameters. A contributing factor to the precision of the lower resolution study is the more extensive solvent model.

References

ADMAN, E. T., SIEKER, L. C. & JENSEN, L. H. (1973). *J. Biol. Chem.* **248**, 3987–3996.

ADMAN, E. T., SIEKER, L. C. & JENSEN, L. H. (1976). *J. Biol. Chem.* **251**, 3801–3806.

ANTONIO, M. R., AVERILL, B. A., MOURA, I., MOURA, J. J. G., ORME-JOHNSON, W. H., TEO, B.-K. & XAVIER, A. V. (1982). *J. Biol. Chem.* **257**, 6646–6649.

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.

BEINERT, H., EMPTAGE, M. H., DREYER, J. L., SCOTT, R. A., HAHN, J. E., HODGSON, K. O. & THOMSON, A. J. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 393–396.

BRÜNGER, A. T. (1988). *X-PLOR Manual.* Version 1.5. Yale Univ., USA.

EMPTAGE, M. H., KENT, T. A., HUYNH, B. H., RAWLINGS, J., ORME-JOHNSON, W. H. & MÜNCK, E. (1980). *J. Biol. Chem.* **255**, 1793–1796.

FRENCH, S. & WILSON, K. (1978). *Acta Cryst.* A**34**, 517–525.

GHOSH, D., FUREY, W. F. JR, O'DONNELL, S. & STOUT, C. D. (1981). *J. Biol. Chem.* **256**, 4185–4192.

GHOSH, D., O'DONNELL, S., FUREY, W. F. JR, ROBBINS, A. H. & STOUT, C. D. (1982). *J. Mol. Biol.* **158**, 73–109.

JENSEN, L. H. (1986). *Iron Sulfur Protein Research*, edited by H. MATSUBARA, Y. KATSUBE & K. WADA, pp. 3 21. Berlin: Springer-Verlag.

JENSEN, L. H. (1990). *Acta Cryst.* B**46**, 650–653.

JOHNSON, M. K., BENNETT, D. E., FEE, J. A. & SWEENEY, W. V. (1987). *Biochim. Biophys. Acta*, **911**, 81–94.

KISSINGER, C. R., ADMAN, E. T., SIEKER, L. C. & JENSEN, L. H. (1988). *J. Am. Chem. Soc.* **110**, 8721–8723.

KISSINGER, C. R., SIEKER, L. C., ADMAN, E. T. & JENSEN, L. H. (1991). *J. Mol. Biol.* **219**, 693–715.

KUNDROT, C. E. & RICHARDS, F. M. (1987). *Acta Cryst.* B**43**, 544–547.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 803–805.

MARTÍN, A. E., BURGESS, B. K., STOUT, C. D., CASH, V. L., DEAN, D. R., JENSEN, G. M. & STEPHENS, P. J. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 598–602.

MORGAN, T. V., STEPHENS, P. J., DEVLIN, F. J., STOUT, C. D., MELIS, K. A. & BURGESS, B. K. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 1931–1935.

STEPHENS, P. J., JENSEN, G. M., DEVLIN, F. J., MORGAN, T. V., STOUT, C. D., MARTÍN, A. E. & BURGESS, B. K. (1991). *Biochemistry*, **30**, 3200–3209.

STOUT, C. D. (1979). *Nature (London)*, **279**, 83–84.

STOUT, C. D. (1988). *J. Biol. Chem.* **263**, 9256–9260.

STOUT, C. D. (1989). *J. Mol. Biol.* **205**, 545–555.

STOUT, C. D., GHOSH, D., PATTABHI, V. & ROBBINS, A. H. (1980). *J. Biol. Chem.* **255**, 1797–1800.

STOUT, G. H., TURLEY, S., SIEKER, L. C. & JENSEN, L. H. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 1020–1022.

SUSSMAN, J. L. (1985). *Methods Enzymol.* **115**, 271–303.

SWEENEY, W. V., RABINOWITZ, J. C. & YOCH, D. C. (1975). *J. Biol. Chem.* **250**, 7842–7847.

THOMSON, A. J., ROBINSON, A. E., JOHNSON, M. K., CAMMACK, R., RAO, K. K. & HALL, D. O. (1981). *Biochem. Biophys. Acta*, **637**, 423–432.

TSUNODA, J. N., YASUNOBU, K. T. & WHITELEY, H. R. (1968). *J. Biol. Chem.* **243**, 6262–6272.

WANG, B. C. (1985). *Methods Enzymol.* **115**, 90–112.

WEAVER, L. H., TRONRUD, D. E., NICHOLSON, H. & MATTHEWS, B. W. (1990). *Curr. Sci.* **59**, 833–837.