Acta Cryst. (1994). D**50**, 744–748

Preliminary Crystallographic Analysis of Glyceraldehyde 3-Phosphate Dehydrogenase from the Extreme Thermophile *Thermus aquaticus*

BY JACK TANNER, RALPH M. HECHT, MARLESE PISEGNA, DALE M. SETH AND KURT L. KRAUSE

Department of Biochemical and Biophysical Sciences, University of Houston, Houston, Texas 77204-5934, *USA*

(*Received* 14 *October* 1993; *accepted* 11 *February* 1994)

Abstract

Crystals have been obtained of glyceraldehyde 3-phosphate dehydrogenase from the extreme thermophile, *Thermus aquaticus*. This enzyme is stable and active at 363 K, thus its three-dimensional structure should add insight into the structural basis of protein thermostability. Large high-quality crystals were grown using isopropanol and polyethylene glycol at pH 8.4. They crystallize in the orthorhombic space group $P2_12_12_1$ with cell dimensions a = 144.77 (6), b = 148.77 (5), c = 149.50 (7) Å, and diffract to beyond 2.8 Å. The volume of the unit cell and the packing observed in other GAPDH structures suggest that there are two tetramers per asymmetric unit. With 300 kDa/asymmetric unit expected in this form, its solution represents a challenging molecular replacement problem. A low-resolution data set has been recorded and used to carry out self-rotation, cross-rotation and Patterson-correlation refinement calculations. We found that the Q molecular axes of both tetramers are approximately coincident with the crystallographic a axis, and the non-crystallographic symmetry relating the two tetramers is approximately a rotation of 90° about the a axis.

Introduction

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) is a tetrameric enzyme of molecular weight 145 000 that serves an important function in glycolysis and gluconeogenesis. It reversibly catalyzes the oxidation and phosphorylation of glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate. The availability of homogeneous preparations has enabled investigators to elucidate its enzymatic properties as well as its primary through quaternary structures (Harris & Waters, 1976). Several GAPDH crystal structures have been solved, including the enzymes from lobster tail muscle (Moras *et al.*, 1975; Murthy, Garavito, Johnson & Rossmann, 1980; Lin *et al.*, 1993), human skeletal muscle (Mercer, Winn & Watson, 1976), the moderate thermophile *Bacillus stearothermophilus* (Skarzynski, Moody & Wonacott, 1987; Skarzynski & Wonacott, 1988), *Bacillus coagulans* (Griffith, Lee, Murdock & Amelunxen, 1983) and *Trypanosoma brucei* (Vellieux *et al.*, 1993). These structures have been used to address issues such as the mechanism of catalysis, the conformational changes induced by the binding of NAD, the inherent symmetry of the functional enzyme, the use of non-crystallographic symmetry in crystal structure determination, and structure-based drug design. In addition, GAPDH from thermophilic organisms serves as a model for the study of protein thermostability (Perutz, 1978; Argos *et al.*, 1979; Walker, Wonacott & Harris, 1980; Merkler, Farrington & Wedler, 1981; Menendez-Arias & Argos, 1989). We report the crystallization and preliminary X-ray analysis of GAPDH from the extreme thermophile *Thermus aquaticus*.

The *T. aquaticus* GAPDH is stable at 363 K and exhibits a half-life greater than 30 min at 371 K (Hocking & Harris, 1973). Under these conditions other GAPDHs are rapidly denatured. For example, GAPDH from the thermophile *B. stearothermophilus* is totally inactivated at 363 K (Suzuki & Harris, 1971). The basis for the extreme thermostability of *T. aquaticus* GAPDH has been inferred from sequence homology to the *B. stearothermophilus* enzyme (Walker, Wonacott & Harris, 1980). These two enzymes share 70% sequence identity (Hocking & Harris, 1980; Hecht, Garza, Lee, Miller & Pisegna, 1989). It is hypothesized that the increased thermostability of *T. aquaticus* GAPDH is caused by extra surface salt links and increased intersubunit hydrophobic contacts that prevent penetration of water into the protein's interior (Walker, Wonacott & Harris, 1980). The structure of GAPDH from *T. aquaticus* should provide new information on the structural basis of protein thermostability and serve as a test of existing theories.

