Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The structure of human liver fructose-1,6-bisphosphate aldolase

Andrew R. Dalby,[a] Dean R. Tolan[b] and Jennifer A. Littlechild[a]*

[a]Schools of Chemistry and Biological Sciences, University of Exeter, Stocker Road, Exeter EX4 4QD, England, and [b]Department of Biology, Boston University, 5 Cummington Street, Boston, MA 02215, USA

Correspondence e-mail: j.a.littlechild@exeter.ac.uk

Received 6 February 2001
Accepted 26 July 2001

PDB Reference: human liver aldolase, 1qo5.

The X-ray crystallographic structure of the human liver isozyme of fructose-1,6-bisphosphate aldolase has been determined by molecular replacement using a tetramer of the human muscle isozyme as a search model. The liver aldolase (B isozyme) crystallized in space group $C2$, with unit-cell parameters a = 291.1, b = 489.8, c = 103.4 Å, α = 90, β = 103.6, γ = 90°. These large unit-cell parameters result from the presence of 18 subunits in the asymmetric unit: four catalytic tetramers and a dimer from a fifth tetramer positioned on the twofold crystallographic axis. This structure provides further insight into the factors affecting isozyme specificity. It reveals small differences in secondary structure that occur in regions previously determined to be isozyme specific. Two of these regions are at the solvent-exposed enzyme surface away from the active site of the enzyme. The most significant changes are in the flexible C-terminal region of the enzyme, where there is an insertion of an extra α-helix. Point mutations of the human liver aldolase are responsible for the disease hereditary fructose intolerance. Sequence information is projected onto the new crystal structure in order to indicate how these mutations bring about reduced enzyme activity and affect structural stability.

1. Introduction

Type I fructose-1,6-bisphosphate aldolase is an enzyme of the glycolytic and gluconeogenic pathways. It is characterized by Schiff-base formation at a catalytic lysine residue in the active site of the enzyme. Aldolase catalyses the reversible cleavage of fructose-1,6-bisphosphate (FBP) or fructose-1-phosphate to dihydroxyacetone phosphate and either glyceraldehyde-3-phosphate or glyceraldehyde, respectively. Aldolase occurs as three different tissue-specific isozymes in mammals (Penhoet *et al.*, 1966). The A isozyme is found in muscle and erythrocytes, the B isozyme is found in the liver and kidney and the C isozyme is found in the brain. As well as being tissue-specific, these isozymes have different catalytic activities with respect to the natural substrates fructose-1,6-bisphosphate and fructose-1-phosphate. The human muscle isozyme catalyses the cleavage of fructose-1,6-bisphosphate at an increased rate when compared with the liver isozyme, which has no substrate preference. This has resulted in the suggestion that the liver enzyme is associated with gluconeogenesis and the muscle enzyme with glycolysis.

Isozyme-specific regions (ISRs) of the vertebrate aldolase isozyme sequences have been studied by constructing chimeric enzymes between the muscle and liver enzymes (Kitajima *et al.*, 1990; Takasaki & Hori, 1992; Motoki *et al.*, 1993; Kusakabe *et al.*, 1994). The chimeric enzymes showed modest (5–10-fold)

effects on either k_{cat} or K_m and there were some contradictory results where one non-ISR region was modified and showed similar kinetic changes to the ISR chimeras. From these studies, however, four regions (ISR1–4) were proposed to be important for determining isozyme specificity. The first three form a distinct group as they are all clustered on exon 3 of the human FBP-aldolase gene (Fig. 1). The fourth is the flexible C-terminal region which exhibits a large degree of sequence divergence amongst the type I FBP-aldolases, suggesting that this region should be considered separately (Rottmann *et al.*, 1987). Further, detailed analysis of individual isozyme-specific residues for aldolase B and C have been identified (Berardini *et al.*, 1997, 1999), many of which are located within the first three ISR regions. The C-terminal region has been extensively studied (Dreschler *et al.*, 1959; Takahashi *et al.*, 1989; Berthiaume *et al.*, 1991, 1993). It is presumed to be a flexible region, as it is disordered in most aldolase X-ray crystallographic structures (Sygusch *et al.*, 1987; Choi *et al.*, 1999; Dalby *et al.*, 1999). This region has been associated with catalytic activity inferred from studies of C-terminal cleavage (Dreschler *et al.*, 1959; Hartman & Brown, 1976), site-directed mutagenesis (Takahashi *et al.*, 1989; Berthiaume *et al.*, 1991, 1993) and X-ray crystallography (Blom & Sygusch, 1997). Although changes affect activity, particularly the high activity towards fructose-1,6-bisphosphate for the muscle isozyme, it is not essential for catalysis and is more likely to be involved in substrate specificity.

Aldolase has also been associated with nucleotide binding and it was believed that NADH was an allosteric effector of the liver isozyme (Sytnik *et al.*, 1991). It is possible that this binding is at the active site because of the nature of the natural substrate, a sugar phosphate; however, there is also the possibility that the B isozyme may possess a second binding site as the effects of AMP were shown to be non-competitive.

The three-dimensional structure of aldolase has been determined from rabbit muscle (Sygusch *et al.*, 1987), human muscle (Gamblin *et al.*, 1991; Dalby *et al.*, 1999), *Drosophila melanogaster* (Hester *et al.*, 1991) and *Plasmodium falciparum* (Kim *et al.*, 1997). The enzyme is homotetrameric, with a subunit molecular weight of 36 kDa. The subunits are arranged in the tetramer with a local 222 symmetry. The enzyme is an eight-stranded α/β barrel, where the Schiff-base lysine is located in the central cavity of the barrel. The substrate cleft is at the C-terminal end of the β-strands and the N-terminal end of the barrel is blocked by an extra helix that does not form part of the eight-stranded α/β-barrel motif. Structures of complexes of the muscle isozymes with fructose-1,6-bisphosphate (Dalby *et al.*, 1999), dihydroxyacetone phosphate (Blom & Sygusch, 1997) and of an inactivated mutant with fructose-1,6-bisphosphate (Choi *et al.*, 1999) have also been determined. These complex structures have clearly identified the phosphate-binding sites in the enzyme, but they have also suggested that the enzyme has more than one mode of substrate binding and that the substrate may migrate from an initial binding site during catalysis.

