

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/13630136>

# Crystal structure of *Saccharomyces cerevisiae* cytosolic aspartate aminotransferase

ARTICLE *in* PROTEIN SCIENCE · JUNE 1998

Impact Factor: 2.85 · DOI: 10.1002/pro.5560070614 · Source: PubMed

---

CITATIONS

21

---

READS

20

## 5 AUTHORS, INCLUDING:



Constance Jeffery

University of Illinois at Chicago

48 PUBLICATIONS 2,036 CITATIONS

[SEE PROFILE](#)



Sharon Doonan

University of the West of Scotland

101 PUBLICATIONS 1,351 CITATIONS

[SEE PROFILE](#)



Gregory A. Petsko

Brandeis University

481 PUBLICATIONS 25,282 CITATIONS

[SEE PROFILE](#)



Dagmar Ringe

Brandeis University

270 PUBLICATIONS 13,350 CITATIONS

[SEE PROFILE](#)

# Crystal structure of *Saccharomyces cerevisiae* cytosolic aspartate aminotransferase

CONSTANCE J. JEFFERY,<sup>1,4</sup> TOM BARRY,<sup>2</sup> SHAWN DOONAN,<sup>2,3</sup>  
GREGORY A. PETSKO,<sup>1</sup> AND DAGMAR RINGE<sup>1</sup>

<sup>1</sup>Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254–9110

<sup>2</sup>Department of Biochemistry, University College Cork, Lee Maltings, Prospect Row, Cork, Ireland

(RECEIVED September 23, 1997; ACCEPTED February 24, 1998)

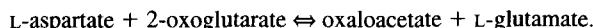
## Abstract

The crystal structure of *Saccharomyces cerevisiae* cytoplasmic aspartate aminotransferase (EC 2.6.1.1) has been determined to 2.05 Å resolution in the presence of the cofactor pyridoxal-5'-phosphate and the competitive inhibitor maleate. The structure was solved by the method of molecular replacement. The final value of the crystallographic *R*-factor after refinement was 23.1% with good geometry of the final model. The yeast cytoplasmic enzyme is a homodimer with two identical active sites containing residues from each subunit. It is found in the “closed” conformation with a bound maleate inhibitor in each active site. It shares the same three-dimensional fold and active site residues as the aspartate aminotransferases from *Escherichia coli*, chicken cytoplasm, and chicken mitochondria, although it shares less than 50% sequence identity with any of them. The availability of four similar enzyme structures from distant regions of the evolutionary tree provides a measure of tolerated changes that can arise during millions of years of evolution.

**Keywords:** aspartate aminotransferase; crystal structure; maleate; pyridoxal phosphate

Given that analyses of whole genome sequences depend on sequence similarity for assignment of biological function to open reading frames, structural studies of the same protein from many distantly related proteins are extremely important. Such studies help to define the level of similarity that can be tolerated by various folds with retention of function. They also serve to define the kinds of tertiary and secondary structural changes that occur at low levels of sequence identity.

L-Aspartate aminotransferase (L-Asp AT, EC 2.6.1.1) is an excellent candidate for comparative studies of sequence and structure conservation. It is a key metabolic enzyme that links amino acid metabolism to carbohydrate metabolism through catalysis of the reversible transamination reaction.



(1)

It also plays a role in ureogenesis, and is involved in transferring reducing equivalents into the mitochondria via the malate/aspartate shuttle. These multiple important roles have resulted in L-Asp AT being conserved throughout evolution. In addition to there being homologous forms of L-Asp AT in different species, the enzyme is found in homologous cytoplasmic and mitochondrial forms in eukaryotes, as well as in additional homologous forms in plant plastids. From the over 25 available sequences, there are forms of L-Asp AT that share as much as 80% sequence identity and as low as 20% sequence identity (Winefield et al., 1995).

