Volume 292, number 1,2, 237–242 FEBS 10305 November 1991

ADONIS 001457939101013J

The crystal structure of fructose-1,6-bisphosphate aldolase from *Drosophila melanogaster* at 2.5 Å resolution

Gerko Hester[1], Olga Brenner-Holzach[2], Franco A. Rossi[1], Martina Struck-Donatz[1], Kaspar H. Winterhalter[1], Jan Derk G. Smit[1,*] and Klaus Piontek[1]

[1]*Laboratorium für Biochemie I, Eidgenössische Technische Hochschule Zürich, Universitätstr. 16, CH-8092, Zürich, Switzerland* and [2]*Biochemisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland*

Received 13 August 1991

The structure of fructose-1,6-bisphosphate aldolase from *Drosophila melanogaster* has been determined by X-ray diffraction at 2.5 Å resolution. The insect enzyme crystallizes in space group $P2_12_12_1$ with lattice constants a=86.60, b=116.80, c=151.58 Å. Molecular replacement with rabbit muscle aldolase as a search model has been employed to solve the structure. To improve the initial phases real space averaging, including phase extension from 4.0 to 2.5 Å, has been applied. Refinement of the atomic positions by molecular dynamics resulted in a crystallographic R-factor of 0.214. The tertiary structure resembles in most parts that of the vertebrate aldolase from rabbit muscle. Significant differences were found in surface loops and the N- and C-terminal regions of the protein. Here we present the first aldolase structure where the functionally important C-terminal arm is described completely.

Crystal structure; Glycolytic enzyme; Fructose-1,6-bisphosphate aldolase; Substrate specificity; *Drosophila melanogaster*

1. INTRODUCTION

Fructose-1,6-bisphosphate (FBP) aldolase (EC 4.1.2.13) is a glycolytic enzyme catalyzing the reversible aldol cleavage of fructose-1,6-bisphosphate to dihydroxyacetone phosphate and D-glyceraldehyde-3-phosphate. FBP aldolase from *Drosophila melanogaster* (strain Sevelen, wild type, pupea) forms a Schiff base between the substrate and a lysyl residue [1] and belongs therefore to the class I aldolases [2]. Contrary to class I the class II aldolases use a divalent metal ion as cofactor [3]. The vertebrate aldolases (class I) comprise three isozymes. They refer to A (muscle), B (liver) and C (brain) depending on the tissue from which they are isolated [3]. The A and C form show a higher substrate specificity for FBP over fructose-1-phosphate (F1P), while the liver isozyme has no such preference.

Abbreviations: DALD, fructose-1,6-bisphosphate aldolase from *Drosophila melanogaster*; RMALD, rabbit muscle aldolase; HMALD, human muscle aldolase; RLALD, rabbit liver aldolase; FBP, fructose-1,6-bisphosphate; F1P, fructose-1-phosphate; r.m.s., root-mean-square.

Sequence numbering: Throughout this paper the sequence numbering of FBP aldolase from rabbit muscle will be applied.

**Present address:* Pharma division, Research Technologies, Hoffman-La Roche Ltd., CH-4002 Basel, Switzerland.

Correspondence address: K. Piontek, Laboratorium für Biochemie I, Eidgenössische Technische Hochschule Zürich, Universitätstrasse 16, CH-8092 Zürich, Switzerland. Fax: (41) (1) 261 5677.

FBP aldolases are tetramers of four identical subunits with a molecular weight of about 158 kDa. Each subunit of the insect enzyme consists of 360 amino acids. The primary structure of DALD is reported in [4]. Compared to vertebrate aldolases the insect enzyme shows 62–70% identical amino acids [4]. Based on biochemical properties and primary sequence alignments it was suggested [4] that DALD is closest to the muscle type aldolases of vertebrates. DALD is able to form enzymatically active interspecies hybrids with vertebrate aldolases [5].

The C-terminal part of the enzyme is essential for the catalytic activity [6]. Enzymatic activity is abolished when the last amino acid, a completely conserved tyrosine, is cleaved off. It is therefore suspected that the C-terminus mediates the alignment and/or attachment of the substrate during catalysis. On the other hand sequence homology is low within the C-terminal residues (345–363) of aldolases. In fact, two of the three deletions in DALD with respect to RMALD are in this part of the polypeptide chain. It has been suggested [6,7] that this low sequence homology is responsible for the different specific activities of aldolases towards their substrates.