Experimental

Gene isolation and protein purification

The *T. aquaticus* GAPDH gene was isolated as a 2.2 kb *Bam*HI fragment from a genomic library of strain YT1 (Hecht, Garza, Lee, Miller & Pisegna, 1989). The earlier sequence (Hocking & Harris, 1980) was confirmed and several ambiguities were resolved. The plasmid containing the *T. aquaticus* GAPDH gene failed to rescue the *E. coli* strain DF221 (gap_{am}), a GAPDH amber mutant because of the presence of a putative upstream transcriptional terminator. When most of this region was

ISSN 0907-4449

eliminated, the shortened *T. aquaticus* GAPDH gene enabled the DF221 to grow on glucose. However, the truncated *E. coli* GAPDH was found to copurify with the *T. aquaticus* GAPDH. To bypass this problem, we used an *E. coli* strain, W3CG, made null for GAPDH (Ganter & Pluckthun, 1990).

The W3CG host containing the *T. aquaticus* plasmid was cultured at 310 K overnight as outlined in Ganter & Pluckthun (1990). The cells (8 × 500 ml) were collected, resuspended in 100 ml of 10 m*M* Tris–HCl pH 7.6, 1 m*M* EDTA, 1 m*M* β-mercaptoethanol, and disrupted in a French press at 5000 psi. The supernatant was removed after a 277 K centrifugation at 12 000 rev min^{-1} for 30 min in a Sorvall SS34 rotor and then heated for 30 min at 363 K. The denatured proteins were removed by centrifugation (12 000 rev min^{-1} for 30 min) and the supernatant was applied to a chromatofocusing column containing 25 ml of polybuffer exchanger, PBE 94 (Pharmacia) equilibrated with 25 m*M* His–HCl pH 6.2. The column was washed with 10–20 volumes of the equilibration buffer. The *T. aquaticus* GAPDH was eluted with Pharmacia Polybuffer 74 pH 4.0 and collected into tubes containing 1/10th volume of 1.0 *M* Tris–HCl pH 8.7.

SDS–PAGE of the purified protein showed a single protein at 36 000. The encoded GAPDH gene expressed an enzyme that exhibited the same thermostability as the wild-type enzyme, *i.e.* stable at 363 K and a 50% loss of enzyme activity resulted after 30 min at 373 K. This indicates that the thermostability of the expressed *T. aquaticus* GAPDH gene is derived from its primary sequence rather than factors intrinsic to the environment of its normal host.

Crystallization of GAPDH

Crystallization at room temperature was accomplished as follows. First, a 20 mg ml^{-1} solution of the enzyme was prepared in 100 m*M* NaCl, 1 m*M* β-mercaptoethanol, 1 m*M* EDTA, 2 m*M* NAD, 2.5 m*M* Tris, pH 7.7. We verified that this solution was monodisperse with respect to enzyme by quasi-elastic light scattering using a Biotage Model 801 Molecular Size Detector. Crystals were grown using the method of sitting-drop vapor diffusion by mixing equal volumes of the protein solution and a reservoir solution containing 15% isopropanol, 20–24% polyethylene glycol 3000 (Fluka) and 0.1 *M* Hepes pH 8.2–8.4. In 2–4 weeks, crystals grew as clear bricks with the largest crystals having dimensions 0.6 × 0.4 × 0.2 mm (Fig. 1).

X-ray data collection

Suitable crystals were mounted in a glass capillary along with a small amount of mother liquor and sealed with wax. We collected low-resolution native data using a FAST area detector on a Rigaku RU-200 rotating anode operating at 50 kV and 180 mA. Graphite monochromatization was employed along with a 0.5 mm collimator. The data set used for the rotation-function calculations described below was collected in two scans of 0.1° frames. The first scan was around c^* and extended over 130° whereas the second scan of 80° was around a^*. The frame exposure time was 60 s, and the crystal-to-detector distance and detector θ angle were 146.3 mm and −15°, respectively. The data-collection temperature was 289 K. These data were collected and processed using *MADNES* (Messerschmidt & Pflugrath, 1987) and *XSCALE* (Kabsch, 1988).