L-Asp AT is a pyridoxal-5'-phosphate (PLP) dependent enzyme with a complex reaction mechanism (Kirsch et al., 1984; Arnone et al., 1985a). During catalysis, the enzyme goes through several intermediate steps to alternate between a PLP form, which contains the cofactor covalently linked via a Schiff base linkage to the epsilon amino group of a lysine residue, and a pyridoxamine form of the enzyme. A detailed understanding of the complex reaction mechanism, catalytic residues, and role of the PLP cofactor in this enzyme has been greatly aided by comparison of the structure and function of aspartate aminotransferases from several organisms. Crystal structures are available for *Escherichia coli* L-Asp-AT (Smith et al., 1989; Kamitori et al., 1990; Jager et al., 1994; Okamoto et al., 1994), the mitochondrial and cytosolic isozymes from chicken (Ford et al., 1980; Borisov et al., 1985; Harutyunyan et al., 1985; McPhalen et al., 1992; Malashkevich et al., 1995) and the cytosolic isozyme from pig (Arnone et al., 1985b). These proteins share a

Reprint requests to: Dagmar Ringe, Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254–9110; e-mail: ringe@binah.cc.brandeis.edu.

Abbreviations: PLP, pyridoxal-5'-phosphate; PMSF, phenylmethyl sulfonyl fluoride; PEG 4,000, polyethylene glycol with average molecular weight of 4,000; L-Asp-AT, L-aspartate aminotransferase.

<sup>3</sup>Current address: Department of Life Sciences, University of East London, Romford Road, Stratford, London E15 4LZ, England.

<sup>4</sup>Current address: Department of Physiology, Tufts University Medical School, Boston, Massachusetts 02111.

similar three-dimensional fold although they have only about 40% amino acid sequence identity (except for the two cytosolic isozymes, which share 83% sequence identity). A comparison of these structures with the L-Asp-AT structure from a unicellular eukaryote would provide more information about the evolution of this family of proteins and add to our understanding of how different sequences can fold into similar structures. The cytosolic L-Asp-AT isozyme from the yeast *Saccharomyces cerevisiae* has been purified and sequenced (Cronin et al., 1991). Sequence comparisons indicate it shares less than 50% amino acid identity with any of the four L-Asp aminotransferases for which structures are available (Table 1). We report here the high resolution crystal structure of the *S. cerevisiae* cytoplasmic aspartate aminotransferase.

## Results

### Structure determination

Crystals of *S. cerevisiae* L-Asp aminotransferase were obtained in the presence of PLP and maleate and diffracted to 2.05 Å resolution with unit cell parameters of  $a = 130.31$  Å,  $b = 134.63$  Å, and  $c = 98.75$  Å,  $\alpha = \beta = \gamma = 90^\circ$ , as described previously (Jeffery et al., 1998). A space group of P2<sub>1</sub>2<sub>1</sub>2 or P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> was determined from the presence of systematic absences in the diffraction data set. The predicted subunit molecular mass of 45,287 Da gives a Matthews constant of 2.4 Å<sup>3</sup>/Da, which is consistent with four subunits per asymmetric unit (Matthews, 1968).

Molecular replacement using the AMoRe program package (Navaza, 1994) was successful in determining the locations of two dimers in the asymmetric unit. Each subunit contains 412 amino acids (1,648 amino acid residues per asymmetric unit), a pyridoxal phosphate cofactor covalently attached to lysine 258, a maleate inhibitor bound in the active site, and approximately 200 water molecules (overall 715 water molecules). Statistics describing the refined protein model are given in Table 2, and residue conformations are indicated in the phi-psi plot in Figure 1. These results indicate good geometry for the final model (Morris et al., 1992). There is one nonglycine residue that is found in a disallowed region in all four subunits, but it has clear electron density. This residue, Ser296, is also found to be in a disallowed region of the phi-psi plot in the L-aspartate aminotransferase structures from other species. Its side chain is able to form hydrogen bonds with the side chain of Asn297 and the side chain of Arg292. Arg292 in turn forms hydrogen bonds with the carbonyl oxygens of the maleate inhibitor. It appears that the strained conformation of Ser296 has been conserved because this series of hydrogen bonds holds Arg292, a key active site residue, in position to interact with substrate.

