Acta Cryst. (1995). D**51**, 318–330

Structure of the Trigonal form of Recombinant Oxidized Flavodoxin from *Anabaena* 7120 at 1.40 Å Resolution*

BY BRIAN M. BURKHART AND B. RAMAKRISHNAN

Biological Macromolecular Structure, Departments of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA

HONGGAO YAN,† ROSS J. REEDSTROM‡ AND JOHN L. MARKLEY

Department of Biochemistry, College of Agriculture and Life Sciences, 420 Henry Hall, University of Wisconsin-Madison, Wisconsin 53706, USA

NEIL A. STRAUS

Department of Botany, University of Toronto, Toronto, Ontario M5S 1A1, Canada

AND M. SUNDARALINGAM§

Biological Macromolecular Structure, Departments of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA

(*Received* 21 *January* 1994; *accepted* 14 *October* 1994)

Abstract

The oxidized recombinant flavodoxin from the cyanobacterium *Anabaena* 7120 has been crystallized in a trigonal form. The recombinant protein has an identical primary structure to that purified directly from *Anabaena*, which functions as a substitute for ferredoxin in an iron-deficient environment for electron transfer from photosystem I to ferredoxin–$NADP^+$ reductase. X-ray data to 1.40 Å were collected on a Siemens area detector. Of the 311 379 reflections collected, 36 069 reflections were unique in space group $P3_121$ ($a = 55.36$, $c = 102.59$ Å) with an R_{merge} of 3.8%. The structure was solved by molecular replacement using coordinates from the wild-type monoclinic structure previously solved in this laboratory [Rao, Shaffie, Yu, Satyshur, Stockman & Markley (1992). *Protein Sci.* **1**, 1413–1427]. The structure was refined with *X-PLOR* and *SHELXL*93 to a crystallographic R factor of 13.9% for 32 963 reflections with $I > 2\sigma(I)$. The final structure contains 2767 atoms including 31 flavin mononucleotide (FMN) atoms, 299 water molecules, and one sulfate ion. The protein is comprised of a central five-stranded β-sheet surrounded by five helices and binds a single molecule of FMN at the C-terminus of the sheet. The trigonal protein structure and the crystal packing are compared with the monoclinic wild-type protein. Helix $\alpha 3$ in this structure is less distorted than in the monoclinic structure and shows additional hydrogen bonds in the N-terminal portion of the helix. The trigonal structure is extensively hydrogen bonded in three major areas with neighboring molecules compared with five regions in the monoclinic structure, but using significantly fewer hydrogen bonds to stabilize the lattice. There are several hydrogen bonds to the amide groups from water molecules several of which stabilize and extend the ends of the β-sheet.

Introduction

Flavodoxins are small (14–23 kDa) acidic α/β proteins. They have been found in a wide variety of prokaryotes including strict anaerobes (*A. vinelandii*, Klugkist, Voorberg, Haaker & Veeger, 1986), facultative anaerobes (*E. coli*, Vetter & Knappe, 1971), and cyanobacteria (*A. nidulans*, Smillie & Entsch, 1971). They have also been found in some eukaryotic algae (*C. crispus*, *C. fusca*, Fitzgerald, Husain & Rogers, 1978; Zumft & Spiller, 1971). In the photosynthetic bacteria, *Anabaena* 7120, they replace ferredoxins in the transfer of electrons from photosystem I (PS I) to ferredoxin–$NADP^+$ reductase under iron-deficient conditions (Sykes & Rogers, 1984). In other organisms, they are variously involved in the metabolism of such diverse molecules as nitrogen (Klugkist *et al.*, 1986), methionine (Fujii & Huennekens, 1974) and pyruvate (Ogata, Miyazaki & Akagi, 1986).

* Presented at the American Crystallographic Association meeting, Pittsburgh, Pennsylvania, August 9–14, 1993; Abstract PC29.

† Present address, Department of Biochemistry, Michigan State University, 302 Biochemistry Building, East Lansing, MI 48824, USA.

‡ Present address, School of Pharmacy, University of Wisconsin-Madison, Room 4320 Chamberlin Hall, Madison, WI 53706, USA.

§ Author to whom correspondence should be sent.

Each flavodoxin molecule non-covalently binds a single molecule of flavin mononucleotide (FMN) as a prosthetic cofactor with no iron–sulfur clusters or other metal centers present. FMN acts as the storage center while transporting electrons between PS I and FNR, delocalizing the electrons added upon reduction. The importance of iron in flavodoxin regulation and expression is discussed elsewhere in detail (Mayhew & Ludwig, 1975; Mayhew & Tollin, 1993).

The FMN molecule can exist in three different oxidation states: oxidized (ox), the one-electron reduced semiquinone (sq) and the two-electron reduced hydroquinone (hq), but *in vivo* shuttles between the sq and hq reduced states. In the long-chain flavodoxins, the ox/sq reduction potential (E_2) is similar to that for free FMN in solution at pH 7.0 (−238 mV), but the sq/hq couple (E_1) is approximately 2.5 times lower. The net effect of the protein is to separate the potentials E_2 and E_1, creating a sizeable energy gap upon oxidation or reduction, with the sq form stabilized over the hq form. The mechanisms by which flavodoxins acting as scaffolds of the FMN molecule modulate the reduction potentials, stabilize the reduced forms of FMN and insure that the electrons are transferred between the proper molecules are unknown.

The flavodoxin from the cyanobacterium *Anabaena* 7120 is a 169-amino-acid (19 kDa) protein (Leonhardt & Straus, 1989). It is homologous to other long-chain flavodoxins, a family of proteins whose members have varied redox potentials, which differ from those of free FMN. For example, E_1, the redox potential for the first (oxidized to semiquinone) reduction, is −196 mV in the flavodoxin from *Anabaena* 7120 (Paulsen, Stankovich, Stockman & Markley, 1990), −50 mV in the flavodoxin from the cyanobacterium *Synechoccus lividus* (Crespi, Norris, Bays & Katz, 1973), and −238 mV in free flavin (Draper & Ingraham, 1960). A better understanding of how the details of protein structure moderate the potential of the flavin cofactor has been an ongoing goal in this work.

The general structure (Fig. 1) of all the flavodoxins are similar in that they contain a central five-stranded β-sheet core surrounded by five helices, three on one side and two on the other. Two distinct structural forms of the protein are found in nature, long chain and short chain. The long-chain flavodoxins contain a loop of approximately 20 amino acids inserted in the middle of the fifth strand without significantly disrupting the hydrogen-bonding network of the β-sheet. The protein serves as a framework to bind FMN strongly, with the isoalloxazine rings located near the carboxyl-terminal portion of the sheet. Several protein loops envelope FMN especially enclosing the polar pteridine rings. The 5′-phosphorylated ribityl chain extends into the interior of the protein forming an extensive system of hydrogen bonds with the N-terminus of helix 1 and anchoring the cofactor to the protein. The FMN isoalloxazine ring system is completely buried with only the methyl groups of the benzene ring exposed to bulk solvent. The association constants vary as a function of the oxidation state of FMN with the sq form binding approximately 100 times more strongly than either the ox or hq forms. However, there is significant variability in these binding constants among the different species studied (Mayhew & Ludwig, 1975). The sequence of this FMN-binding pocket appears to be largely conserved (Ludwig & Luschinsky, 1993) despite the fact that most of the hydrogen bonds are with the main-chain atoms.

Our previous structural studies of this protein, using X-ray crystallography (Rao *et al.*, 1992) and high-resolution multidimensional multinuclear NMR spectroscopy (Stockman, Krezel, Markley, Leonhardt & Straus, 1990), utilized flavodoxin purified directly from large-scale cultures (80 l batches) of *Anabaena* 7120 (Stockman, Westler, Mooberry & Markley, 1988). Since genetic manipulation of *Anabaena* 7120 is difficult, we decided to express the cloned gene in *E. coli*, in order to pursue site-specific structural mutations in the flavodoxin and to facilitate the production of isotropically labeled protein for NMR studies.

Three long-chain oxidized flavodoxin structures are known (Smith *et al.*, 1983; Fukuyama, Matsubara & Rogers, 1992; Rao *et al.*, 1992), the latter from the photosynthetic bacterium *Anabaena* 7120 was reported by our laboratory at 2.0 Å resolution in the monoclinic form. During our efforts to obtain crystals of the present recombinant protein in the same crystal form (monoclinic) for further study, we grew the new trigonal crystal form that diffracted to 1.40 Å resolution. These crystals display a new packing motif for long-chain flavodoxins.

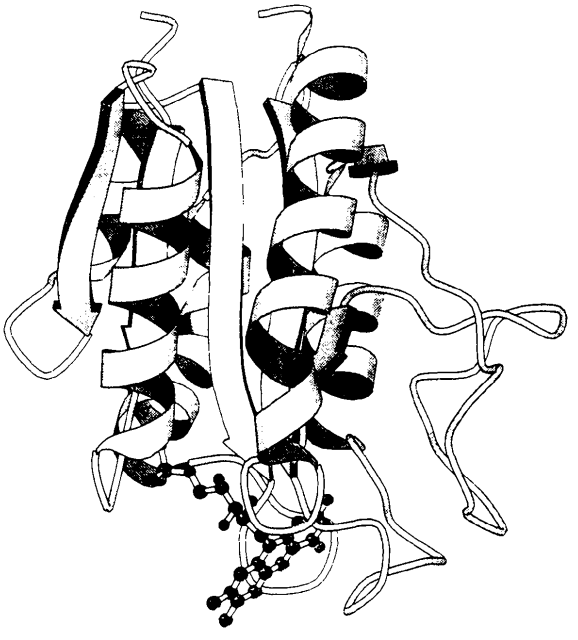

Fig. 1. Schematic ribbon drawing of *Anabaena* 7120 flavodoxin drawn with the graphics display package *MOLSCRIPT* (Kraulis, 1991).