Additional data were collected on an AFC-5R diffractometer and on an Enraf–Nonius precession camera mounted on a FR590 sealed-tube machine. The AFC-5R, which is equipped with a 600 mm 2θ arm, was mounted on the RU-200 opposite to the FAST. Standard Rigaku graphite monochromatization was employed and a 0.5 mm collimator was used. We analyzed the AFC-5R data using *TEXSAN* (Molecular Structure Corporation, 1989). The FR590 was operated at 40 kV and 45 mA. Data were collected using a one-dot collimator and built-in nickel filter.

Rotation-function calculations

Cross-rotation-function searches were conducted using *X-PLOR* version 3.0 (Brünger, 1990, 1992) on a Cray YMP-EL computer. The search model was the *B. stearothermophilus* GAPDH tetramer oriented so that the *P*, *Q* and *R* molecular twofold axes were coincident with the *a*, *b* and *c* axes, respectively. The rotated Patterson map was calculated following placement of the search model in an orthorhombic box with dimensions 110 × 90 × 100 Å. The Patterson integration radius was 5–23 Å, and the 6–4 Å shell of the area-detector data set was used. The asymmetric unit of rotational space, $0 \le \theta_1 < 2\pi$, $0 \le \theta_2 \le \pi/2$, $0 \le \theta_3 < \pi$ (Rao, Jih & Hartsuck, 1980), was searched using an angular grid interval of 3°. Unless noted otherwise, all rotation angles are Euler angles in the *ZXZ* convention (Goldstein, 1980).

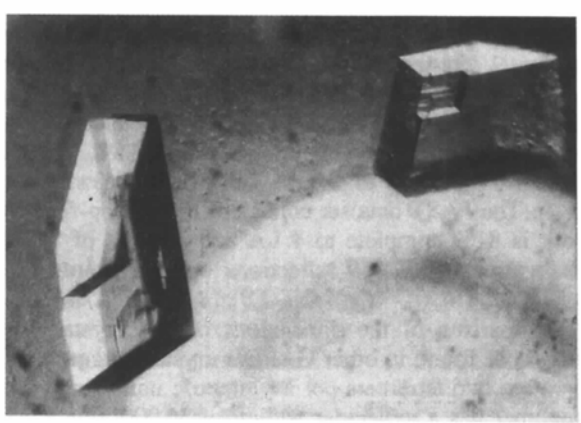

Fig. 1. Crystals of glyceraldehyde 3-phosphate dehydrogenase from *Thermus aquaticus* grown in isopropanol and polyethylene glycol 3000. The larger crystal has dimensions 0.6 × 0.4 × 0.2 mm.

Patterson-correlation (PC) refinement was performed in *X-PLOR* (Brünger, 1990), and consisted of 20 steps of rigid-body optimization of the tetramer orientation against an effective energy function that is proportional to the standard linear correlation coefficient between the square of the normalized observed structure-factor amplitudes and the square of the normalized calculated structure-factor amplitudes (Brünger, 1992).

Results and discussion

We have described the crystallization of GAPDH from *T. aquaticus* using a mixture of polyethylene glycol and isopropanol. It is of interest to note that all previous diffraction-quality GAPDH crystals have been obtained using ammonium sulfate as the precipitating agent. Whereas Hocking & Harris (1973) and Harris, Hocking, Runswick, Suzuki & Walker (1980) reported growth of microcrystals of *T. aquaticus* GAPDH using ammonium sulfate, our crystallization experiments *versus* ammonium sulfate failed. The exact role of the isopropanol in the crystallization is unclear. While we obtained small crystals using polyethylene glycols without isopropanol, the large crystals were always grown with isopropanol. The fact that our crystals are grown in a medium of relatively low ionic strength may be advantageous in that we should be able to elucidate the role of surface salt links in enzyme thermostability without the interference of high solvent salt concentration. Walker, Wonacott & Harris (1980) note that not all of the possible surface salt links are observed in the *B. stearothermophilus* structure suggesting that artifacts due to the electrostatic influence of the ammonium sulfate in the crystallization mother liquor may be important.