Crystals of DALD have been grown and characterized previously [8]. The three-dimensional structures of rabbit muscle aldolase and human muscle aldolase have been determined recently at 2.7 and 3.0 Å resolution [7,9], respectively. The basic feature of the FBP aldolase subunit is the eight stranded α/β-barrel found before also in other glycolytic enzymes [10–12]. Information on

the conformation of the C-terminus from human muscle aldolase is lacking. In addition, parts of this region in the rabbit muscle aldolase structure are not well determined [7], leaving an uncertainty concerning its conformation. Even though these two known structures were not compared to each other the high degree of identity (99%) [4] implies identical tertiary structure [13].

Considering the expected structural differences between DALD and the vertebrate muscle type aldolases, especially in the functionally important C-terminal part, an X-ray structure determination of DALD has been carried out. This paper is focussed on the crystal structure determination of an insect aldolase (DALD) at 2.5 Å resolution and the comparison to the vertebrate aldolase of rabbit muscle.

2. EXPERIMENTAL

2.1. *Crystallization*

DALD crystallizes [8] in space group $P2_12_12_1$ with a=86.60 Å; b=116.80 Å; c=151.58 Å. With one tetramer per asymmetric unit a V_m value of 2.4 Å^3/Da is calculated. This is well within the normal range for globular proteins [14]. Density measurements confirmed this result [8].

2.2. *Data collection and processing*

Data were collected on oscillation photographs at the Synchotron Radiation Source (SRS), in Daresbury, UK. During data collection crystals were held at 4°C. The data were processed using the MOSCO program package [15,16]. Scaling, merging and postrefinement were carried out with the CCP4 program suite. The results are summarized in Table I.

2.3. *Structure determination and refinement*

The DALD structure was solved by molecular replacement methods using the RMALD structure as model. The coordinates of the RMALD structure were kindly provided by Prof. J. Sygusch (University of Sherbrooke, Québec, Canada).

Rotation functions [17] were employed to determine the molecular symmetry of the tetrameric DALD molecule. Self-rotation functions (6–4 Å shell) revealed the orientation of three local twofold axes displaying 222-symmetry. They are close to the crystallographic axes (respectively 12.0°, 6.0° and 13.4° from the a,b,c axes) and constitute an orthogonal set.

Cross-rotation functions (6–4 Å shell) were calculated to determine the orientation of the DALD structure in the unit cell. The RMALD molecule was placed in a $P1$ cell with a=b=c=120 Å, α=β=γ=90°. The local 2-fold axes were oriented along the cell edges. They were in turn superimposed on the directions corresponding to the peaks found in the self-rotation function. This resulted for each superposition of a local 2-fold axis in improved orientations for the two other axes. The above procedure was successfully applied to find the orientation of the DALD molecule in the unit cell with respect to the RMALD molecule.

The position of the DALD molecule in the unit cell was determined by applying the translation function of the program package MERLOT [18]. The model was oriented according to the results from the cross-rotation function studies and placed at the origin of the same $P1$ cell as above. The relevant Harker sections of the translation function showed single high peaks of 11–13 r.m.s. resulting in an unique solution. The obtained orientation and position were refined using a R-factor minimisation procedure as implemented in one of the programs of the MERLOT program package. The refined molecular replacement solution gave a R-factor (defined as $\Sigma(||F_o|-|F_c||)/\Sigma|F_o|$) of 45.4% in the 6–3 Å resolution range. The packing of the tetramers (based on the RMALD model) in the DALD unit cell was checked and showed only very few bad contacts related to non-conserved amino acids.

A $2F_o$–F_c map based on data up to 3 Å resolution was calculated. The map was of such a quality that differences from the RMALD structure could be seen, even though there were ambiguous areas.

The presence of four chemically identical subunits in the asymmetric unit provided the possibility to improve the quality of the phases by applying the technique of real space averaging [19]. Initially only phases up to 4.0 Å resolution were accepted from the model phasing. R-Factor calculations as a function of resolution indicated that phases beyond this value were not reliable. Phases were iteratively refined and resolution was extended from 4.0 to 3.0 Å in six steps. The molecular envelope within which the electron density was averaged was constructed on the basis of main chain and C_β-atoms of the RMALD model. Electron density outside this region was set to zero.