Mutations of human aldolase B are responsible for the autosomal recessive disease hereditary fructose intolerance (Tolan, 1995). Previously, fructose intolerance had been diagnosed by the procedures of either fructose challenge or liver biopsy, which expose the patients to high risk (Laméire *et al.*, 1978). The identification of the mutations associated with the condition has allowed new non-invasive tests to be developed (Cross & Cox, 1989; Tolan & Brooks, 1992). The most common mutation, accounting for 57% of the fructose-intolerance cases worldwide, is the mutation of alanine 149 to proline (Cross *et al.*, 1988; Tolan, 1995). There is a global distribution of this mutation and Brooks and Tolan have suggested that subjects carrying this change are all descended from a single mutation event (Brooks & Tolan, 1993). This mutation is in close proximity to the active site of the enzyme. Other mutations have been associated with structural defects that lead to dissociation of the tetramer (Rellos *et al.*, 2000), loss of substrate specificity (Rellos *et al.*, 1999) or loss of enzyme stability (Brooks & Tolan, 1994; Rellos *et al.*, 2000). The determination of the crystal structure of human liver aldolase provides an opportunity to suggest the structural basis for fructose intolerance.

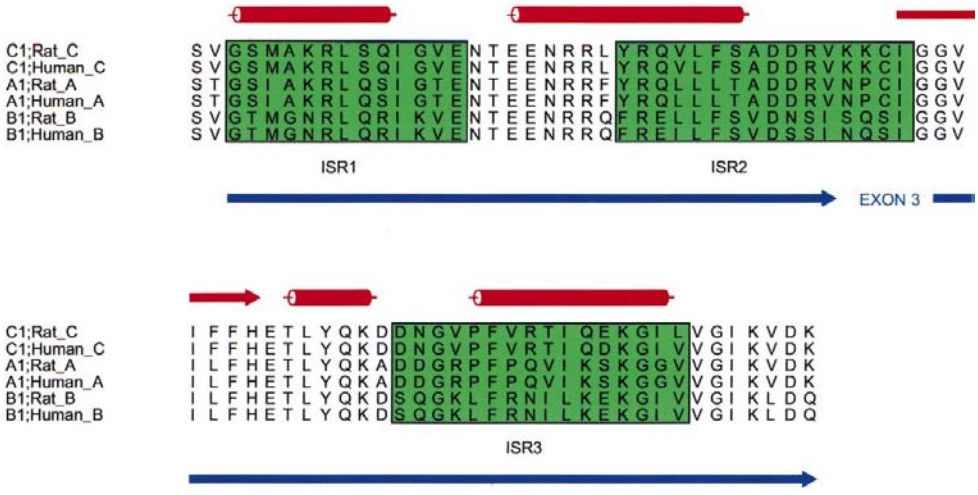

Figure 1
Sequence alignment of exon 3 from human and rat fructose-1,6-bisphosphate aldolases. This alignment indicates the positions of the first three isozyme specific regions within mammalian aldolase. Secondary-structural elements have been assigned from the human muscle aldolase crystal structure. This figure was prepared using the program *ALSCRIPT* (Barton, 1993).

2. Materials and methods

2.1. Crystallization and data collection

Human liver aldolase was overexpressed in *E. coli* and purified as described previously (Beernink & Tolan, 1992; Doyle & Tolan, 1995; Dalby *et al.*, 1996). Crystals were grown using the

Table 1
Summary of data-collection and model-refinement statistics.

Resolution range† (Å)	20–2.5 (2.64–2.5)
Completeness† (%)	70.6 (55.8)
R_{sym}†‡	0.122 (0.438)
$\langle I\rangle/\sigma(I)$†	7.4 (1.4)
$I > 3\sigma(I)$ (%)	65.0
Resolution where $\langle I\rangle/\sigma(I) = 2.5$ (Å)	2.82
Redundancy†	2.4 (2.2)
Unique reflections	339612
B factor of data from Wilson plot (Å^2)	42.6
Final R factor†§ (%)	22.4 (26.8)
Free R factor (5.0% total data)† (%)	27.6 (34.3)
No. of protein atoms	48042
Average B factor, protein (Å^2)	18.4
No. of water molecules	1169
Average B factor, waters (Å^2)	25.7
R.m.s. deviations from ideality (target values are given in parentheses)	
Bond lengths (Å)	0.022 (0.010)
Bond angles (Å)	0.063 (0.020)
1–4 neighbours (Å)	0.082 (0.030)
Planar groups (Å)	0.021 (0.008)
Torsion angles (°)	
Planar	5.6 (6.0)
Staggered	22.5 (10.0)
Orthonormal	35.2 (20.0)
B-factor correlation (Å^2)	
Main-chain bond	5.7 (4.0)
Main-chain angle	7.3 (7.0)
Side-chain bond	11.2 (9.0)
Side-chain angle	12.6 (9.0)

† Values in parentheses are for the outer resolution shell. ‡ $R_{sym} = \sum_h \sum_j |\langle I_h\rangle - I_j(h)|/\sum_h \sum_j I(h)$, where $I(h)$ is the intensity of reflection h. $\sum_h$ is the sum over all reflections and $\sum_j$ is the sum over j measurements of the reflection. § $R_{cryst} = \sum \big||F_O| - |F_c|\big|/\sum |F_o|$.