Refined unit-cell parameters for the *T. aquaticus* GAPDH crystals as determined with the AFC-5R were as follows: a = 144.77 (6), b = 148.77 (5), c = 149.50 (7) Å. The unit-cell axes, b and c are similar but precession photographs of the three principal zones revealed *mmm* and not 4/*mmm* symmetry. Precession photographs displaying the 0*kl* and *hk*0 zones are shown in Fig. 2. Systematic absences found along all three reciprocal axes combined with the *mmm* Laue group have allowed for identification of the space group as $P2_12_12_1$. When mounted on the FAST, crystals diffracted strongly to 3.1 Å with some diffraction to 2.8 Å. The FAST data set collected for rotation-function work is 81% complete to 4.1 Å and consists of 87 842 observations of 27 349 reflections with an overall R_{sym} $(= \Sigma_{hkl}\Sigma_{i=1}^{N}|I_{hkl,i} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_{i=1}^{N} I_{hkl,i})$ of 2.7%.

Comparison of the dimensions of this crystal form with those found in other GAPDH crystals suggests that there are two tetramers per asymmetric unit. Since each monomer has a molecular weight of 36 000, this would result in a V_m value of 2.8 Å^3 Da^{-1}, yielding 56% solvent in the unit cell (Matthews, 1968). The asymmetric unit would then contain almost 300 kDa, thus creating a large molecular replacement calculation, the results of which we discuss below.

Analysis of the rotation-function data has allowed us to determine the orientation of these two tetramers in the unit cell. The prominent features of the rotation function occurred near the $\theta_3 = 0°$ level, and we show this level in the contour plot in Fig. 3. There are several peaks with nearly the same rotation-function value in Fig. 3,

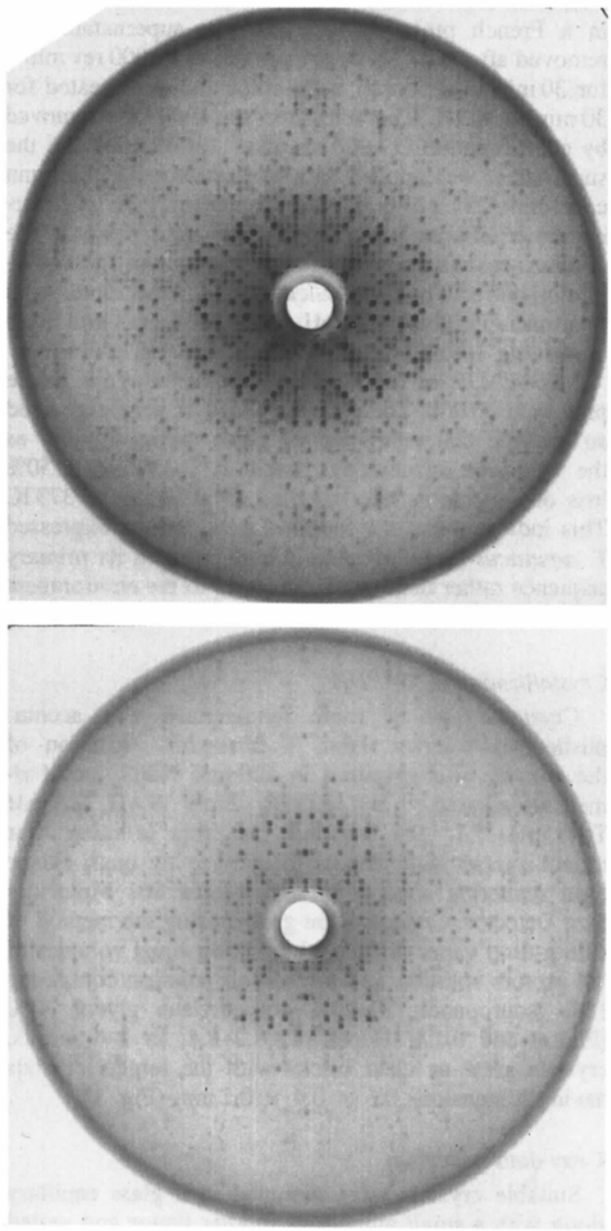

Fig. 2. Precession photographs of the 0*kl* zone (top) and *hk*0 zone (bottom). The crystal-to-film distance was 100 mm and the precession angle was 12.4°. The exposure time was 24 h. b^* is horizontal in each photo, while c^* is vertical in the 0*kl* zone and a^* is vertical in the *hk*0 zone.