For the averaging procedure the program package RMOL was used, kindly provided by Dr. B. Rees (IMBC, Strasbourg, France). RMOL is an extended version of a program written by Dr. J.E. Johnson (Purdue University, West Lafayette, USA) [20]. We introduced a modification of the program package RMOL so a weighting scheme as suggested by Rayment [21] could be applied. The calculation of these weights is based on the agreement of the observed amplitudes to the calculated amplitudes from the Fourier transform of the averaged, solvent leveled map. The weights were then applied to the observed amplitudes and used together with calculated phases to compute new electron density maps.

The averaged electron density map up to 3.0 Å resolution was easily interpreted in terms of the DALD sequence. A model was built including all residues but the last 15 amino acids. Inspection of the electron density map together with the envelope used for the averaging showed that this portion of the structure was mostly outside the area of the envelope. An electron density map based on phases from this partial model confirmed this. Thus an improved envelope was computed and averaging continued with phase extension from 3.0 to 2.5 Å in five steps. The final R-factor was 20.0% in the 30.0–2.5 Å resolution shell. The overall r.m.s. phase change from the initial phases was 85.7°. The result of this procedure was an averaged electron density map of high quality with interpretable density for most of the missing C-terminal part. Only the last three amino acids could not be assigned to this density.

The model was refined using the program XPLOR [22]. Three rounds of molecular dynamics with simulated annealing and refinement cycles were performed with non-crystallographic symmetry constraints. After the second round the last three missing residues could also be assigned to a $2F_o$–F_c electron density map. The current model of DALD comprises all 360 residues and 112 water molecules per subunit with an overall R-factor of 21.4% for all the observed 51 148 reflections between 8 and 2.5 Å resolution. The model has good geometry with r.m.s. deviations from ideal geometry of 0.005 Å and 1.456° for bond lengths and angles, respectively. Individual temperature factor refinement was included at this stage. The average B-factor is 14.4 Å^2. Further refinement of the model is now in progress.

Table I

Data collection and processing of *Drosophila* aldolase

No. of photographs	120
Resolution (Å)	30.0–1.9
No. of measurements	377 901
No. of unique reflections	123 654
The theoretically possible no. of reflections (%)	87.0[a]
Scaling R-factor[b] (%)	9.7

[a]The percentage of the theoretically possible no. of reflections up to 2.01 Å is 96.1%.

[b] $$R = \frac{\sum_h |(\langle I_h \rangle - I_h)|}{\sum_h I_h} \times 100\%$$

where $\langle I_h \rangle$ is the mean of the I_h observations of reflection h.