We report here the structure of this protein as well as structural and packing-motif comparisons with the monoclinic polymorph as well as with other flavodoxins, both long- and short-chain forms.

Experimental procedures

Materials

Sequenase was obtained from United States Biochemicals. The plasmid pGEM-3Zf(−) was purchased from Promega. The plasmid pGPl-2 was a gift of Dr Richard Burgess (University of Wisconsin, Madison, USA). The *E. coli* strain AW737 (phototrophic strain) was from the collection of Dr Julius Adler (University of Wisconsin, Madison, USA). Sephadex G-50 resin and DEAE–cellulose DE-53 were from Pharmacia and Whatman, respectively. The ammonium sulfate and Tris buffer used in the crystallization was purchased from Aldrich.

Cloning

Restriction-enzyme digestion, ligation and bacterial transformation were carried out as described by Maniatis, Fritsch & Sambrook (1982). DNA sequencing was performed using the chain-termination method of Sanger, Niklen & Coulson (1977) following the protocol provided by United States Biochemicals with Sequenase.

Expression and purification

10 ml of LB medium containing 100 mg ml^{-1} each of the ampicillin and kanamycin was inoculated with a single colony or 2–3 µl of frozen stock of *E. coli* AW737 containing pGEM-AVF and pGPl-2 and grown overnight with good aeration (200 rev min^{-1} shaker) at 303 K. A 0.5 ml aliquot of this overnight culture was inoculated into 350 ml of a rich (LB) or minimal (M9) medium in a 2.8 l Fernbach flask and incubated in a temperature-controlled shaker bath at 303 K with supplemental aeration of (0.45 µm) filtered air. When an optical density (OD) at 600 nm of 1.0–1.3 was reached, the temperature of the bath was increased to 315 K for 0.5 h then lowered to 313 K, the incubation was then continued for 6–12 h. Cells were harvested by centrifugation for 5 min at 2500*g*. The pellet was then resuspended in 40 ml of low-salt buffer (LSB: 50 m*M* phosphate pH 7.6, 100 m*M* KCl, 5 m*M* EDTA) and divided into two 20 ml aliquots. Each aliquot was sonicated (power setting 30, Sonifer Cell Disruptor Model W185 with microtip), and the resulting lysate was then passed through a French press (Amicon French Pressure Cell) at a force of 700 PSIG and a flow rate of approximately 10 ml min^{-1}. Unbroken cells were pelleted by centrifugation for 10 min at 7000*g*. The supernatant was decanted and stored at 273 K. This procedure (resuspend, sonicate, French press, centrifuge, decant) was repeated twice more with each pellet, and all the supernatants were pooled. Repeated disruption was necessary in order to minimize loss of expressed protein in the pelleted fractions 5–10 mg of FMN was added to the pooled supernatants. This material was then loaded onto a 1.5 × 15 cm column of Whatman DE-53 anion exchange resin at 277 K which had previously been equilibrated with LSB. After washing with one column volume of LSB, an intense orange band remained just below the upper surface of the resin bed. The column was developed with a 700 ml gradient of 100–300 m*M* KCl in the same phosphate buffer, and 10 ml fractions were collected. The band eluted in a small number of fractions. All of the orange-colored fractions were pooled and concentrated to a small volume on an Amicon stirred-cell pressure-filtration concentrator with a 10 000 M_r cutoff filter (YM10). This concentrated solution was purified further and desalted on a 1.5 × 40 cm column of Sephadex G-50 resin. The orange eluant was dialyzed against 3 l of 100 µ*M* phosphate buffer (pH 7.6). The dialysate was lyophilized and yielded 20–30 mg of purified flavodoxin.

Crystallization

Crystals were grown by the hanging-drop vapor-diffusion method from 10 µl droplets at 277 K, with each droplet containing equal parts of protein (80 mg ml^{-1} in 200 m*M* Tris–HCl at pH 8.0) and 80% saturated ammonium sulfate (Table 1), conditions similar to those used to obtain the monoclinic crystals (Rao *et al.*, 1992), but using 3.3 *M* ammonium sulfate instead of 3.0 *M*. The protein salted out forming a fluffy precipitate from which large bright orange crystals averaging 0.3 × 0.3 × 0.4 mm in size appeared in approximately 1–2 weeks. The crystals were well formed and hexagonal in shape.

Data collection and processing

Data were collected at 263 K from a single crystal on a Siemens four-circle area detector utilizing Cu $K\alpha$ radiation ($\lambda = 1.54$ Å) from a Maxscience rotating-anode source (50 kV, 90 mA). The detector was mounted 125 mm from the crystal along the 2θ arm. A total of 17 data sets were collected by varying 2θ, χ, φ and ω. Data were collected in 0.2° frame^{-1} steps with a 90 s exposure time per frame. The crystal remained very stable during the entire data collection (nearly one week) with a fall-off intensity occurring only at the end of the data-collection period. The 17 data runs were processed and merged together using the program *XENGEN* 2.0 (Howard, Nielsen & Xuong, 1985) yielding an $R_{\rm merge}$ of 3.8%. The space group was determined to be $P3_121$ with cell constants $a = 55.36$, $c = 102.59$ Å, and one molecule in the asymmetric unit. A total of 311 379 reflections were collected of which 36 069 were unique with 32 963 reflections having intensity greater than $2\sigma(I)$ (Table 1).

Table 1. *Crystal and data statistics comparison for trigonal and monoclinic flavodoxins and completeness of trigonal data*

	Trigonal crystal (present work)	Monoclinic crystal (Rao *et al.*, 1992)
Space group	$P3_121$	$P2_1$
Cell constants	$a = b = 55.36$ Å	$a = 48.0$ Å
	$c = 102.59$ Å	$b = 32.0$ Å
	$\alpha = \beta = 90°$	$c = 51.6$ Å
	$\gamma = 120°$	$\alpha = \gamma = 90°$
	$Z = 6$	$\beta = 92°$
		$Z = 2$
Cell volume (Å^3)	272290	79210
V_m (Å^3 Da^{-1})	2.26	2.06
Total reflections collected	311379	24329
Unique reflections	36069	9795
Merging R factor (%)	3.8	7.0

Completeness of data in trigonal flavodoxin

Resolution (Å)	Number of indepedent reflections* $I \geq 0\sigma(I)$	$I \geq 2\sigma(I)$
∞–8.00	240 (98.8)	238 (97.3)
8.00–2.38	7366 (98.0)	7313 (97.8)
2.38–1.90	7257 (99.8)	7151 (97.8)
1.90–1.66	7210 (99.9)	6941 (97.3)
1.66–1.51	7172 (99.8)	6501 (95.6)
1.51–1.40	6774 (94.3)	4817 (90.0)
∞–1.40†	36019 (99.1)	32961 (90.0)

* Numbers in parentheses are the percentages of possible reflections.
† The total number of reflections is higher than the breakdown indicates because there are 50 reflections with resolution higher than 1.40 Å. The reflections were used in the refinement.

Structure solution

The structure was solved by molecular replacement in real space using the program *X-PLOR* 3.0 (Brünger, 1988). 1181 reflections ($F_o \geq 2.0$) from 15 to 4.5 Å and the atomic coordinates (residues 2–169) from the 2.0 Å monoclinic structure were used to calculate the Patterson maps. The rotation function gave a single solution at $\theta_1 = 349.5$, $\theta_2 = 78.5$ and $\theta_3 = 340.5°$ that was used in a translation search over the asymmetric unit using 1 Å grid. The model derived from the molecular replacement had a starting R factor of 38.9% which dropped to 35.2% after rigid-body refinement. The correctness of the molecular-replacement solution was confirmed in the difference map by the presence of clean electron density for FMN which had been omitted from the molecular replacement.

Refinement

Following rigid-body refinement, omit maps clearly showed density for the N-terminal residue, S1, which was built into the model. The molecule was then annealed by initially setting the temperature at 3000 K and slow-cooling in 25 K steps to 300 K using the 15 140 reflections in the range 8.0–1.85 Å. At each step in the cooling process, the model was equilibrated with 5 fs of molecular-dynamics simulations. Following the final round of annealing, the molecule was subjected to 120 cycles of positional refinement, lowering the R factor to 23.4%. The thermal parameters were maintained at 15.0 Å^2 during the course of the annealing.

Difference 'omit' maps ($F_o - F_c$), in which ten-residue segments were excluded from phasing, showed that several regions of the molecule required manual manipulation to fit the electron density correctly. The model was rebuilt residue-by-residue using these maps on our ESV-10 using *FRODO* (Jones, 1985) and refined after every 30 residues. Two regions, residues 27–30 and 134–136, required extensive refitting to fit the electron density. The electron density was well defined over all regions in omit maps at this resolution. After fitting was completed, several rounds of solvent picking and refinement were performed until 120 waters and the sulfate ion were found in difference maps.