and they appear mainly in two groups, near $\theta_1 = 90$ and 270°. Additional cross-rotation searches in which we varied the amount of diffraction data used, the length of the Patterson integration radius and the size of the box for calculation of the rotated Patterson map produced contour plots similar to that of Fig. 3. The lists of the highest ten peaks resulting from these calculations had many orientations in common; however, the order of the peaks differed from list to list.

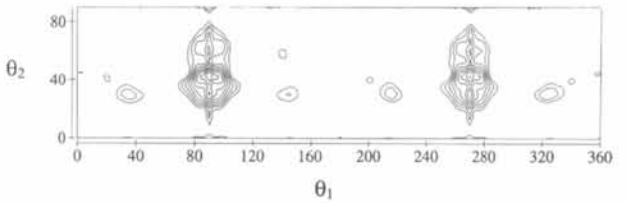

Fig. 3. Contour plot of the $\theta_3 = 0$ level of a conventional cross-rotation function. Only the asymmetric unit of rotational space for θ_1 and θ_2 is shown. See text for details of the calculation.

Since the cross-rotation-function peak list was sensitive to the input parameters, we performed Patterson-correlation (PC) refinement (Brünger, 1990) of the highest peaks of the first cross-rotation function. The rotation-function peaks were arranged in clusters using *X-PLOR* cluster analysis with $\varepsilon = 0.35$. Clustering reduces the number of peaks to be refined by grouping peaks that differ by less than about 15° in the same cluster. The highest 200 'clustered' peaks were input to PC refinement. Refinement of the 200 peaks using the 6–4 Å shell of diffraction data required 25 h of Cray YMP-EL central processor time. For comparison, a typical cross-rotation function required 30 min of central processor time.

The results of PC refinement are shown in Fig. 4, where we plot the rotation-function value before PC refinement (Fig. 4*a*) and the PC value after refinement (Fig. 4*b*) as functions of the peak index. By construction, the rotation-function value decreases with increasing peak index. Choice of the correct rotation-function peaks is hampered by the fact that the rotation-function values of first peak and the 200th peak differ by a factor of only 1.1 (Fig. 4*a*). After PC refinement, however, five peaks clearly stood out above the noise level by factors of 2–5 (Fig. 4*b*). The refined orientations of these five peaks are listed in parentheses beside their respective peaks in Fig. 4(*b*).