hanging-drop method from conditions similar to those previously reported (Dalby *et al.*, 1996). Droplets containing protein at a concentration of 15 mg ml^{-1}, 20%(*v*/*v*) saturated $(NH_4)_2SO_4$, 30 m*M* imidazole buffer pH 7.0, 30 m*M* KCl, 3 m*M* magnesium acetate, 0.1 m*M* EDTA were equilibrated over wells containing 1 ml of 35% saturated $(NH_4)_2SO_4$. Crystals were harvested into mother liquor containing 45% saturated $(NH_4)_2SO_4$, 30 m*M* imidazole buffer pH 7.0, 30 m*M* KCl, 3 m*M* magnesium acetate and 0.1 m*M* EDTA. Crystal growth was not reproducible and resulted in only three crystals suitable for X-ray analysis. These crystals were irregular shards the largest of which had dimensions 0.5 × 0.5 × 0.2 mm.

Data were collected at 120 K at the SRS, Daresbury on station 9.5 at wavelength 0.84 Å using an ADSC Quantum4 CCD detector from a single crystal in a cryobuffer containing 35%(*v*/*v*) glycerol added to the mother liquor. The completeness of the data was lower than usual from synchrotron data collections because loss of cooling resulted in the destruction of the crystal during collection. Data collected from the other aldolase crystals was of poor quality with diffuse scattering and could not be merged with that from the best crystal.

The crystals belong to space group *C*2, with unit-cell parameters a = 291.1, b = 489.8, c = 103.4 Å, α = 90, β = 103.6, γ = 90°. The asymmetric unit contains 18 subunits corresponding to four and a half enzymatic tetramers, giving a solvent content of 79%. Data were processed using *DPS*, *MOSFLM* and *SCALA* (Rossmann & van Beek, 1999; Leslie, 1999; Collaborative Computational Project, Number 4, 1994). The data statistics are given in Table 1. These indicate weak diffraction resulting from the large unit-cell parameters and the short exposure times required to collect a complete data set in a given time window with small angles of rotation. However, the data were good enough to solve the structure.

2.2. Molecular replacement

Molecular replacement was performed using a tetramer of the human muscle aldolase as the search model using the program *AMoRe* (Navaza, 1994). An all-atom model of the tetramer was used including the flexible C-terminal region. The final solution contained five tetramers in the unit cell, corresponding to the ten highest solutions of the rotation function.

The last tetramer found had a contact distance to its symmetry-related molecules of 0 Å from packing analysis. This molecular twofold axis lies exactly on the twofold crystallographic symmetry axis and so only two of the subunits lie within the asymmetric unit.

Initial phases were calculated from the molecular-replacement solution using *SIGMAA* and 18-fold NCS averaging was performed using *DM* (Read, 1986; Cowtan, 1994). The resulting electron density was examined to verify the solutions. Electron density was clear and interpretable and unbroken except for ISR2 and ISR3 and the C-terminus. Residue changes could often be seen, as could small displacements of surface-loop regions. Rigid-body refinement using *REFMAC* was performed to improve the molecular-replacement solutions, which showed small overall displacements within the averaged density, and averaging was repeated on the improved models (Murshudov *et al.*, 1997).

2.3. Model building and refinement

Most of the molecule could be rebuilt from the averaged maps using the program *O* (Jones *et al.*, 1991). In molecular replacement model bias is a factor and so this initial map was used to rebuild as much of the model as possible including all of the amino-acid changes from the search-model sequence. After the first rebuilding cycle new averaged maps were generated which showed improved density in the ISRs, where there was a significant displacement of and change in secondary structure between the muscle and liver enzymes (Fig. 2). This allowed the complete molecule to be rebuilt except for the flexible C-terminal region, where density was still poor. Refinement was performed with strict NCS restraints for the core of the molecule using *REFMAC*; the C-terminus was left without NCS restraints (Murshudov *et al.*, 1997). As refinement was restrained and not constrained all of the subunits were rebuilt individually. Solvent molecules were added using *ARP* and 15 sulfate ions included where the electron density suggested them (Lamzin & Wilson, 1993).

3. Results and discussion

3.1. Quality of the model

The crystal structure of human liver aldolase was solved using molecular replacement in a previously unreported *C*2 crystal form. The asymmetric unit consisted of 18 subunits: four catalytic tetramers and half of a fifth tetramer centred on the crystallographic twofold axis. Final refinement of the structure resulted in an *R* factor of 22.4% and a free *R* factor of 27.6% to 2.5 Å resolution (5% of the data were used as a test set). Statistics for data collection and refinement are given in Table 1. These show the results of a difficult data collection with low completeness. The weak data and the high proportion of partial observations resulting from the small oscillation range required for this large unit cell gave a higher $R_{\rm sym}$ than is usual and the data set as a whole can be considered as poor.

The last three residues were missing from all 18 of the subunits. The preceding flexible C-terminal region was visible in some of the subunits but could not be built into all of them. For this reason, the subunits were each rebuilt separately and strict NCS restraints were not applied to this region of the model. In some cases the first two or three amino acids at the N-terminus were also omitted. A C^{α} trace of one of the subunits is shown in Fig. 3. Stereochemical analysis of the model showed that 4785 amino-acid residues were in the most favoured regions of the Ramachandran plot, 672 in additional allowed regions, 60 in generously allowed regions and 41 in disallowed regions (Ramakrishan & Ramachandran, 1965). The flexible C-terminal region and ISRs 2 and 3 were the regions where disallowed amino acids were found to occur (Table 2). Some subunits had very poor density at the N- and C-termini, resulting in these residues being omitted from the final structures.

Table 2
Residues in disallowed regions.