The data were then extended to 1.4 Å and the structure was refitted and refined with *X-PLOR* 3.0 in iterative sessions as previously described. Restraints on B factors were not used. A total of 16 side chains were modeled in at least two distinct confomations with equal occupancies for each conformer, E16, S17, V18, R23, D24, E25, L33, N58, E61, S64,* E67, S71, S110, S158, S165 and E166, with one existing in three conformations. The R factor for this model was 16.3% for 35 779 data with unit weights between 8.0 and 1.4 Å.

The final refinements were carried out using the *SHELXL*93 program (Sheldrick, 1993), which refines against F^2 instead of F. The program also uses a diffuse solvent continuum model according to Babinet's principle (Langridge *et al.*, 1960). Because the data quality and the resolution were high, inclusion of H atoms as 'riding atoms' on the heavier protein atoms and anisotropic refinement were tested as further extensions to the refinement. An R_{free} (Brünger, 1992) test conducted within *SHELXL*93 was performed to explore these options. These separate tests showed that inclusion of scatterers for H atoms as 'riding atoms' on protein atoms was reasonable whereas anisotropic refinement was not. The R_{free} value was lowered by 1.4% by including H-atom scatterers whereas the R_{free} was lowered by only 0.2% by including anisotopic thermal parameters. Therefore, 'riding model' H atoms on protein atoms were used in the final stages of the refinement, while anisotropic refinement was not performed. Any H atom attached to a group with a rotational degree of freedom (*i.e.* hydroxyl and methyl H atoms) were initially placed according to maxima in the electron density and the torsional angle refined. H atoms were not placed on water molecules. In addition, group occupancies for the multiply ordered side chains were refined with the occupancies constrained to sum to unity.

The final model consists of 3072 scatterers in 169 residues (16 with alternate side-chain positions), FMN, 299 water molecules (with refined occupancy and

* This residue was modelled with the hydroxyl group staggered in three positions about χ_1 by approximately 120°.

thermal parameters), and a sulfate ion for a total of 7380 refined parameters. The R factor was 13.9% for the 32963 reflections with $I > 2\sigma(I)$ and 14.7% or all 36069 data. The reflection weighting scheme used was that of *SHELXL*93 where $w = 1/[\sigma^2(F_o^2) + (0.2984P)^2 + 462.0744P]$ and $P = (F_o^2 + 2F_c^2)/3$. The average error in the coordinates as determined by the method of Luzzati (Luzzati, 1952) is 0.11 Å and the r.m.s. deviations from ideality (Engh & Huber, 1991) are 0.007 Å and 0.9° for bonds and angles, respectively. Some representative electron-density plots are show in Fig. 2 to demonstrate the excellent quality of the maps. The refined coordinates have been deposited with the Protein Data Bank (Bernstein *et al.*, 1977).*

Results and discussion

Flavodoxin overproduction

Plasmid pGEM-AVF was constructed by ligating the *Hin*dIII fragment from a plasmid that contained the gene for flavodoxin (Leonhardt & Straus, 1989), into the *Hin*dIII site of plasmid pGEM-3Zf(−) downstream of the T7 promoter. On screening the resulting plasmids for direction of insert, it was found that all of the 16 clones that contained the flavodoxin gene had it inserted in the same direction: all inverted relative to the *lac* promoter that transcribes across this site. As a result, the flavodoxin gene is not expressed by the *lac* promoter. Since statistical (50/50) directionality of insertion is expected, this unusual result suggests that the flavodoxin is toxic to *E. coli*, even in the small amounts produced by leakage from this promoter. Osborne, Chen & Mathews (1991) reached similar conclusions about the toxity of the *E. coli* flavodoxin gene to *E. coli* when present on a multicopy plasmid. An inverted orientation with respect to the *lac* promoter positions the flavodoxine gene appropriately for expression from the T7 promoter, which is the promoter used for protein expression, so the selection ensures the correct direction for our purposes. The overexpression plasmid (pGEM-AVF) was transformed successfully into cells containing pGPl-2. The second plasmid (pGPl-2) contained both the gene for T7 RNA polymerase, under control of the λ P_L promoter, and the gene for a temperature-sensitive λ repressor protein (c1857) (Tabor & Richardson, 1985). This construct allows induction of T7 RNA polymerase by shifting the growth temperature, which in turn causes the flavodoxin gene on pGEM-AVF to be transcribed.†

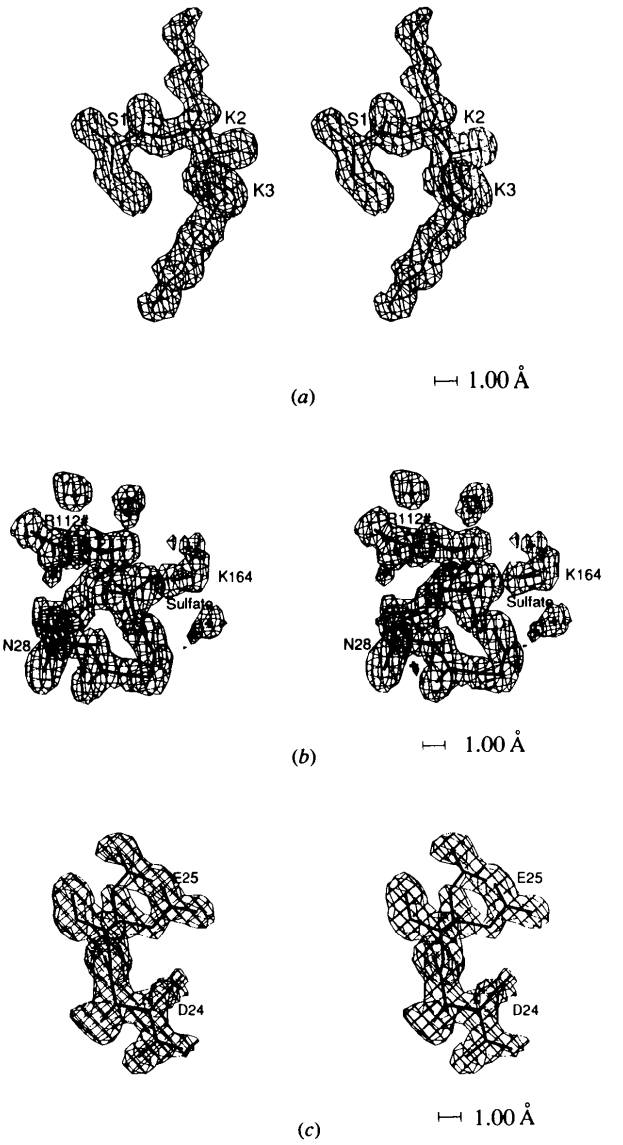

Fig. 2. Stereo drawings of three regions of the molecule illustrating the excellent quality of the electron density. All electron densities are from difference omit maps of their respective regions. (*a*) The tripeptide S1-K2-K3 at the N-terminus contoured at 2.0σ. This area was considerably more mobile in the monoclinic structure and had poor electron density. In the trigonal crystal, this area is a region of intermolecular contact, *via* water molecules, and consequently is much better ordered. (*b*) The tetrapeptide segment G27-N28-D29-V30 comprising the turn between α1 and β2 with bound sulfate ion contoured at 2.0σ. This region was also well ordered in the trigonal lattice as a result of intermolecular contacts. However, in the monoclinic crystal this region was highly mobile and contained ill defined electron density The sulfate dianion is countered by two full positive charges from K164 and R112 of a neighboring molecule, while N28 donates two partial positive charges from its backbone and side-chain amides. (*c*) Stereo drawing of the dipeptide D24-E25 from α1 with electron density contoured at 1.5σ. This helix is exposed to one of the large solvent channels running along *c*. Consequently many of the residues along this helix have multiply ordered side chains as demonstrated here.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1RCF, R1RCFSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0369).

† We had no success in transforming purified pGPI-2 DNA into cells that already contained pGEM-AVF or in transforming both plasmids simultaneously. The few colonies recovered all failed to produce flavodoxin upon temperature induction.

When the apoprotein produced was reconstituted with exogenous FMN, the product was found to be indistinguishable from flavodoxin isolated directly from *Anabaena* 7120 by NMR spectroscopy (J. B. Olson, R. A. Chylla, H. Yan, F., Abildgaard, E. S. Mooberry & J. L. Markley, unpublished results) and reduction potential measurements (K. E. Paulsen & M. T. Stankovic, personal communication).

β-Sheet: structure and hydration

The main β-sheet of flavodoxin comprises the core of the molecules (Fig. 3*a*). It is composed of five parallel β-strands β1(4–9), β2(30–36), β3(48–56), β4(81–89) and β5(115–117/141–143) in the relative orientation 2–1–3–4–5 as defined by the main-chain hydrogen-bonding pattern. There are 26 hydrogen bonds holding the strands together with a mean distance of 2.96 (0.14) Å [and mean C=O···N angle of 160 (7)°]. Nine water molecules (Wat1, Wat3, Wat4, Wat6, Wat10, Wat56, Wat85, Wat170 and Wat233) play significant roles in stabilizing the termini of the β-strands.