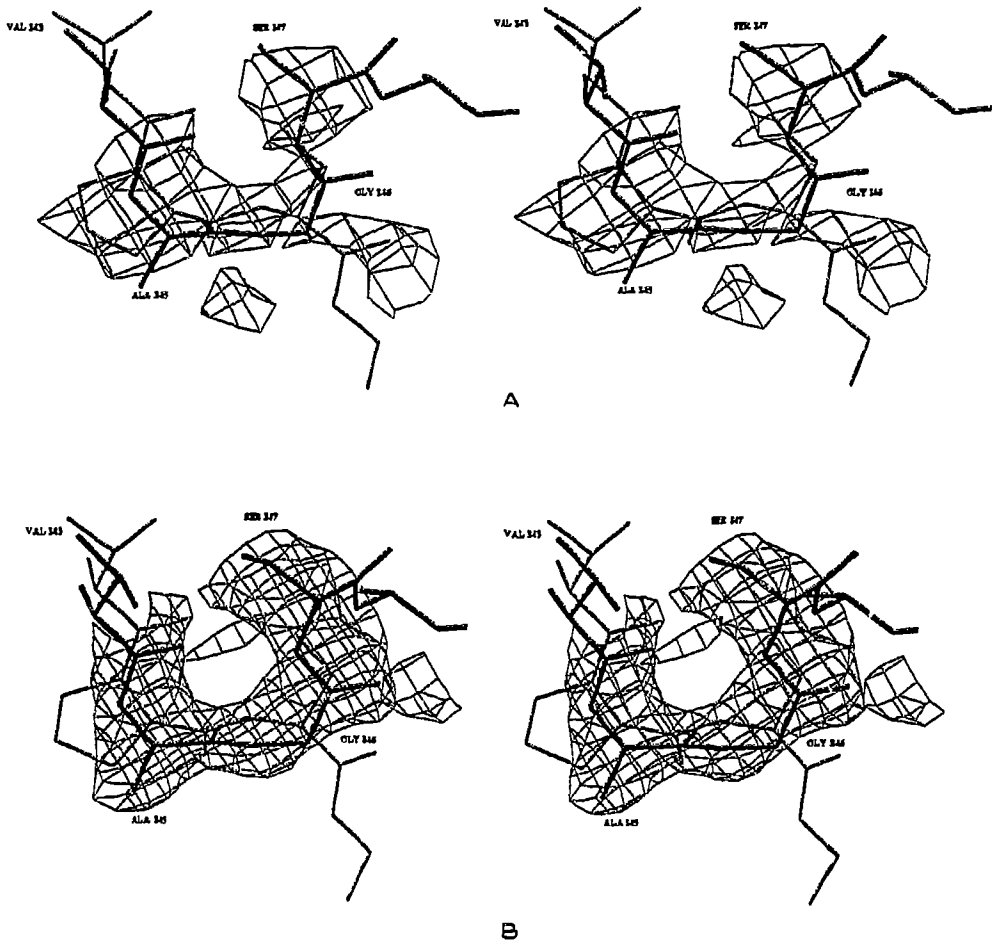

Fig. 1. Residues Val-343 to Ala-348 of DALD (thick lines) and the equivalent residues of RMALD (thin lines). This part forms the beginning of the C-terminal arm. (A) Electron density map in this region before real space averaging. (B) Electron density map after real space averaging. Both maps are contoured at 1.5 σ.

3. RESULTS AND DISCUSSION

Molecular replacement is usually complicated by the bias of the model structure used to calculate a first set of approximate phases. The errors introduced into the electron density can lead to interpretational conflicts in regions where the two structures differ. Here the electron density of the model structure might be still visible in addition to electron density for the unknown structure.

Real space averaging and phase extension from 4.0 to 2.5 Å, which allowed ab initio phasing in this range of resolution, was successfully applied here to completely eliminate the model bias. It is demonstrated in the region of the electron density (Fig. 1A,B) corresponding to the beginning of the C-terminal arm. This part of the polypeptide chain consists of residues 345–363 and shows considerable structural differences between the two aldolases.

The core of the DALD structure forms an eight stranded α/β barrel (Fig. 2) as first found in triose phosphate isomerase (TIM) [10]. The overall fold of the two aldolases is rather similar. The r.m.s. deviation (Fig. 3) of the main chain positions between one subunit of DALD and RMALD is 3.89 Å. If residues 1–9 and 345–363 (the two regions where the structures deviate most) are left out this value becomes 0.75 Å. When only residues 345–363 are used for the comparison the r.m.s. deviation is 22.75 Å.

Antibodies formed against DALD do not react with vertebrate muscle aldolases [23]. There are two regions likely to be responsible for this: residues 235–245 and residues 83–94. These are loops on the surface of the protein which are significantly different in the two structures (Figs. 3 and 4).

The biggest r.m.s. deviations between DALD and RMALD for the region formed by residues 235–245 are found for residues 239–243. There the sequence differs considerably [4]: in DALD residue 239 is missing and residues 240–243 are Ala, Lys, Lys, Asn. For RMALD the sequence for residues 239–243 is Cys, Thr, Gln, Lys, Tyr.