Monomer	Residues
A	His2
B	Ser68, Ile69
C	None
D	His2, Gln89, Ser355
E	His2, Ser67, Thr345, Ser352, Ser355
F	His2, Arg3, Thr345, Ser347, Ser352, Ser355
G	None
H	Ser67, Ser347, Ser355
I	Ser67, Ile69
J	Ser348, Ser355
K	His2, Arg3, Ser348, Thr358, Ala359
L	Arg3
M	Arg3, Thr353
N	None
O	None
P	His2, Ser355
Q	Arg3, Gln89, Thr345, Ala350, Ser355
R	Arg3, Ser355

3.2. The flexible C-terminus

In the first crystal structure of aldolase to be determined, that of rabbit muscle, the C-terminus was omitted because there was no clear electron density (Sygusch *et al.*, 1987). Subsequently, structures from human muscle and *D. melanogaster* and a higher resolution structure from rabbit muscle have all determined the position of this flexible C-terminus (Hester *et al.*, 1991; Blom & Sygusch, 1997). The positions have been shown to be different in the three crystal structures (Blom & Sygusch, 1997). This region begins at residue 344 and from this point the structures give different orientations for the terminus, but all of them are loop regions without any secondary structure and all of them drift towards or over the active-site cleft.

In the human B aldolase structure there is an extra short α-helix from residue 349 to residue 354 in three of the subunits (Figs. 4*a* and 4*b*). Two of these subunits occur within one of the tetramers as is clearly seen in Fig. 4*b*. The C-terminal helix is not visible in the other subunits because of the inherent flexibility of this region, which results in poor electron density. The observation of different structural states within this tetramer is consistent with previous biochemical results that have examined aldolase flexibility through spectroscopy and have also suggested that at low temperature the aldolase active site is half occupied (Kochman & Dobryszycki, 1991; Grazi & Trombetta, 1984). The main part of the α/β-barrel is well determined except for two surface helices whose flexibility seems to be related to that of the C-terminus. So if there are two states for the enzyme, one binding and one unavailable for binding, it is possible that it is a result of the motions of the C-terminus.

The additional helix lies to one side of the active-site cleft and is aligned to the helices surrounding the central barrel with the C-terminus projecting over the active-site cleft (Fig. 4*a*). As in the case of the muscle isozyme, the interactions between the C-terminus and the main α/β-barrel are *via* a conserved hydrophobic residue at position 356 (that is, conserved in all of the aldolase isozymes). It is possible that the last three amino acids, which are not visible in the density, project even further towards the active site. Although the C-terminus has been reported as having some catalytic function, owing to its orientation and flexibility, that function cannot be defined at present (Dreschler *et al.*, 1959). A complete understanding of its role in aldolase activity requires further study. In the uncomplexed structure described here it is not in close enough proximity to be directly involved in catalysis, but it may become more tightly associated with the active site on substrate binding or Schiff-base formation.

3.3. Sulfate-binding sites

15 well defined sulfate molecules were built into the final model of human liver aldolase. The crystals were grown from ammonium sulfate and it is not unusual for the sulfates from crystallization to associate with the phosphate-binding sites of enzymes because of electrostatic interactions. In this case, eight of the sulfates are found in the active-site clefts, but the other sulfate ions are found at the extensive dimer interfaces

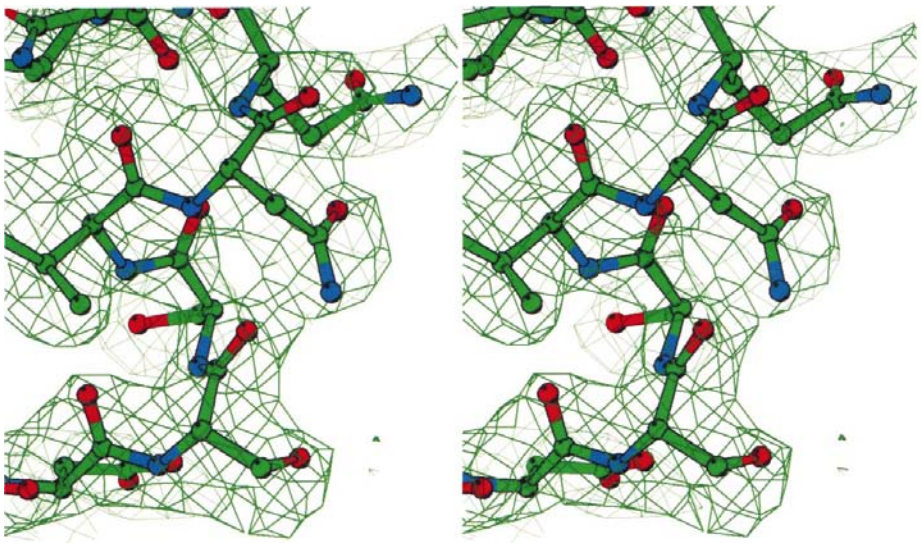

Figure 2
A stereo diagram of the final electron density for ISR2. The $2F_o - F_c$ electron density is shown contoured at 1σ for the *N* subunit. This figure was prepared using the program *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997).