The juxtaposition of five β-strands creates a total of eight crevices, four gaps at each end of the β-sheet. Each of these gaps contains one or more bridging water molecules that form hydrogen bonds with one or more backbone atoms. In one case, the interactions utilize side-chain atoms to complete the bridging. The net effect of these waters is to extend the β-sheet in each of the gap regions, which is reflected in the φ/ψ backbone torsion angles. The crevice created by the C-terminal portions of the β1 and β3 strands contains no such bridging water; however this region is stabilized by the extensive hydrogen-bonding network of the FMN phosphate.

The N-terminal β1/β3 gap contains Wat10 (Fig. 3*b*), bridging the carbonyl group of K3 and the amide of Y49 while simultaneously bridging the carbonyl of S1 and the amide N atom of Q48. The water molecule is bound to the four ligands in a nearby tetrahedral geometry. These interactions probably play a role in ordering S1 in the lattice. The β1/β2 gap is bridged by Wat85, connecting the carbonyl O atoms of K2 and V30, while β3/β4 gap is filled by Wat233 which interacts with the carbonyl groups of Q48 and G80. Another solvent molecule, Wat56, caps the β4/β5*a* gap bridging the carbonyl group of G113 with the amide of G80 and the carbonyl group of K81.

Wat6 binds in the C-terminal gap between strands β4 and β5*b*. It extends this region of the β-sheet by bridging the amide N atom of E145 and the carbonyl O atom of G89 as described (Sekharudu & Sundaralingam, 1993). This water also forms a hydrogen bond with the O4′ of the FMN ribityl chain. Although this water molecule is present in both polymorphs of *Anabaena* 7120 structures, it is not present in the structures of either of the other long-chain flavodoxins. The C-terminal crevice between strands β1 and β2 and in part between β1 and β3 is stabilized by Wat4 which hydrogen bonds to the amide of T10 and the side-chain carboxylates of D35 and D65. Strand β3 contains a proline at residue 55 which disrupts the hydrogen-bonding pattern of the β-sheet. This proline ring is puckered such that the C^{β} atom is displaced from the plane of the other four ring atoms by approximately 0.5 Å resulting in the C^{β}-*exo* ($_{\beta}E$) type of puckering, which is 'enantiomerically' related to the C(2)-*endo* (^{2}E) puckering seen in the furanose sugar rings of B-DNA (Westhof & Sundaralingam, 1983). In several other flavodoxins, this proline is replaced by a serine instead. The presence of a proline at this particular sequence position in the long-chain flavodoxins seems to influence the positioning of the 58–60 loop for interaction with N5 of FMN and for positioning the side chain of W57 above the FMN ring. This placement allows $N^{\varepsilon 1}$ of W57 to form a hydrogen bond with the phosphate of FMN thus tilting at an angle of about 40° to the FMN ring system instead of being parallel, as is Y94. Structurally, the β-sheet morphology is maintained by two bound waters that stabilize this region (Fig. 3*c*). Two solvents, Wat3 and Wat170, lie in the gap created when the side chain and peptide plane of P55 rotate ~ 90° away from their 'normal' positions in a parallel β-sheet. These two water molecules form numerous interactions, including hydrogen bonds with the carbonyl group of C54 and T88, creating a two-water bridge ($C54^{C=O}\cdots$Wat3$\cdots$Wat170$\cdots{}^{O=C}$T88). In addition, Wat170 makes interactions with O2 of FMN and the amide N atom of T56. In all, Wat170 makes seven close contacts with nearby polar atoms in an arrangement that is pseudo-octahedral, such as that seen in metal binding, though the hydrogen-bonding distances (2.69–3.31 Å) as expected are longer than the 2.2–2.4 Å seen in metal-binding sites. These hydrogen bonds probably take the place of the two hydrogen bonds lost by the presence of P55, including a conserved serine hydroxyl hydrogen bond with the phosphate. There are two other longer interactions on either side of these extended water bridge that may be stabilized by the presence of these waters. The first interaction between the carbonyl group of C54 and the amide N atom of T88 has a distance of 3.94 Å and a C=O···N angle of only 110°. The second, between T56 N and T88 O, has a distance of 3.50 Å and C=O···N angle of 166° which is clearly more favorable. A further stabilization of this region arises as a result of the FMN ribityl chain which makes interactions with the O^{γ} atom of T56 and the carbonyl O atom of T88. Although the poor geometry of these interactions suggests that they are not strong hydrogen bonds, they probably contribute to the overall stability of the region through electrostatic interactions.

Strand β5 is split in two by the insertion of the long loop (118–140). However, the hydrogen-bonding pattern of the β-sheet is maintained by Wat1. It forms hydrogen bonds with the amide N atom of Y85 and the carbonyl O atoms of V117 and V139, effectively tying the three spatially close but sequentially distant strands together.

This water is present in all of the other long-chain flavodoxin structures.

Helices: structure and hydration

Five helices (13–26, 40–46, 63–76, 100–113 and 149–167) surround the β-sheet in the flavodoxin molecule with helices 1 and 5 on one side of the β-sheet and helices 2–4 on the other. The 54 hydrogen bonds comprising these helices have a mean O···N_{i+4} distances of 3.09 (0.16) Å and a mean C=O···N angle of 150 (7)°. Three of the five helices are composed entirely of α-helix, with the other two being comprised wholly or in part of 3_{10} helix.

Helix 1 is entirely α-helical in character. The residues in this helix make one intermolecular and five intramolecular hydrogen bonds with other protein atoms. Four of these five intramolecular hydrogen bonds are with residues involved in phosphate interactions, forming a highly intertwined pocket stabilizing FMN binding. Six of the 13 carbonyl groups in this helix are hydrated by one or more water molecules. A large proportion of the

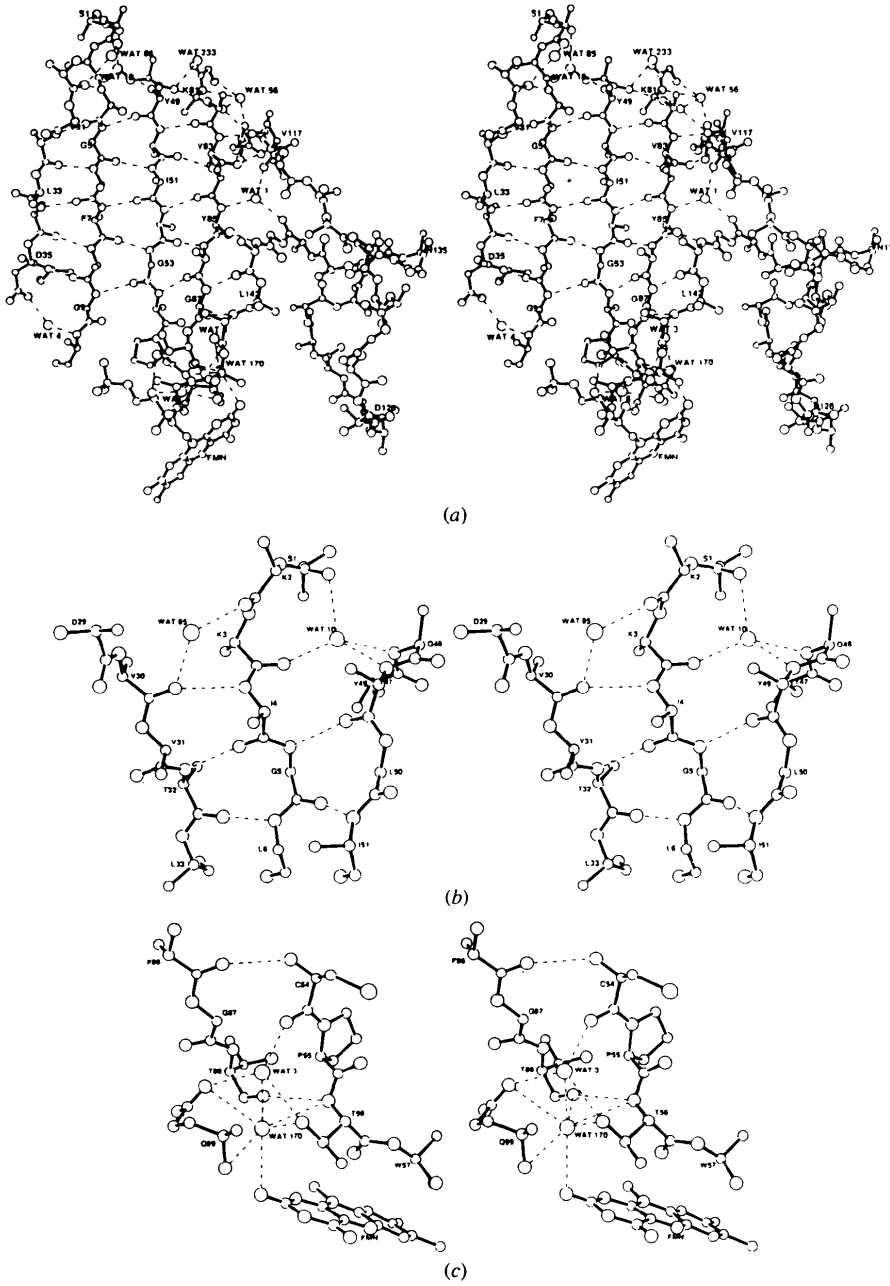

Fig. 3. Portions of the β-sheet, illustrating bound water molecules. (*a*) Stereo drawing of the central β-sheet in flavodoxins with the nine structural waters. The view is similar to that in Fig. 1. Shown for illustration are the FMN cofactor which binds at the bottom of the sheet, and the long loop that interrupts β5. (*b*) Stereo drawing of the N-terminal region of the β-sheet in the vicinity of the N terminus. The N terminus, S1, is found in clear electron density. The ordering of this residue, as well as better ordering of K2 and K3 may occur as a result of the bound water Wat10. This water forms four hydrogen bonds with neighbouring polar atoms, two H-atom donors (amides of Q48 and Y49) and two H-atom acceptors (carbonyls of S1 and K3). The geometry of binding is nearly tetrahedral. (*c*) Stereo drawing of structural waters in the vicinity of P55. The side chains of residues F86, W57, and the ribityl chain of FMN have been omitted for clarity. The integrity of the β-sheet, deformed by the presence of a proline at this point, is maintained by the numerous interactions created by the insertion of waters 3 and 170 in the 'hole' formed by the rotation of the proline peptide plane away from the plane of the β-sheet. Wat170 forms a total of seven interactions with neighboring polar atoms.

residues with alternate side-chain conformations are members of this helix and include residues E16, S17, V18, R23, D24 and E25.