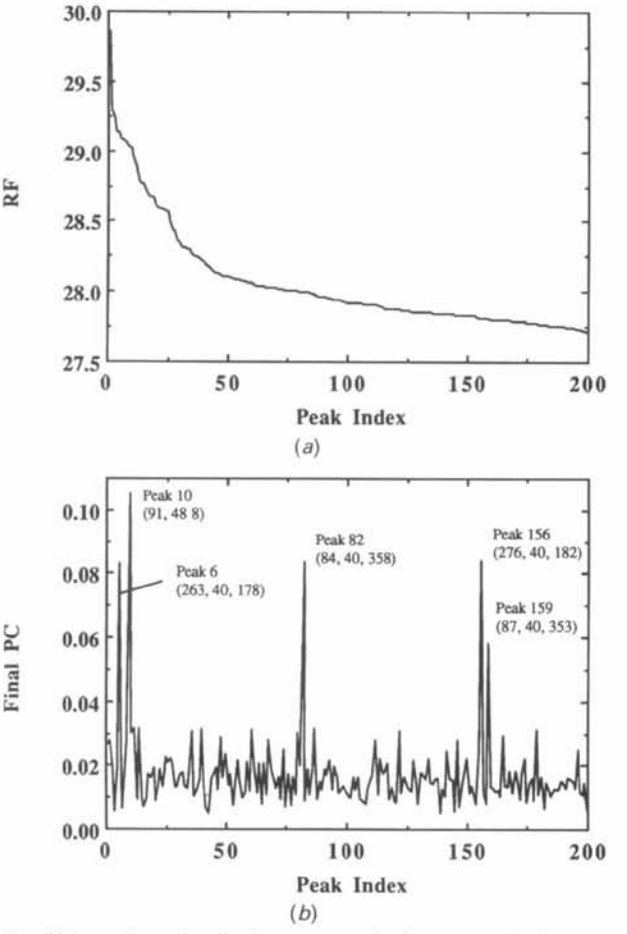

Fig. 4. Comparison of results from a conventional cross-rotation-function search (*a*), and from Patterson-correlation refinement (*b*). The conventional rotation-function value (RF) and the final Patterson-correlation value (PC) are plotted as functions of the cross-rotation-function peak index. The numbers in parentheses in plot (*b*) are the Euler angles of selected peaks after Patterson-correlation refinement.

Inspection of the refined angles listed in Fig. 4(*b*) shows that three peaks, 10, 82 and 159, represent the same orientation. Since peak 10 possesses the largest PC value, we believe that it represents the orientation of one of the tetramers in the asymmetric unit. If the tetramer search model is placed in this orientation (91, 48, 8°), then the Q molecular twofold axis is nearly coincident with the crystallographic a axis. Thus, the P and R molecular axes lie almost in the plane formed

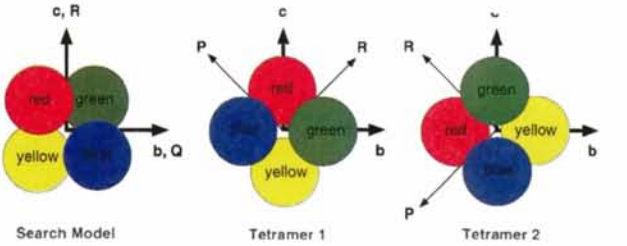

Fig. 5. Diagram explaining the non-crystallographic symmetry. Following Moras *et al.* (1975), the monomers are represented by red, green, blue and yellow circles. Thick arrows represent crystallographic axes and thin arrows represent molecular twofold axes. The crystallographic a axis is perpendicular to the page and points toward the viewer. In the search model, the P molecular axis is aligned with the crystallographic a axis. In tetramer 1 and tetramer 2, the Q axis is nearly coincident with the a axis.

by the b and c axes. If the crystallographic coordinate system is rotated so that the positive a axis points toward the viewer, then the R molecular axis will be offset from the positive c axis by a clockwise rotation of approximately 45°, and the P molecular axis will be offset from the positive c axis by a counterclockwise rotation of approximately 45° (Fig. 5, tetramer 1).

The two remaining peaks with high PC values, 6 and 156, have similar refined Euler angles and nearly identical refined PC values. If the search model is oriented according to either of these peaks, then the Q molecular axis is nearly coincident with the crystallographic a axis (Fig. 5, tetramer 2). Thus, both tetramers in the asymmetric unit have their Q molecular twofold axes nearly coincident with the a axis. The two tetramers in the asymmetric unit are related by a rotation of approximately 90° about the crystallographic a axis (Fig. 5).

The non-crystallographic symmetry determined from the cross-rotation-function calculations is in agreement with the results of a self-rotation function, calculated using the 15–4 Å shell of diffraction data, a Patterson integration radius of 5–40 Å and an angular grid with spacing of 5°. The highest non-self peak of this rotation function was at spherical polar angle (ψ, φ, κ) = (90, 0, 90°), which is a rotation of 90° about the a axis, and had a height of 2σ above the mean.

It is interesting to note that the correct solutions to the rotation problem were not among the highest five peaks of the cross-rotation function. Subsequent Patterson-correlation refinement appears to have suppressed noise peaks and led to the present solutions.

The authors wish to thank Dr Pluckthun for his gift of strain W3CG. This work was supported by grants from the National Science Foundation (to RMH), the US Army Research Office (to RMH), the National Institutes of Health (to KLK) and the W. M. Keck Foundation. An academic research infrastructure program grant from the National Science Foundation is also acknowledged.