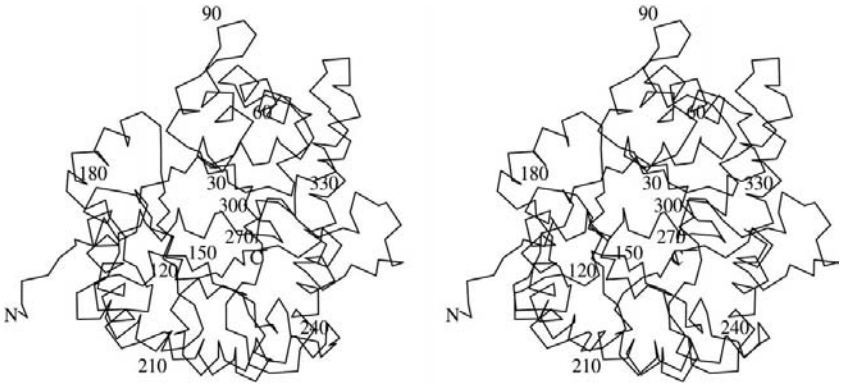

Figure 3
A stereoview of the C^{α} trace for a subunit of human liver aldolase. Every 30th residue is numbered. This figure was prepared using the program *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997).

as defined by Gamblin *et al.* (1990). This dimer interface is made up of helical interactions, unlike the minor dimer interface which consists of a β-hairpin from each of the subunits (Fig. 4*b*). The binding environment for the sulfates at the dimer interface is made up of amino-acid side chains from both subunits and because of the local 222 symmetry of the aldolase homotetramer there will be four of these binding sites associated with each tetramer. This signifies only partial occupancy of the potential sulfate-binding sites, as there are 18 active sites and 18 dimer interface sites that could be occupied.

The sulfate sites at the dimer interfaces lie each on a surface cleft between the two subunits with a separation of 33 Å. These binding sites are composed of four amino-acid residues that are conserved throughout the three aldolase isozymes. Lys12 in the N-terminal helix and Gln202 from helix *E* in one subunit and The254 and Arg258 from helix *F* in the second chain form hydrogen bonds with the sulfate O atoms. This environment would bind any anionic substrate, including the enzyme's natural substrate, so it is not clear why the enzyme should possess an alternative binding site so far from the active site. It is possible that this is the site for the reported association of nucleotides, rather than the active-site cleft (Sytnik *et al.*, 1991).

The sulfate positions in the active site are not always the same as the 1-phosphate binding site, but have a stronger interaction with the Schiff-base lysine Lys229 and with Arg303. Binding of phosphate to the Schiff-base lysine was also found in the structure of an inactive mutant bound to the natural substrate (Choi *et al.*, 1999). In that case, the Schiff-base lysine was hydrogen bonded to the 1-phosphate group of the substrate and was not in close proximity to the C2 carbonyl group where Schiff-bond formation takes place. The positions of the sulfates in this crystal structure are consistent with the conclusion of that study that initial binding of the substrate is by electrostatic interactions with the positively charged lysines in the active site, including the Schiff-base lysine, and that subsequently the substrate is reorientated within the active site to allow Schiff-base formation and cleavage of the carbon–carbon bond.

3.4. Substrate specificity

The active-site residues of all three aldolase isozymes are well conserved. Four regions were identified as isozyme specific by Kitajima *et al.* (1990). The first is between amino acids 37 and 49, the second and third are described below and the fourth is the flexible C-terminus. None of the amino acids from these regions lies within the active site as determined from the crystal structures of enzyme and natural substrate (Dalby *et al.*, 1996). Although the region 37–49 had been thought to be involved in the binding of 6-phosphate in initial crystallographic studies of human muscle aldolase (Gamblin *et al.*, 1991), recent studies have implicated the region around Arg303 in the binding of the 6-phosphate group (Choi *et al.*, 1999). However, this residue is conserved in both the muscle and liver isozymes and so cannot explain the differences in substrate preference between the two isozymes. One of the amino acids that differs between the liver and muscle isozymes is Lys41. The basic lysine residue in the muscle isozyme is replaced by a neutral polar asparagine in the liver form. In the structure of human muscle aldolase complexed with the natural substrate Lys107 and not Lys41 is involved with 6-phosphate binding, but in the complex struc-