The second helix, residues 40–46, the shortest of the helices in this structure, is entirely 3_{10} in character. The tightly wound nature of this helix causes the carbonyl groups to form bent hydrogen bonds with the corresponding N_{i+3}. This bending allows three of the seven carbonyl groups to be hydrated by one or more water molecules. The 3_{10} folding does not appear to be an artifact of the packing because this region does not make any intermolecular hydrogen bonds in the crystal lattice. There are two hydrogen bonds between backbone atoms and side-chain atoms of this helix. One of these hydrogen bonds, between the carbonyl group of N45 and N^{ζ} of K81, anchors this helix to the third strand of the β-sheet.

The third helix shows considerable change from the monoclinic structure. The helix is best defined by including residues 63–76 as defined by the backbone torsion angles. The helix begins as a normal α-helix, but tightens at the carboxyl terminus, becoming more like a 3_{10} helix after the helix kinks at L69. This is in contrast to the monoclinic form in which the N-terminal region was as it is here, but degenerated into random coil at the transition where the helix became 3_{10} in this structure, most likely as a result of packing interactions in this vicinity. There are a considerable number of hydrogen bonds involving side-chain atoms: those that help stabilize this helix by interacting with itself and those that bind the helix to the sheet structure. The side chain of W66 ($N^{\varepsilon 1}$) helps bind the helix to the β-sheet while the side-chain atoms from residues Q63, S71 and E72 form hydrogen bonds with backbone atoms of the helix, possibly stablizing a bend of approximately 35° at residue L69. The biphasic nature of the helix is demonstrated by an examination of the hydration pattern. The portion of the helix that is primarily α-helix has only two of the seven carbonyl groups hydrated, *i.e.* hydrogen bonded to a water molecule. But, after the transition occurs, five of the seven carbonyl groups are hydrated as seen in the other 3_{10} helix. This region lies near a crystallographic twofold and helix lies nearly parallel to it. Three residues make significant contacts with the twofold-related molecule and all three of these residues (S64, E67 and S71) also adopt alternate side-chain conformations. The alternate form of E67 is essential for maintaining the twofold (Fig. 4). The twofold-related $O^{\varepsilon 1}$ of the major forms (1.97 Å), as well as the $O^{\varepsilon 1}$, of the minor forms (1.16 Å) are too close to exist simultaneously. Although the crystallographic twofold is maintained, neither form may exist unless the corresponding residue on the symmetry-related molecule is in the other conformer, *i.e.* disordered, in order to maintain reasonable van der Waals interactions in this region. Since there were a large number of intermolecular forces present in the monoclinic crystal, the secondary structure shown in the trigonal crystal may more closely resemble that present in solution.

The fourth helix is entirely α-helical being comprised of residues 100–113. There are numerous interactions involving this helix including three between the side-chain atoms of R112 and the backbone carbonyl groups of helix 3. Two hydrogen bonds anchor this helix to the β-sheet while two other hydrogen bonds hold the long loop against this helix. The hydration of the helix is fairly uniform with every third carbonyl being hydrated, which

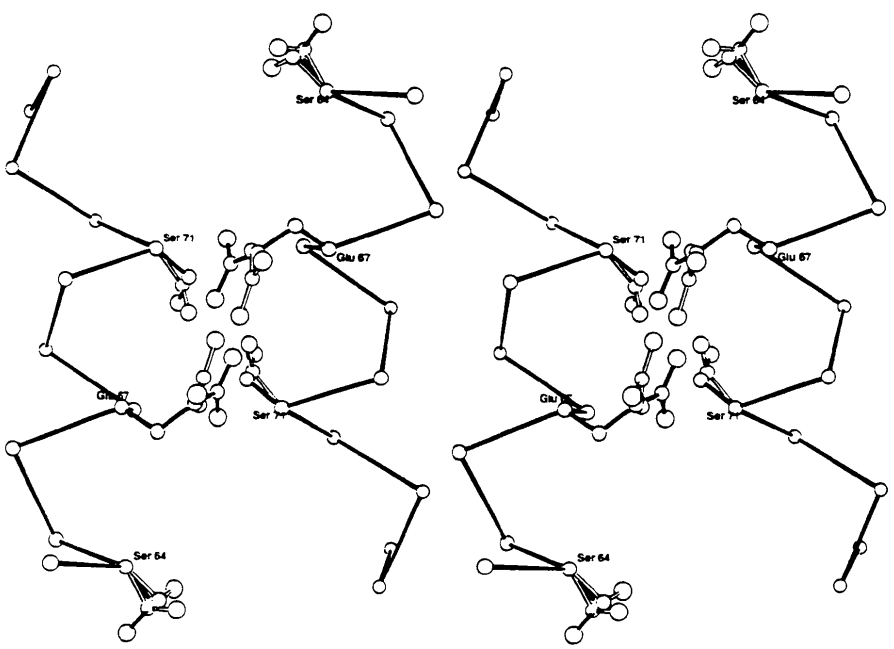

Fig. 4. Stereo drawing illustrating the disordered residues S64, E67 and S71. Both positions of the disordered side chains are shown with the lower occupancy conformer drawn in hollow bonds.

are also those that are on the solvent-exposed side of the helix.

The fifth helix of flavodoxin begins are residue 148 and continues to the C-terminus, although residue L169 appears to have only minimal association with the helix. In addition, the first turn of the helix may be partially 3_{10} in character as judged by the hydrogen bonding. There are many side-chain hydrogen bonds formed mostly with atoms of the long loop, although there are several interactions among the members of the helix. Three alternate side chains occur in this region, S158, S165 and E166. There are also several symmetry-related hydrogen bonds especially with the C-terminus. The helix is heavily hydrated with 12 of the 21 carbonyl groups, excluding the C-terminus, bearing a water molecule. The pattern is fairly regular with those carbonyl groups exposed to solvent being hydrated, while those on the interior are not. The C-terminus is also heavily hydrated in addition to making five hydrogen bonds with symmetry-related molecules.

The long loop; structure and hydration

The long loop is the defining feature of the long-chain flavodoxins, but is does not play an obvious role in modulating the redox potentials (Ludwig & Luschinsky, 1993). In the crystal, the bulk of the long-loop is exposed to one of the large solvent channels. Consequently, many of the residues have high mobility including residues D126 and N135 whose side-chain electron density is weak and ill defined. Yet, much of the loop region is stabilized by several intra-loop hydrogen bonds, two of which involve W120 and F138 and form a short β-sheet-like region. However, this sheet-like domain does not interact with the main β-sheet of the molecule.

Turns: structure and hydration

A total of 14 turns have been identified in this crystal structure (Table 2). Eight turns connect the regular secondary-structure elements, while the other six exist in regions of random coil. Three contiguous turns have been identified in the 91–100 region and three non-adjacent turns have been identified in the long loop (residues 118–138). The majority, nine, of the turns are type I (I′) (Wilmot & Thornton, 1988). The remainder are of type II (II′) except for the $3_{10}/\beta 3$ turn which is in the less common type VIII conformation and may be a degenerate continuation of the 3_{10} helix. All of the turn regions are exposed to solvent with every polar atom not already involved in hydrogen bonding participating in hydrogen bonds with water molecules.

The 26–29 turn is of considerable interest in this crystal structure because of the presence of a bound sulfate ion in this region (Fig. 2*b*). This turn was disordered in the monoclinic structure as well as in the *A. nidulans* structure. However, in this case, this region is very well ordered. The ordering arises as a result of the packing in this crystal which places this loop in close contact with the $\alpha 3/\beta 5$ loop of a symmetry-related flavodoxin. This interaction traps a sulfate ion, further stabilizing the residues in these regions (Table 3).