References

ARGOS. P., ROSSMANN, M. G., GRAU, U. M., ZUBER, H., FRANK, G. & TRATSCHIN, J. D. (1979). *Biochemistry*, **18**, 5698–5703.

BRÜNGER, A. T. (1990). *Acta Cryst.* A**46**, 46–57.

BRÜNGER, A. T. (1992). *X-PLOR Manual*. Version 3.0. Yale Univ., New Haven, Connecticut, USA.

GANTER, C. & PLUCKTHUN, A. (1990). *Biochemistry*, **29**, 9395–9402.

GOLDSTEIN, H. (1980). *Classical Mechanics*, 2nd ed., pp. 143–148. Reading, Massachusetts: Addison-Wesley.

GRIFFITH, J. P., LEE, B., MURDOCK, A. L. & AMELUNXEN, R. E. (1983). *J. Mol. Biol.* **169**, 963–974.

HARRIS, J. I., HOCKING, J. D., RUNSWICK, M. J., SUZUKI, K. & WALKER, J. E. (1980). *Eur. J. Biochem.* **108**, 535–547.

HARRIS, J. I. & WATERS, M. (1976). *The Enzymes*, Vol. 13, edited by P. D. BOYER, pp. 1–50. New York: Academic Press.

HECHT, R. M., GARZA, A., LEE, Y., MILLER, M. D. & PISEGNA, M. A. (1989). *Nucleic Acids Res.* **17**, 10123.

HOCKING, J. D. & HARRIS, J. I. (1973). *FEBS Lett.* **34**, 280–284.

HOCKING, J. D. & HARRIS, J. I. (1980). *Eur. J. Biochem.* **108**, 567–579.

KABSCH, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

LIN, Z., LI, J., ZHANG, F., SONG, S., YANG, J., LIANG, S. & TSOU, C. (1993). *Arch. Biochem. Biophys.* **302**, 161–166.

MATTHEWS, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

MENENDEZ-ARIAS, L. & ARGOS, P. (1989). *J. Mol. Biol.* **206**, 397–406.

MERCER, W. D., WINN, S. I. & WATSON, H. C. (1976). *J. Mol. Biol.* **104**, 277–283.

MERKLER, D. J., FARRINGTON, G. K. & WEDLER, F. C. (1981). *Int. J. Pept. Protein Res.* **18**, 430–442.

MESSERSCHMIDT, A. & PFLUGRATH, J. W. (1987). *J. Appl. Cryst.* **20**, 306–315.

Molecular Structure Corporation (1989). *TEXSAN. TEXRAY Structure Analysis Package*. Version 5.0. MSC, 3200A Research Forest Drive, The Woodlands, TX 77381, USA.

MORAS, D., OLSEN, K. W., SABESAN, M. N., BUEHNER, M., FORD, G. C. & ROSSMANN, M. G. (1975). *J. Biol. Chem.* **250**, 9137–9162.

MURTHY, M. R. N., GARAVITO, R. M., JOHNSON, J. E. & ROSSMANN, M. G. (1980). *J. Mol. Biol.* **138**, 859–872.

PERUTZ, M. F. (1978). *Science*, **201**, 1187–1191.

RAO, S. N., JIH, J.-H. & HARTSUCK, J. A. (1980). *Acta Cryst.* A**36**, 878–884.

SKARZYNSKI, T., MOODY, P. C. E. & WONACOTT, A. J. (1987). *J. Mol. Biol.* **193**, 171–187.

SKARZYNSKI, T. & WONACOTT, A. J. (1988). *J. Mol. Biol.* **203**, 1097–1118.

SUZUKI, K. & HARRIS, J. I. (1971). *FEBS Lett.* **13**, 217–220.

VELLIEUX, F. M. D., HAJDU, J., VERLINDE, C. L. M. J., GROENDIJK, H., READ, R. J., GREENHOUGH, T. J., CAMPBELL, J. W., KALK, K. H., LITTLECHILD, J. A., WATSON, H. C. & HOL, W. G. J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 2355–2359.

WALKER, J. E., WONACOTT, A. J. & HARRIS, J. I. (1980). *Eur. J. Biochem.* **108**, 581–586.