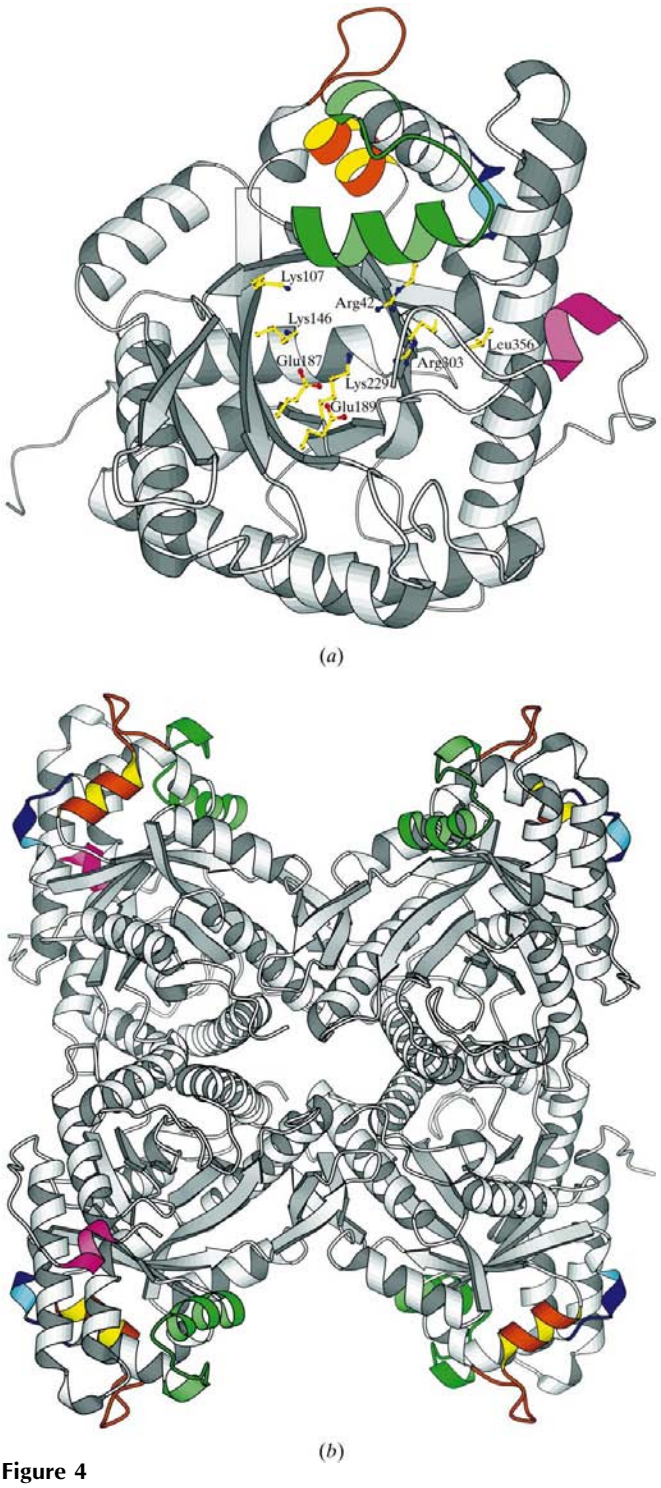

Figure 4
(*a*) A ribbon diagram representation of the aldolase subunit. Amino-acid side chains involved in catalysis are drawn in ball-and-stick representation, including the conserved hydrophobic residue from the C-terminal region (Leu356). The regions where there is a high r.m.s. deviation between the human muscle and human liver isozymes within ISR1, ISR2 and ISR3 are coloured blue, green and orange, respectively. The extra C-terminal α-helix is coloured magenta. The figure was prepared using the program *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997). (*b*) A ribbon diagram of the aldolase tetramer. The colour scheme is the same as that for the subunit. Only two C-terminal helices can be seen in the tetramer as the other C-termini could not be modelled and the amino-acid residues are omitted from the density. This figure was prepared using the program *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997).

ture of the cleavage inactive mutant the initial binding mode of the substrate is associated with Lys41 (Dalby *et al.*, 1999; Choi *et al.*, 1999). The enzymes therefore seem to have two distinct substrate-binding states, one for recognition and a second for catalysis. In the recognition binding state the isozyme specific residues of the human liver aldolase seem to determine the enzyme's substrate specificity, whereas the active-site residues are conserved in all three isozymes and in this binding state the isozymes are not discriminating.

3.5. Isozyme-specific regions

Aldolase varies from the usual eight-stranded α/β-barrel with the insertion of extra helices between strands 1 and 2 and also between strands 2 and 3 which break up the alternating α/β motif (Fig. 4*a*). These additional helices correspond to the regions 35–46, 66–71 and 92–100. These regions align closely to the isozyme-specific regions (ISRs) identified by Kitajima *et al.* (1990) and referred to as IGS1, IGS2 and IGS3 in their study, which correspond to residues 37–49, 58–73 and 88–103 in the human liver sequence (Kitajima *et al.*, 1990). In this study, the regions are referred to as ISR1, ISR2 and ISR3, respectively (Fig. 1). Further mutagenesis studies have provided stronger evidence that ISR2 contains a large block of aldolase B specific residues, Phe58, Glu/Asp60 and Ser68 (Berardini *et al.*, 1999) and that ISR3 contains a block of aldolase A specific residues, Arg91 and Gly102 (Berardini, 1998). The position of these ISRs within the crystal structure can be seen in Figs. 4(*a*) and 4(*b*). The density for each of the 18 subunits rebuilt during the structure determination varies most in quality in these regions and the flexible C-terminal region. Where a subunit had poor density in these regions it also had a correspondingly high *B* factor or amino-acids in the unfavourable regions of the Ramachandran plot, indicating disorder in these regions in the crystal structure. In subunits where the density for ISR2 and ISR3 was poor, the density at the C-terminus for that subunit was also poor and often untraceable for the last 20 amino-acids. Correspondingly subunits with continuous density in ISR2 and ISR3 had traceable C-terminal electron density.

When the structure was rebuilt from the human muscle aldolase, ISR2 required substantial rebuilding of the main chain of the protein, which would not have been expected in such closely related enzymes. This resulted in a 3_{10}-helix in ISR2 of the human liver isozyme. The human muscle aldolase coordinates do not contain a 3_{10}-helix in ISR2 and this could have indicated a source of isozyme specificity. Examination of the other determined structures of fructose-1,6-bisphosphate aldolase from *D. melanogaster* (Hester *et al.*, 1991), *P. falciparum* (Kim *et al.*, 1997) and rabbit muscle (Blom & Sygusch, 1997) showed that the helix is present in all of these structures between residues 66 and 71. The rabbit muscle aldolase has particularly high sequence identity to human muscle aldolase and so this suggests that the loss of the 3_{10}-helix in the human muscle aldolase structure is an artefact. This may have resulted from crystal packing in the tightly packed high crystal symmetry form (Dalby *et al.*, 1999).

Table 3
Summary of the root-mean-square deviations in C^{α} positions between the human liver and human muscle isozymes.

Region	Average r.m.s. (Å)	σ r.m.s. (Å)
High r.m.s. deviation region within ISR1	1.80	0.40
High r.m.s. deviation region within ISR2	1.74	0.79
High r.m.s. deviation region within ISR3	1.58	0.66
C-terminus	2.58	1.98
Complete sequence	0.78	0.80
Complete sequence excluding ISR 1-4 and the C-terminus	0.53	0.35

An analysis of the root-mean-square deviations (r.m.s.) of the C^{α} coordinates of the human liver and human muscle isozymes showed five regions with high r.m.s. deviations (Table 3): the C-terminus, which has well documented flexibility, and also the regions residues 35–54, 63–69, 84–100 and 238–244 (Figs. 4*a* and 4*b*). In all of these regions the r.m.s. deviation was greater than 1 Å and there was a sharp increase in the r.m.s. deviation from neighbouring residues. Three of these regions show a high degree of overlap with the ISRs and the additional helices as defined above. Examination of the secondary-structural elements shows that the α-helix in ISR3 is displaced in the human liver aldolase structure relative to the previously determined muscle aldolase structure. This indicates a rigid-body movement of the secondary structure with an accompanying rearrangement of the rest of the loop so that the displacement does not affect those secondary-structural elements either preceding or following ISR3.