Table 2. *Turn regions in trigonal flavodoxin*

Residues	$(\varphi, \psi)_2$ (°)	$(\varphi, \psi)_3$ (°)	Type
T10-Q11-T12-G13	−110/−50	−107/15	I
F26-G27-N28-D29	64/−123	−88/1	II′
V36-S37-Q38-A39	−68/−20	−110/−1	I
N45-D46-Y47-Q48	−72/−10	−129/155	VIII
N58-I59-G60-E61	63/74	86/12	I′
F78-N79-G80-K81	−62/136	90/2	II
Q91-I92-G93-Y94	−86/−47	−77/−24	I
Y94-A95-D96-N97	−63/−23	−93/−3	I
N97-F98-Q99-D100	−70/116	68/23	II
S110-Q111-R112-G113	−56/−37	−92/13	I
T122-D123-G124-Y125	−67/137	94/7	II
S130-K131-A132-L133	−76/−4	−88/−3	I
R134-N135-G136-K137	54/46	88/−1	I′
E145-D146-N147-Q148	−81/−36	−118/−26	I

Table 3. *Sulfate interactions in trigonal flavodoxin*

Sulfate atom	Protein atom	Distance (Å)	Symmetry operation
OS1	N28 N	3.31	—
	R112 $N^{\eta 2}$	2.67	5 1 1 1
	K164 N^{ζ}	3.21	—
OS2	Water 67	3.27	5 1 1 1
	Water 113	2.73	—
OS3	N28 $N^{\delta 2}$	3.08	—
	R112 $N^{\eta 1}$	2.99	5 1 1 1
	Water 225	3.31	5 1 1 1
OS4	N28 N	3.25	—
	Water 160	2.71	—

Symmetry code: 5 1 1 1 = $1-x, y-x+1, \frac{4}{3}-z$.

Flavin mononucleotide: conformation and binding

A molecule of FMN is non-covalently bound to flavodoxin. The FMN-binding site, which has been described in the monoclinic paper in great detail, is largely unchanged in this structure. A total of 22 hydrogen bonds are formed between FMN and the protein or water atoms. The xylene portion of the isoalloxazine ring system is oriented so that the methyl groups are exposed to solvent while the remainder of the rings are buried in the interior of the protein. Three major regions of the protein interact with FMN, T11–T15, T56–G60, and T88–Q99. Additionally, D146 makes a hydrogen bond with one of the ribityl hydroxyl groups.

FMN is anchored to the protein by the 5′-phosphorylated ribityl chain. The anionic phosphate O atoms form a total of ten hydrogen bonds with protein atoms. Six of these hydrogen bonds are with the backbone amide N atoms of residues T11–T15 forming the phosphate-binding pocket. The ribityl hydroxyls form five hydrogen bonds, two of which are with bound water molecules (Wat6 and Wat27).

The isoalloxazine ring system forms seven hydrogen bonds, one of which is with a bound water molecule, Wat170. The two atoms that can undergo protonation upon reduction, N and N5, are involved in hydrogen bonds. The acceptor N5 atom makes a long (3.60 Å) but

favorable linear interaction with the donor amide N atom of I59. Whereas, the amide N atom of D90 forms a hydrogen bond with N1 of FMN. In the monoclinic structure this hydrogen bond is bifurcated between N1 and O2, but in this structure the interaction seems to be stronger to N1 (3.15 Å) than O2 (3.36 Å). The presence of the hydrogen bonds at these two positions may account for some of the differences in electron potentials between the long-chain and short-chain flavodoxins (Rao *et al.*, 1992; Fukuyama *et al.*, 1992). Hydrogen bonds between protein atoms and these two atoms of FMN appear to be present in all of the long-chain structures but not necessarily in the short-chain structures.

Aside from the two methyl C atoms of FMN which are solvent accessible, the immediate environment of the ring system is highly apolar. The rings are sandwiched between the side chain of Y94 on the *si* face and the benzene portion of the side chain of W57 on the *re* side. These rings make angles with the general plane of the isoalloxazine rings of 2 and 40°, respectively. Although W57 is inclined to FMN, the two rings eclipse one another along the long axes of the planar groups. This is not the case with Y94, whose long axis is rotated relative to that of FMN. This orientation of Y94 places O^{η} very near, 3.32 Å, to the C6—C7 bond with the C^{β}—O^{η} aligned approximately parallel to the C7—C7M bond. The isoalloxazine system is not entirely planar despite being restrained to be planar. However, similarly restrained groups such as the side chains of histidine, tryptophans, phenylalanines and tyrosines completely retain their planarity and are flat to within 0.01 Å. The ring system is significantly bent forming a butterfly-like structure, bending along the N5···N10 pseudo-bond at a dihedral angle of approximately 8° (Fig. 5). In addition, the dimethylbenzene ring is somewhat non-planar. Although the six atoms, N10, N5, C9a, C5a, C9 and C6 are coplanar within 0.01 Å, the four atoms most exposed to solvent, C7, C7M, C8 and C8M are displaced from the above plane by as much as 0.10 Å. The displacement of these four atoms gives the benzene ring a puckered appearance, with C7 displayed above the plane and C8 displaced below. The methyl groups of these two atoms behave anomalously as well. The C8 methyl group is displaced below the plane, while that of C7 bends back towards the plane in an sp^3 type of arrangement.

The C7 methyl group of FMN makes a close contact with the carbonyl O atom of G80 in a 3_1 symmetry-related molecule. The carbon–oxygen distance is only 3.29 Å and makes an angle of 118° with the carbon–oxygen double bond of the carbonyl. This interaction may be indicative of a possible C—H···O hydrogen bond between the methyl C atom and the backbone of a symmetry-related molecule. This interaction is especially interesting because G80 is at the N-terminus of strand β4, whose C-terminus forms the hydrogen bond between the amide N atom of D90 and N1 of FMN. The resulting series of these interactions run the length of the *c* axis in the crystal (Fig. 6).

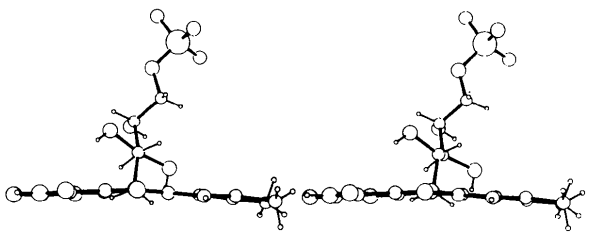

Fig. 5. Detail of the FMN cofactor illustrating the butterfly bend of approximately 8° on the remote side opposite the sugar-phosphate side chain.

Description of the packing

The tightness of the molecular packing in this crystal varies depending on the region of the molecule under examination (Fig. 7). The packing is such that the N-terminal region of the β-sheet of the reference molecule is in direct contact with the C-teminal region of the symmetry-related molecule above. This novel head-to-tail type of packing which is not seen in the monoclinic or any other flavodoxin structure is favorable for satisfying the polararity of the β-sheet as well as the α-helices whose dipoles are arranged in an opposite fashion. 24 amino acids (Table 4) form direct hydrogen bonds with a symmetry-related molecule as the molecules pack along the 3_1 screw. The regions of direct

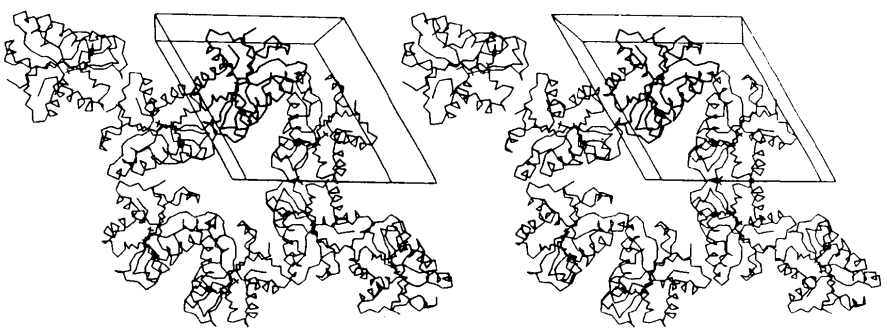

Fig. 6. Lattice interactions in the crystal link symmetry-related FMN molecules *via* the β4 strand of the central β-sheet. The FMN N1–D90 N distance is 3.16 Å and the FMN C7M–G80^{symm} O distance is 3.29 Å suggesting a possible C—H···O interaction.

contact include both the N- and C-termini of the protein and much of the N-terminal region of the β-sheet area around them. Other regions of direct contact include the C-terminal region of the β-sheet, including I59 and E61 which are members of the loop curling around N5 of FMN, as well as much of the area surrounding it including the N-termini of several of the α-helices. The only other significant region of direct contact is along the twofold axis, which runs nearly parallel to the helical axis of $\alpha 2$ creating several interactions among the three side chains, S64, E67 and S71, in addition to ordering them in two or more conformations. In contrast, the monoclinic structure displays fewer contacts but in more areas, including many contacts in the long loop and in the vicinity of $\alpha 3$. It is interesting that despite the completely different manner in which the two crystal systems pack, the molecules display relatively minor differences in their structures.

Several more interactions are formed *via* water bridges and the bound sulfate ion. In the regions of close intermolecular contact, as described above, all of the protein atoms are very well defined in the electron density. In addition, these regions have nearly all of the possible water molecules well defined. However, the trigonal lattice creates large solvent channels, approximately 20 Å in diameter, that run parallel to *c* through the length of the crystal (Fig. 7). The exteriors of $\alpha 1$, 3_{10}, $\alpha 3$, the tip of the long loop and $\alpha 4$, are the regions of the protein that contact this water hole and are considerably more mobile than other areas. The residues of these areas have higher *B* values, multiply ordered side chains, and consequently fewer water molecules were found bound in these regions.