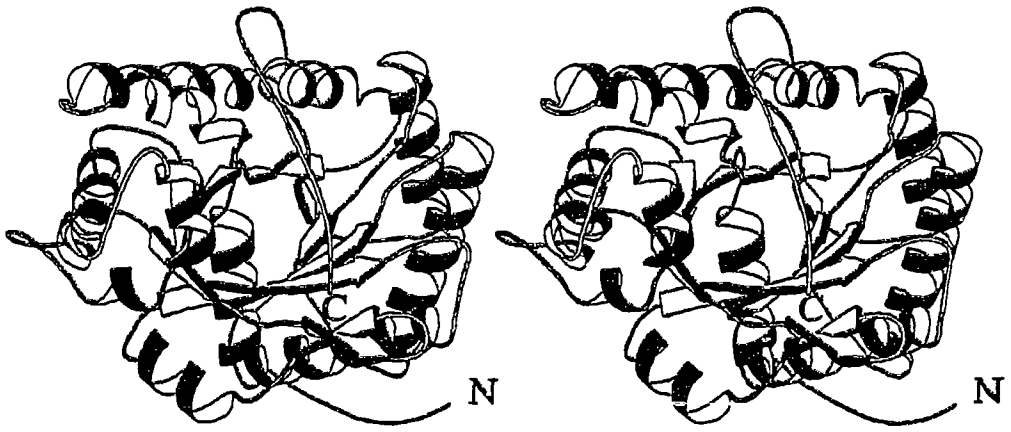

Fig. 2. Ribbon diagram of one subunit of *Drosophila* aldolase. The view is into the active site cavity. The C-terminus (C) is positioned over the active site cleft in the middle of the picture. The N-terminus (N) is at the lower right corner of the picture.

The second candidate for the antigenic differences is a loop formed by residues 83–94. The sequence and conformation difference is less pronounced as for residues 235–245.

One explanation for the different conformations in these two loops could be crystallographic contacts. But examination of crystal packing shows that the intermolecular contacts for the two loop regions resemble each other. The different conformations are therefore not necessarily caused by packing effects.

Based on the current results the different antigenic properties of DALD compared to RMALD cannot definitively be assigned, but the loop formed by residues 235–245 seems the most probable candidate.

In DALD the arm-like polypeptide chain, formed by residues 345–363, folds over the active site cleft, thus

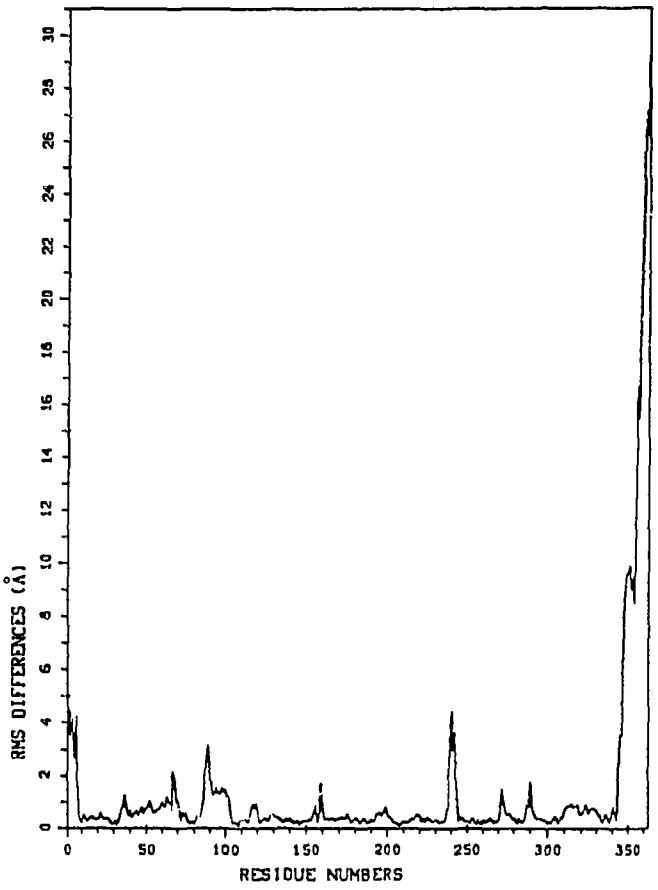

Fig. 3. Comparison of main chain atoms of RMALD with corresponding positions of DALD.