3.6. Fructose intolerance mutations

The most frequent mutation causing hereditary fructose intolerance (HFI) is Ala149Pro, accounting for 57% of cases (Tolan, 1995). This residue is in one of the surrounding α-helices making close contacts with the central β-barrel. When the sequence is mutated in the liver aldolase crystal structure the mutated structure makes bad contacts with Glu187 and Pro188. Both of these residues form part of strand 5. Glu187 is particularly important as it is associated with enzyme catalysis and is one of the two acidic amino-acid side chains lining the β-barrel (Choi, Morris & Tolan, personal communication).

Two other mutations responsible for HFI are also associated with strand 5 (Santamaria *et al.*, 1999; Cross *et al.*, 1990). The deletion of amino acids 182 and 183 that form part of the surface loop will also perturb strand 5 that follows these residues. The mutation Ala174Asp also affects this strand, disrupting interactions with the β-barrel and creating bad contacts with Pro184. The concentration of these mutations around strand 5 and the catalytic significance of the amino acids explains the reduced activity of these enzyme mutants, but also suggests that they do not disrupt the overall enzyme structure significantly as had been previously suggested (Rellos *et al.*, 2000).

Of the other mutations associated with HFI, Arg303Trp is the only one of the point mutations that directly affects the active site of the enzyme. Arg303 forms part of the phosphate-binding site of the enzyme and is also associated with positioning the flexible C-terminus (Dalby *et al.*, 1999; Choi *et al.*, 1999). The substitution of tryptophan alters the charge properties of the phosphate-binding site significantly, although the increase in residue size can be accommodated as the residue lies on the enzyme surface. The last well characterized point mutation is Asn334Lys. This region of the C-terminus had been associated with isozyme specificity. This mutation results in the weakening of helix–helix interactions with neighbouring helices owing to the loss of a hydrogen bond and the presence of bad contacts with residues 279 and 280. This may result in a loss of structural integrity, thereby reducing enzyme function.

3.7. Oligomeric state

Recent studies have suggested that HFI mutations in aldolase can affect the oligomeric state of the enzyme (Rellos *et al.*, 2000). Previously, Tolan and coworkers had shown that active monomers of a double mutant of aldolase A can be produced (Beernink & Tolan, 1996). The crystal structure of *P. falciparum* aldolase also suggested that there were significant differences in the extent of buried surface area on formation of the tetramer (Kim *et al.*, 1997). This buried area was divided into components for each dimer within the tetramer and in *P. falciparum* aldolase the extensive dimer interface buried much less surface area than in human muscle (1600 $Å^2$ compared with 2800 $Å^2$). In the current structure, 1800 $Å^2$ is buried at this dimer interface. This is much closer to the figure for the *P. falciparum* aldolase than that for the human muscle. This difference in the malarial enzyme had been attributed to the deletion of Leu291 in the malarial aldolase sequence (Kim *et al.*, 1997). This results in a 6 Å displacement of the surrounding amino acids when compared with the *D. melanogaster* and human muscle structures. The similar result found here for human liver aldolase, which does not have this deletion, suggests that the associations at this interface cannot be solely attributed to the loop region containing Leu291.

4. Conclusions

The determination of the crystal structure of a second aldolase isozyme has allowed the direct comparison of structural differences between the muscle and liver enzymes that can be specifically correlated with the ISRs. The structure also allows a more precise definition of the ISRs which previously were only loosely defined by the chimeric enzyme studies (Kitajima *et al.*, 1990). This study relates sequence variation to structural variation and shows how regions of a protein structure can evolve while maintaining the overall stability of the protein fold (Chothia & Lesk, 1986). This has allowed a deeper understanding of the structural and functional basis of isozyme specificity.

The liver aldolase structure also has implications for studying the disease hereditary fructose intolerance. This structure provides an insight into how several selective point mutations in this enzyme can be responsible for the cause of this human disease. Crystallographic studies of the mutants involved in the disease have been initiated and it is hoped that this will lead to an understanding of the disease at a molecular level.

The authors would like to thank Emma McGhie and Sharon Doyle for their assistance in purifying the enzyme and also Miroslav Papiz and Michail Isupov for help with the data collection. This work was supported by a BBSRC postdoctoral fellowship to ARD.

References

Barton, G. J. (1993). *Protein Eng.* **6**, 37–40.

Beernink, P. T. & Tolan, D. R. (1992). *Protein Expr. Purif.* **3**, 332–336.

Beernink, P. T. & Tolan, D. R. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 5374–5379.

Berardini, T. Z. R. (1998). Thesis. *Evolution of the Vertebrate Aldolase Isozymes and the Aldolase Gene Family.* Boston University, Boston, MA, USA.

Berardini, T., Amsden, A. B., Penhoet, E. E. & Tolan, D. R. (1999). *Comput. Biochem. Physiol.* **122**, 53–61.

Berardini, T. Z., Drygas, W. M., Callard, G. V. & Tolan, D. R. (1997). *Comput. Biochem. Physiol* **117**, 471–476.

Berthiaume, L., Loisel, T. P. & Sygusch, J. (1991). *J. Biol. Chem.* **266**, 17099–17105.