The packing of the molecule in the trigonal lattice represents an entirely new type of packing for the long-chain flavodoxins in terms of the placement of the flavin rings relative to one another. In the structure of *A. nidulans*, *C. crispus* and the monoclinic form of *A.* 7120 flavodoxins, the exposed C7 and C8 methyl groups point toward the exposed portion of the long loop of a symmetry-related molecule with an average distance of 3.5 Å. However, in the present structure these methyl groups point toward the N-terminal portion of the β-sheet with a much shorter contact distance (3.29 Å). The presence of this interaction could be indicative of preferential binding between two flavodoxins or even this flavodoxin and one of its redox partners.

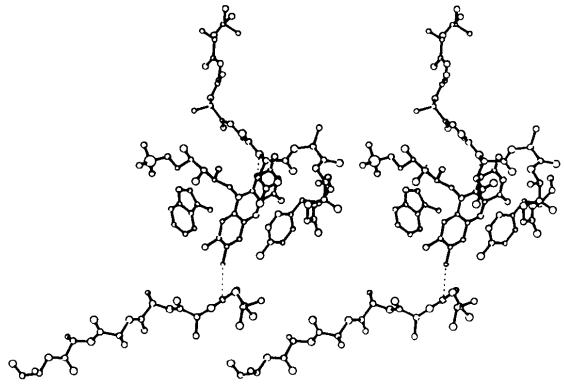

Fig. 7. Packing of *Anabaena* 7120 flavodoxin in the trigonal lattice. The stereoview is down *c*. Note the wide (~20 Å in diameter) solvent channels between the stacked rows of molecules giving several regions free access to bulk solvent.

Table 4. *Interactions between symmetry-related molecules in trigonal flavodoxin*

Protein atom*	Symmetry atom	Distance (Å)	Symmetry operation
S1 N	E107 $O^{\epsilon 2}$	2.80	5 1 1 1
Q11 $N^{\epsilon 2}$	E166 O	2.98	3 1 0 −1
D29 $O^{\delta 1}$	Q111 $N^{\epsilon 2}$	2.82	5 1 1 1
D29 $O^{\delta 1}$	R112 $N^{\eta 2}$	2.92	5 1 1 1
V30 N	Q111 O	2.94	5 1 1 1
D35 N	D150 $O^{\delta 2}$	2.81	4 1 0 1
Q48 $N^{\epsilon 2}$	N58 $O^{\delta 1}$	3.02	2 0 −1 0
Q48 $N^{\epsilon 2}$	N58 $O^{\delta 1'}$	3.25	2 0 −1 0
Y49 O^{η}	N58 $O^{\delta 1}$	3.09	2 0 −1 0
N58 $N^{\delta 2'}$	E107 $O^{\epsilon 1}$	3.09	2 0 −1 0
I59 O	K131 N^{ζ}	2.80	4 1 −1 1
E61 $O^{\epsilon 1'}$	K131 N^{ζ}	2.55	4 1 −1 1
E61 $O^{\epsilon 2}$	K131 N^{ζ}	2.83	4 1 −1 1
S64 $O^{\gamma''}$	G168 O	3.10	3 1 0 −1
E67 $O^{\epsilon 1}$	E67 $O^{\epsilon 1}$	1.97	4 1 −1 1
E67 $O^{\epsilon 1}$	E67 $O^{\epsilon 1'}$	2.36	4 1 −1 1
E67 $O^{\epsilon 1'}$	E67 $O^{\epsilon 1'}$	1.16	4 1 −1 1
E67 $O^{\epsilon 1}$	S71 $O^{\gamma'}$	3.18	4 1 −1 1
S71 O^{γ}	S71 O^{γ}	3.26	4 1 −1 1
S71 O^{γ}	S71 $O^{\gamma'}$	3.08	4 1 −1 1
S71 $O^{\gamma'}$	S71 $O^{\gamma'}$	2.31	4 1 −1 1
D74 $O^{\delta 1}$	L169 O^{ter1}	3.33	5 1 0 1
D74 $O^{\delta 1}$	L169 O^{ter2}	2.43	5 1 0 1
D74 $O^{\delta 2}$	L169 O^{ter1}	3.15	5 1 0 1
G80 O	FMN C7M	3.29	3 1 1 −1
K108 N^{ζ}	L169 O^{ter1}	3.00	5 1 0 1
Q111 $N^{\epsilon 2}$	L169 O^{ter1}	3.23	5 1 0 1

Symmetry codes: $2\ 0\ -1\ 0 = -y, x-y-1, \frac{1}{3}+z$; $3\ 1\ 0\ -1 = y-x+1, -x, z-\frac{1}{3}$; $4\ 1\ 0\ 1 = y+1, x, 1-z$; $4\ 1\ -1\ 1 = y+1, x-1, 1-z$; $5\ 1\ 0\ 1 = 1-x, y-x, \frac{4}{3}-z$; $5\ 1\ 1\ 1 = 1-x, y-z+1, \frac{4}{3}-z$.

* Atoms belonging to the alternate positions are designated with a prime (′) or double prime (″).

Mobility and the alternate side-chain conformations

The overall mobility of the atoms in this structure is relatively low averaging 17.7 $Å^2$ for all of the atoms. However, several atoms, in residues E40, K115, D126, N135 and K157, have very high *B* values which probably indicates a low occupancy for these atoms. There are significant differences in the *B*-value profiles between the monoclinic and trigonal structures. The overall lower *B* factors in the monoclinic form may arise as a result of the data being truncated at 2 Å. It is interesting to note that the *B* factors of the trigonal crystal at an intermediate stage in the refinement, using data truncated at 1.8 Å, displayed similar lower values. This behavior of the *B* values, displaying low sometimes negative values when using data truncated at lower resolution than the diffraction limit of the crystal, seems to be general for

many structures refined in this laboratory (Burkhart & Sundaralingam, unpublished results).

Despite the rigid nature of the molecules in this crystal there are several regions that exhibit high mobility including the side chains with high B values and multiple conformations. The plot of average B values (Fig. 8) shows that the solvent-accessible sites of the helical regions have side chains with higher mobility than the non-accessible ones. In addition, the residues of the central β-sheet at the core of the protein, as expected, have the lowest average B values. Several residues have average side-chain values higher than the average B values for solvents (35.3 Å^2), which is approximately two times higher than the average B value for all atoms in the model. These residues (K14, E20, E72, D75, D77, K115, D123, D129, K137 and D153) have large B values for their terminal atoms and although somewhat disordered still have interpretable electron density. It shouuld be noted that all of these residues belong to either regions of turn or the termini of helices and, with the exception of K115, are all exposed to the large solvent channels running parallel to c.

There are 16 side chains (E16, S17, V18, R23, D24, E25, L33, N58, E61, S64, E67, S71, S110, S158, S165 and E166) exhibiting multiple conformations (Fig. 2c and Table 5). The majority of these side chains are localized in a relatively small area with six of the 16 being serines. This represents over 50% of all serines in the protein with all of them in helical regions and exposed to bulk solvent, which may account for their mobility. For example, the 16–18, 23–25 and 32–33 multiples all lie in approximately the same region as do the side chains of N135 and K137 whose side-chain electron densities are weak and B values are high.

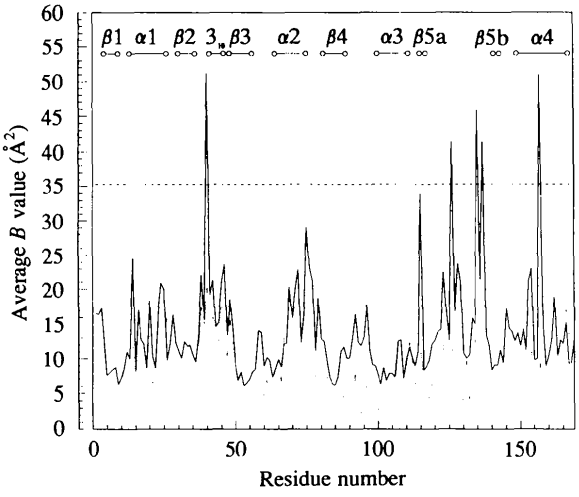

Fig. 8. Temperature factors as a function of residue for the trigonal (solid) and monoclinic (dot) structures. The average temperature factor for each residue is plotted. The secondary-structural elements are shown at the top. The average temperature factor (35.3 Å^2) for the 301 waters (dash) is shown for reference.