**Table 1.** Pairwise % amino acid sequence identity

	Yeast cyt.	<i>E. coli</i>	Chick cyt.	Chick mit.
Yeast cyt.	100	38.6	47.4	42.9
<i>E. coli</i>		100	39.5	41.3
Chick cyt.			100	45.9
Chick mit.				100

**Table 2.** Data collection and refinement statistics

	Data collection	Refinement
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	
Cell dimensions	$a = 130.31$ Å $b = 134.63$ Å $c = 98.75$ Å	
Temperature	4 °C	
Resolution limit	2.05 Å	
Observed reflections		
Total	470,081	
Unique	100,347	
Completeness		
30.0–2.05	93%	
2.14–2.05	88%	
<i>R</i> -merge <sup>a</sup>	12%	
Subunits per asymmetric unit	4	
Resolution range	10–2.05 Å	
<i>R</i> -factor <sup>b</sup>	23.1	
<i>R</i> -free <sup>c</sup>	29.9	
In each asymmetric unit, number of		
Protein subunits	4	
PLP cofactors	4	
Maleic acid molecules	4	
Protein atoms	12,844	
Water molecules	715	
Mean <i>B</i> factor		
Main chain	17.8 Å <sup>2</sup>	
Side chains	20.8 Å <sup>2</sup>	
Water molecules	32.2 Å <sup>2</sup>	
RMSDs from ideal geometry		
Bond distances	0.017 Å	
Bond angles	1.8°	
Torsion angles	23.0°	

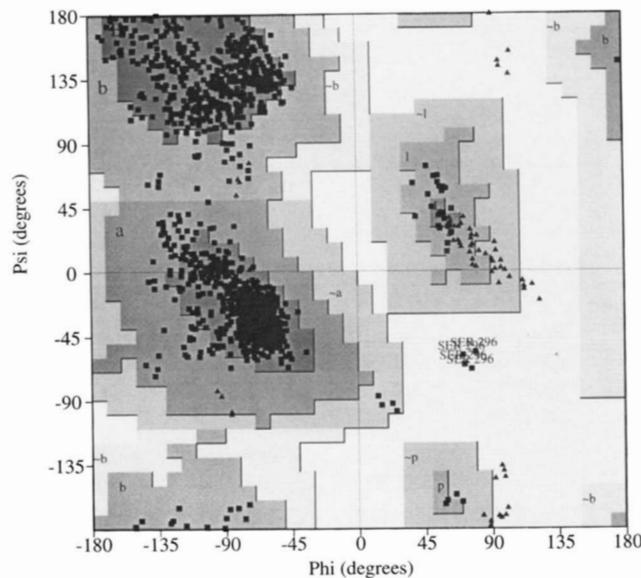
$$^a R\text{-merge} = \frac{\sum |I - \langle I \rangle|}{\sum \langle I \rangle}$$

$$^b R\text{-factor} = \frac{\sum_{hkl} |F_{obs}(hkl) - F_{calc}(hkl)|}{\sum_{hkl} F_{obs}(hkl)}$$

<sup>c</sup> *R*-free as defined by Brunger (1992a), using 10% of the data.

### Adjustment of sequence

When the DNA sequence for the yeast aspartate aminotransferase became available from the Yeast Genome Project (file number AAT2), three areas were found in which the amino acid sequence predicted from the DNA sequence differs from the amino acid sequence previously reported from peptide sequencing (Cronin et al., 1991). Residue 95 was found to be a leucine in peptide sequencing, but is predicted to be a phenylalanine from its DNA sequence. An electron density omit map (Read, 1986; Hodel et al., 1992) indicates that the residue is a leucine. It is possible that the difference in the sequences is due either to a single base mistake in the gene sequencing or a difference between the strains of *S. cerevisiae* used for sequencing and used as the source of protein for the crystal structure. The codon for this residue is reported as TTT, which encodes phenylalanine, but a single base change could result in TTA, TTG, or CTT, which all encode leucine. Residue 95 is located at the surface of the protein, and its side chain is more than



**Fig. 1.** Phi-psi plot of the final model. Allowed regions are indicated by grey shading. Triangles represent glycine residues. Squares represent nonglycine residues. As described in the text, Ser296 was found in a disallowed conformation in each of the four subunits. This figure was prepared by PROCHECK (Laskowski et al., 1993).

3 Å from any other side chain, so it is not at a vital area of the protein structure.

The last observable residue in the model (residue 412) was found to be an alanine in the peptide sequencing, but is predicted to be a threonine from its DNA sequence. The difference electron density is consistent with an alanine at that position. Again, there is only a single base change between the gene sequence, ACT, and a possible codon for alanine, GCT. Residue 413 also differed between the peptide sequence and the sequence predicted from the DNA sequence, but there was no observable electron density to confirm either residue, presumably because the C-terminal residues are disordered.