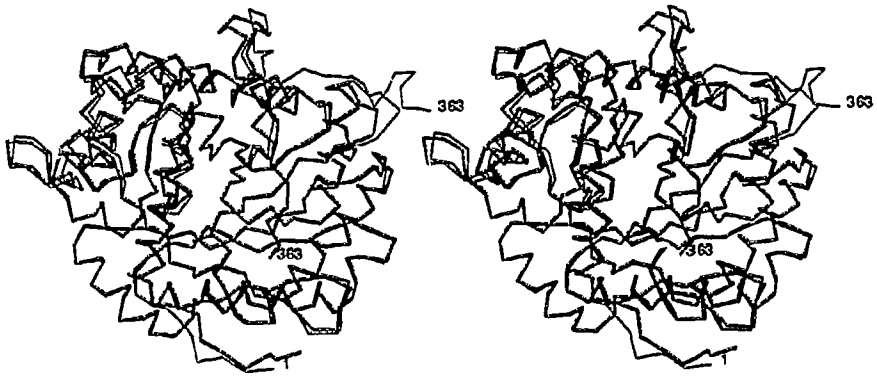

Fig. 4. C_α stereo-plot of one subunit of DALD (thick lines) superimposed with one subunit of RMALD (thin lines).

inhibiting the access of substrate and release of product. RMALD exhibits an open conformation for the equivalent residues (Fig. 5). There the first six residues of the C-terminal part form a short stretch having a similar conformation as in DALD, followed by a sharp turn at residue 350. The polypeptide chain then continues straight, projecting out into the solvent region while in DALD the C-terminus approaches the Schiff base forming residue lysine-229 in the active site cleft.

The region between residues 346 and 352 is rich on alanine and glycine residues in all known aldolase sequences, indicating a high flexibility for this part of the polypeptide chain. A rotation of 70.0° about the C_α–C bond of residue 349 of RMALD brings the C-terminal regions of the two aldolases close together. The r.m.s. deviation for residues 345–363 then becomes only 4.96 Å when the two structures are compared. Therefore, it is highly probable that these residues, particularly Gly-349, serve as a hinge for the C-terminal arm in aldolases. Thus, the two observed conformations in the DALD and RMALD structures might represent the two major conformational states of the C-terminal region during catalysis.

Recently, the aldolase structure from rabbit liver (RLALD) has been determined (J. Sygusch, personal communication). It shows a similar closed conformation of the C-terminal arm as in DALD. The FBP/F1P activity ratio in muscle type aldolases is 50–100 while for the liver type aldolases this value is 1–2 [3]. For DALD the corresponding ratio is 20 [24]. Therefore, the substrate specificity for the insect enzyme can be located between the muscle and liver isozymes. These findings might correlate with the closed conformations of the C-terminal arm in the RLALD and DALD native structures, indicating a lower flexibility in the hinge region around Gly-349. It seems reasonable to conclude that the FBP substrate needs a larger opening than the smaller F1P substrate to approach the active site in aldolases and that this can be achieved faster by the more flexible C-terminal arm of the muscle type aldolases.

As the C-terminal arm is probably responsible for the opening and closing of the active site cleft during the catalytic process, the observed structural differences are possibly relevant for the efficiency of catalysis of aldolases.

A more detailed analysis of the three aldolase structures has to be awaited until the refined coordinates of the rabbit liver aldolase will be available. Furthermore, it should be of interest to compare them to the structure of the human muscle enzyme.

Coordinates of the aldolase from *Drosophila melan-*

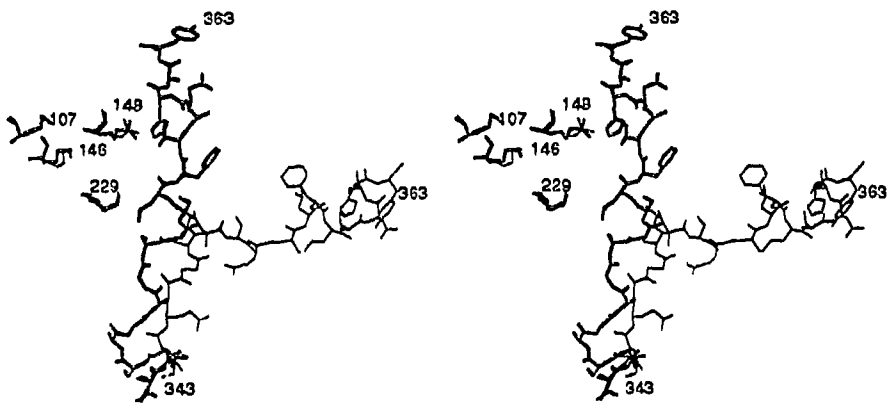

Fig. 5. Polypeptide chain from Val-343 to Tyr-363 of DALD (thick lines) with equivalent residues of RMALD (thin lines). In addition, the side chains of the active site residues Lysine-107, -146, -229 and arginine-148 are shown.

ogaster will be deposited with the Brookhaven Data Base.