Table 5. *Residues exhibiting multiple conformations in trigonal flavodoxin*

Residue	Conformer 1 Occupancy (%)/ average B (Å^2)	Conformer 2 Occupancy (%)/ average B (Å^2)	Conformer 3 Occupancy (%)/ average B (Å^2)
E16	64/24.2	36/17.0	—
S17	58/14.1	42/16.8	—
V18	72/11.3	28/17.2	—
R23	66/17.7	34/21.0	—
D24	46/16.2	54/34.4	—
E25	56/36.2	44/12.2	—
L33	33/6.8	67/19.3	—
N58	58/16.9	42/15.0	—
E61	71/10.0	29/12.5	—
S64	50/7.1	24/11.4	26/8.0
E67	46/9.7	54/16.4	—
S71	62/11.4	38/28.8	—
S110	61/7.2	39/16.2	—
S158	35/10.2	65/25.6	—
S165	36/11.0	64/18.1	—
S166	53/12.7	47/23.8	—

Comparison with the monoclinic structure

Although the origins of the proteins used in the crystallization of recombinant trigonal and natural wild-type monoclinic forms of this flavodoxin are different, chemically they are the same. The r.m.s.d. between the two structures is 0.83 Å for 702 atoms comprising the entire backbone and FMN, except for the four N-terminal atoms of residue S1 which were not modeled in the monoclinic structure. The major areas of deviation include the poorly ordered residues 2–4 at the N-terminus of the monoclinic structure, residues 25–31 in the vicinity of the disordered loop in the monoclinic structure, residues 133–137 at the tip of the long loop, and residues 164–169 at the C-terminus. If these areas are excluded from the superposition, the two structures align with an r.m.s.d. of 0.24 Å. The FMN molecules superpose very well with a mean deviation of 0.24 Å.

Although the data from the monoclinic structure was reported at only 2.0 Å, resolution, restricted by the camera setting for synchrotron data collection, the Matthews coefficient (V_m) (Matthews, 1968) for the monoclinic structure was 2.06 Å^3 Da^{-1} with 39% solvent content. In contrast, the present V_m is 2.26 Å^3 Da^{-1} for the present structure representing a solvent content of 46%. Both of these values are lower than those seen in other flavodoxins. It is expected that comparable high-resolution data may be obtained with good crystals of the monoclinic form. However, the availability of both structures is an even better resource for describing this variety of flavodoxin in as accurate a manner as possible while still demonstrating the conformational flexibility of several regions under molecular contact conditions or free in solution.

Summary

The structures of many flavodoxins (mutants, redox states and those bearing FMN analogues) have been

completed and consequently the general features of the flavodoxin structure are well known (for review see Mayhew & Ludwig, 1975; Mayhew & Tollin, 1993; Ludwig & Luschinsky, 1993; Rao *et al.*, 1992). However, the long-chain flavodoxins present many interesting deviations from the behaviour of the short-chain variety. For example, the effects of the various loop regions on reductive potentials has not been extensively studied. Neither has the importance of the various conserved residues been fully explored *via* mutational studies. This is especially intriguing considering that this protein performs essentially the same function, electron transport, as the ferredoxins, which has nothing in common with it structurally. These conserved regions may not be important in redox-related activities, but may instead be important in the binding of flavodoxin to PS I or FNR. Recent studies with ferrodoxin from *Anabaena* 7120 have implicated key residues important in FNR recognition and possibly in the electron-transfer process (Hurley *et al.*, 1993). If the two proteins can be functionally aligned, regions of the flavodoxin molecule analogous to those identified in ferredoxin may be functionally classified.

This structure of the oxidized form of *A.* 7120 flavodoxin makes an excellent starting model for examination of these reduced states as well as possible site-directed mutant flavodoxins, in part because of its accuracy but also because of its completeness. It is hoped that having such an accurate, well resolved structure of this electron-transport protein will contribute to finding an appropriate electron-transport pathway or mechanism in these flavin-bearing proteins relative to their iron-bearing cousins.

We gratefully thank the National Institutes of Health of the US Public Health Service for partial support of this research through Grant GM17378 and the Ohio Regents Eminent Scholar endowed chair to MS. We also acknowledge The Ohio Supercomputer Center for computer time on the Cray Y-MP/864. Supported in part by USDA Cooperative State Research Service Grant 92-37306-7699 to JLM. HY was supported by an HFSP Postdoctoral Fellowship; RJR was supported in part by a traineeship from the HIH Cellular and Molecular Biology Training Grant (GM07215-16). We would like to thank Drs S. T. Rao and Dr C. Y. Sekharudu for their helpful discussions and Dr George Sheldrick for providing us with the *SHELXL*93 program.

References

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F., BRICE, M. D., ROGERS, J. B., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

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

BRÜNGER, A. T. (1988). *Crystallographic Computing* 4: *Techniques and New Technologies*, edited by N. W. ISAACS & M. R. TAYLOR, pp. 126–140. Oxford: Clarendon Press.

CRESPI, H. L., NORRIS, J. R., BAYS, J. P. & KATZ, J. J. (1973). *Ann. NY Acad Sci.* **222**, 800.

DRAPER, R. D. & INGRAHAM, L. (1960). *Arch. Biochem. Biophys.* **125**, 802–808.

ENGH, R. A. & HUBER, R. (1991). *Acta Cryst.* A**47**, 392–400.

FITZGERALD, M. P., HUSAIN, A. & ROGERS, L. J. (1978). *Biochem. Biophys. Res. Commun.* **81**, 630–635.

FUJII, K. & HUENNEKENS, F. M. (1974). *J. Biol. Chem.* **249**, 6745.

FUKUYAMA, K., MATSUBARA, H. & ROGERS, L. J. (1992). *J. Mol. Biol.* **225**, 775–789.

HOWARD, A. J., NIELSEN, C. & XUONG, N. H. (1985). *Methods Enzymol.* **114**, 452.

HURLEY, J. K., SALAMON, Z., MEYER, T. E., FITCH, J. C., CUSANOVICH, M. A., MARKLEY, J. L., CHENG, H., XIA, B., CHAE, Y. K., MEDINA, M., GOMEZ-MORENO, C. & TOLLLIN, G. (1993). *Biochemistry*, **32**, 9346.

JONES, T. A. (1985). *Methods Enzymol.* **115**, 157–171.

KLUGKIST, J., VOORBERG, J., HAAKER, H. & VEEGER, C. (1986). *Eur. J. Biochem.* **155**, 33.

KRAULIS, P. J. (1992). *J. Appl. Cryst.* **24**, 946–950.

LANGRIDGE, R., MARVIN, D. A., SEEDS, W. E., WILSON, H. R., HOOPER, C. W., WILKINS, M. H. F. & HAMILTON, L. D. (1960). *J. Mol. Biol.* **2**, 38.

LEONHARDT, K. G. & STRAUS, N. A. (1989). *Nucleic Acids Res.* **17**, 4384.

LUDWIG, M. L. & LUSCHINSKY, C. L. (1993). *Chemistry and Biochemistry of Flavoenzymes*, Vol. 3, edited by F. MÜLLER, pp. 427–466. Boca Raton: CRC Press.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802–810.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press.

MATTHEWS, B. M. (1968). *J. Mol. Biol.* **33**, 491.

MAYHEW, S. G. & LUDWIG, M. L. (1975). *Enzymes*, **12**, 57–118.

MAYHEW, S. G. & TOLLIN, G. (1993). *Chemistry and Biochemistry of Flavoenzymes*, Vol. 3, edited by F. MÜLLER, pp. 389–426. Boca Raton: CRC Press.

OGATA, M., MIYAZAKI, F. & AKAGI, T. (1986). *J. Biochem.* **100**, 311.

OSBORNE, C., CHEN, L. & MATHEWS, R. G. (1991). *J. Bacteriol.* **173**, 1729–1737.

PAULSEN, K. E., STANKOVICH, M. T., STOCKMAN, B. J. & MARKLEY, J. L. (1990). *Arch. Biochem. Biophys.* **280**, 68–73.

RAO, S. T., SHAFFIE, F., YU, C., SATYSHUR, K. A., STOCKMAN, B. J., MARKLEY, J. L. & SUNDARALINGAM, M. (1992). *Protein Sci.* **1**, 1413–1427.

SANGER, F., NIKLEN, S. & COULSON, A. R. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 5463.

SEKHARUDU, C. Y. & SUNDARALINGAM, M. (1993). *Water and Biological Macromolecules*, edited by E. WESTHOF, pp. 148–159. London: Macmillan Press.

SHELDRICK, G. M. (1993). *SHELXL*93. *A Program for the Refinement of Crystal Structures*. Univ. of Göttingen, Germany.

SMILLIE, R. M. & ENTSCH, B. (1971). *Methods Enzymol.* **23**, 504.

SMITH, W. W., PATTRIDGE, K. A., LUDWIG, M. L., PETSKO, G. A., TSERNOGLOU, D., TANAKA, M. & YASUNOBU, K. T. (1983). *J. Mol. Biol.* **165**, 737–755.

STOCKMAN, B. J., KREZEL, A. M., MARKLEY, J. L., LEONHARDT, K. G. & STRAUS, N. A. (1990). *Biochemistry*, **29**, 9600–9609.

STOCKMAN, B. J., WESTLER, W. M., MOOBERRY, E. S. & MARKLEY, J. L. (1988). *Biochemistry*, **27**, 136–142.

SYKES, G. A. & ROGERS, L. J. (1984). *Biochem. J.* **217**, 845–850.

TABOR, S. & RICHARDSON, C. C. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 1074.

VETTER, H. & KNAPPE, J. (1971). *Hoppe-Seyer's Z. Physiol. Chem.* **352**, 433–446.

WESTHOF, E. & SUNDARALINGAM, M. (1983) *J. Am. Chem. Soc.* **105**, 970–976.

WILMOT, C. M. & THORNTON, J. M. (1988). *J. Mol. Biol.* **203**, 221–232.

ZUMFT, W. G. & SPILLER, H. (1971). *Biochem. Biophys. Res. Commun.* **45**, 112.