Berthiaume, L., Tolan, D. R. & Sygusch, J. (1993). *J. Biol. Chem.* **268**, 10826–10835.

Blom, N. & Sygusch, J. (1997). *Nature Struct. Biol.* **4**, 36–39.

Brooks, C. C. & Tolan, D. R. (1993). *Am. J. Hum. Genet.* **52**, 835–840.

Brooks, C. C. & Tolan, D. R. (1994). *FASEB J.* **8**, 107–113.

Choi, K. H., Mazurkie, A. S., Morris, A. J., Ultheza, D., Tolan, D. R. & Allen, K. N. (1999). *Biochemistry*, **38**, 12655–12664.

Chothia, C. & Lesk, A. M. (1986). *EMBO J.* **5**, 823–826.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K. D. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

Cross, N. C. & Cox, T. M. (1989). *Quart. J. Med.* **73**, 1015–1020.

Cross, N. C., de Franchis, R., Sebastio, G., Dazzo, C., Tolan, D. R., Gregori, C., Odievre, M., Vidailhet, M., Romano, V., Mascali, G., Romano, C., Musumeci, S., Steinmann, B., Gitzelmann, R. & Cox, T. M. (1990). *Lancet*, **335**, 306–309.

Cross, N. C. P., Tolan, D. R. & Cox, T. M. (1988). *Cell*, **53**, 881–885.

Dalby, A., Dauter, Z. & Littlechild, J. A. (1999). *Protein Sci.* **8**, 291–297.

Dalby, A., Rawas, A., Watson, H. C. & Littlechild, J. A. (1996). *Protein Pept. Lett.* **3**, 207–212.

Doyle, S. A. & Tolan, D. R. (1995). *Biochem. Biophys. Res. Commun.* **206**, 902–908.

Dreschler, E. R., Boyer, P. D. & Kowalsky, A. G. (1959). *J. Biol. Chem.* **234**, 2627–2634.

Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132.

Gamblin, S. J., Cooper, B., Millar, J. R., Davies, G. J., Littlechild, J. A. & Watson, H. C. (1990). *FEBS Lett.* **262**, 282–286.

Gamblin, S. J., Davies, G. J., Grimes, J. M., Jackson, R. M., Littlechild, J. A. & Watson, H. C. (1991). *J. Mol. Biol.* **219**, 573–576.

Gitzelmann, R., Steinmann, B. & Van den Berghe, G. (1995). *The Metabolic and Molecular Basis of Inherited Disease*, Vol. 1, edited by C. Scriver, A. Beaudet, W. Sly & D. Valle, pp. 905–934. New York: McGraw–Hill Inc.

Grazi, E & Trombetta, G. (1984). *Arch. Biochem. Biophys.* **233**, 595–602.

Hartman, F. C. & Brown, J. P. (1976). *J. Biol. Chem.* **251**, 3057–3062.

Hester, G., Brenner-Holzach, O., Rossi, F. A., Struck, D. M., Winterhalter, K. H., Smit, J. D. & Piontek, K. (1991). *FEBS Lett.* **292**, 237–242.

Jones, T. A., Zou, J.-Y., Cowan, S. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kim, H., Certa, U., Dobeli, H., Jakob, P. & Hol, W. G. (1997). *Biochemistry*, **37**, 4388–4396.

Kitajima, Y., Takasaki, Y., Takahashi, I. & Hori, K. (1990). *J. Biol. Chem.* **265**, 17493–17498.

Kochman, M. & Dobryszycki, P. (1991). *Acta Biochim. Pol.* **38**, 407–421.

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

Kusakabe, T., Motoki, K., Sugimoto, Y., Takasaki, Y. & Hori, K. (1994). *Protein Eng.* **7**, 1387–1393.

Laméire, N., Mussche, M., Baele, G., Kint, J. & Ringoir, S. (1978). *Am. J. Med.* **65**, 416–423.

Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.

Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.

Motoki, K., Kitajima, Y. & Hori, K. (1993). *J. Biol. Chem.* **268**, 1677–1683.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

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

Penhoet, E., Rajkumar, T. & Rutter, W. J. (1966). *Proc. Natl Acad. Sci. USA*, **56**, 1275–1282.

Ramakrishan, C. & Ramachandran, G. N. (1965). *Biophys. J.* **5**, 909–933.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Rellos, P., Ali, M., Vidailhet, M., Sygusch, J. & Cox, T. M. (1999). *Biochem. J.* **340**, 321–327.

Rellos, P., Sygusch, J. & Cox, T. M. (2000). *J. Biol. Chem.* **275**, 1145–1151.

Rossmann, M. G. & van Beek, C. G. (1999). *Acta Cryst.* D**55**, 1631–1653.

Rottmann, W. H., Deselms, K. R., Niclas, J., Camerato, T., Holman, P. S., Green, C. J. & Tolan, D. R. (1987). *Biochimie*, **69**, 137–145.

Santamaria, R., Vitagliano, L., Tamasi, S., Izzo, P., Zancan, L., Zagari, A. & Salvatore, F. (1999). *Eur. J. Hum. Genet.* **7**, 409–414.

Sygusch, J., Beaudry, D. & Allaire, M. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 7846–7850.

Sytnik, A. I., Chumachenko, Y. V. & Demchenko, A. P. (1991). *Biochim. Biophys. Acta*, **1079**, 123–127.

Takahashi, I., Takasaki, Y. & Hori, K. (1989). *J. Biochem. (Tokyo)*, **105**, 281–286.

Takasaki, Y. & Hori, K. (1992). *Protein Eng.* **5**, 101–104.

Tolan, D. R. (1995). *Human Mutation*, **6**, 210–218.

Tolan, D. R. & Brooks, C. C. (1992). *Biochem. Mol. Med.* **48**, 19–25.