Acknowledgements: The authors are indebted to Prof. J. Sygusch for providing the rabbit muscle aldolase coordinates prior to publication. We thank the staff of the S.R.S., Daresbury, U.K. and of the EMBL Outstation, D.E.S.Y., Hamburg, Germany for their assistance in data collection and data processing. O.B.-H. and J.D.G.S. wish to thank Prof. Jeremias Kägi for his continuous interest throughout the biochemical and crystallographic work on *Drosophila* aldolase. The critical reading of the manuscript by Prof. Michael G. Rossmann is greatly appreciated. The financial support to the group is gratefully acknowledged. Specifically salaries were provided by: The Swiss Institute of Technology Zürich (to G.H., M.S.-D. and F.R.), the Roche Research Foundation (to F.R.) and the Swiss National Science Foundation Grant 3100-8390.85 (to G.H., M.S.-D. and F.R.). M.S.-D. acknowledges an EMBO fellowship for a stay at the EMBL Outstation in Hamburg, Germany.

REFERENCES

[1] Brenner-Holzach, O. and Zumsteg, C. (1982) Arch. Biochem. Biophys. 214, 89–101.
[2] Horecker, B.L., Tsolas, O. and Lai, C.Y. (1972) in: The Enzymes, vol. 7 (Boyer, P.D. ed.) pp. 213–258, Academic Press, New York.
[3] Penhoet, E.E., Kochman, M. and Rutter, W.J. (1969) Biochemistry 8, 4396–4402.
[4] Malek, A.A., Hy, M., Honegger, A., Rose, K. and Brenner-Holzach, O. (1988) Arch. Biochem. Biophys. 266, 10–31.
[5] Brenner-Holzach, O. and Leuthardt, D. (1972) Eur. J. Biochem. 31, 423–426.
[6] Takahashi, I., Takasaki, Y. and Hori, K. (1989) J. Biochem. 105, 281–286.
[7] Sygusch, J., Beaudry, D. and Allaire, M. (1987) Proc. Natl. Acad. Sci. USA 84, 7846–7850.
[8] Brenner-Holzach, O. and Smit, J.D.G. (1982) J. Biol. Chem. 257, 11747–11749.
[9] Gamblin, S.J., Cooper, B., Millar, J.R., Davies, G.J., Littlechild, J.A. and Watson, H.C. (1990) FEBS Lett. 262, 282–286.
[10] Alber, T., Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Rivers, P.S. and Wilson, I.A. (1981) Phil. Trans. Roy. Soc. Lond. B293, 159–171.
[11] Stuart, D.I., Levine, M., Muirhead, H. and Stammers, D.K. (1979) J. Mol. Biol. 134, 109–142.
[12] Mavridis, M., Hatada, H., Tulinsky, A. and Lebioda, L. (1982) J. Mol. Biol. 162, 419–444.
[13] Freemont, P.S., Dunbar, B. and Fothergill-Gilmore, L.A. (1988) Biochem. J. 249, 779–788.
[14] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–497.
[15] Arndt, U.W. and Wonacott, A.J. (1977) The Rotation Method in Crystallography, Elsevier, Amsterdam.
[16] Greenhough, T.J. and Helliwell, J.R. (1982) J. Appl. Crystallogr. 15, 493–508.
[17] Rossmann, M.G. and Blow, D.M. (1962) Acta Cryst. 15, 24–31.
[18] Fitzgerald, P.M.D. (1988) J. Appl. Crystallogr. 21, 273–278.
[19] Bricogne, G. (1976) Acta Cryst. A32, 832–847.
[20] Johnson, J.E. (1978) Acta Cryst. B34, 576–577.
[21] Rayment, I. (1983) Acta Cryst. A39, 102–116.
[22] Brünger, A.T., Kuriyan, J. and Karplus, M. (1987) Science 235, 458–460.
[23] Brenner-Holzach, O. and Leuthardt, F. (1969) Helv. Chim. Acta 52, 1273–1281.
[24] Brenner-Holzach, O. and Leuthardt, F. (1971) Helv. Chim. Acta 54, 2809–2820.