Peptide sequencing also predicted there were at least two unknown residues following residue 241 in the protein sequence. They were labeled "XX" in the SwissProt sequence. The DNA sequence predicts that there are five residues at this position, L242a, G242b, V242c, E424d, and K243. The difference electron density confirms the presence of these five residues. This insertion extends alpha helix 10, which is found between beta strands six and seven in the large domain. It does not make a significant difference in the protein structure compared to the other L-Asp AT structures.

#### Overall structure

The yeast aspartate aminotransferase dimer structure is similar to that of the other aspartate aminotransferases for which structures are known. It is a homodimer with two identical active sites containing residues from each subunit (Fig. 2). Each subunit of the homodimer contains a large and small domain. The large domain is comprised of residues 49 to 325 and consists of a central mixed beta sheet surrounded by 12 alpha helices. The beta sheet contains 43 residues arranged in seven strands. The small domain also has an  $\alpha/\beta$  structure and is made up of the N-terminal residues 16 to



**Fig. 2.** Schematic diagram of the dimer. One subunit is shown in dark grey, and the other subunit is shown in white. The PLP cofactor and the maleate bound in the active site are indicated by ball-and-stick models in black. This figure and Figures 3, 4, 5B, and 5C were prepared with MOLSCRIPT (Kraulis, 1991).

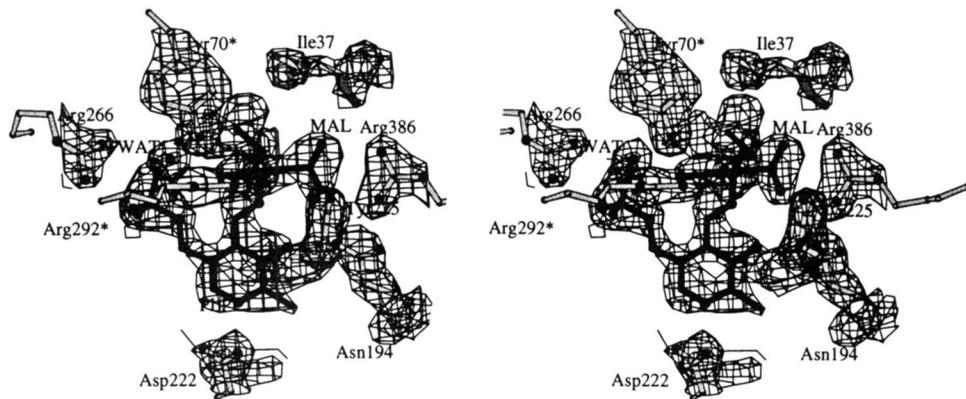
48 and the C-terminal residues 326 to 412. The structure has three short beta strands and five surrounding alpha helices. The N-terminal 14 amino acids form an extended region that wraps across the other subunit of the dimer. The C-terminal 19 residues form an alpha-helix. The protein sequence contains five C-terminal amino acids for which there is no observed electron density presumably due to protein disorder in the crystal. These residues are not included in the final model.

A least-squares minimization overlay of the four subunits found in the asymmetric unit showed that all four subunits have the identical polypeptide chain fold, although they are found in different environments within the crystal. The pairwise root-mean-square deviations (RMSDs) in  $\alpha$ -carbon positions are between 0.3 and 0.4 Å.

The L-Asp aspartate aminotransferases from other species have been shown to adopt two major conformations. Without substrate or inhibitor, the aminotransferases adopt an "open" conformation, in which the small and large domains are distant from each other, that enables the substrates to bind and the products to dissociate. When substrate or inhibitor is bound, the aminotransferases adopt a "closed" conformation, in which the small domain rotates into a position on top of the large domain, closing up the active site pocket and sequestering the substrate from solvent. The yeast cytoplasmic aspartate aminotransferase structure reported here contains a bound maleate inhibitor in each active site, and has a conformation that is similar to the closed forms of the *E. coli* (PDB entry 1amr, with maleate bound), chicken mitochondrial (PDB entry 1ama, with 2-methyl aspartate bound), and chicken cytoplasmic (PDB entry 2cst, with maleate bound) enzymes.

#### Active site

The electron density surrounding the PLP cofactor, maleate inhibitor, and surrounding amino acid side-chain and main-chain atoms was clear and allowed unambiguous positioning of all the groups in the active site (Fig. 3). The active site of the yeast enzyme resembles the active sites of the other available aspartate aminotransferase structures. Many of the residues forming the active site,



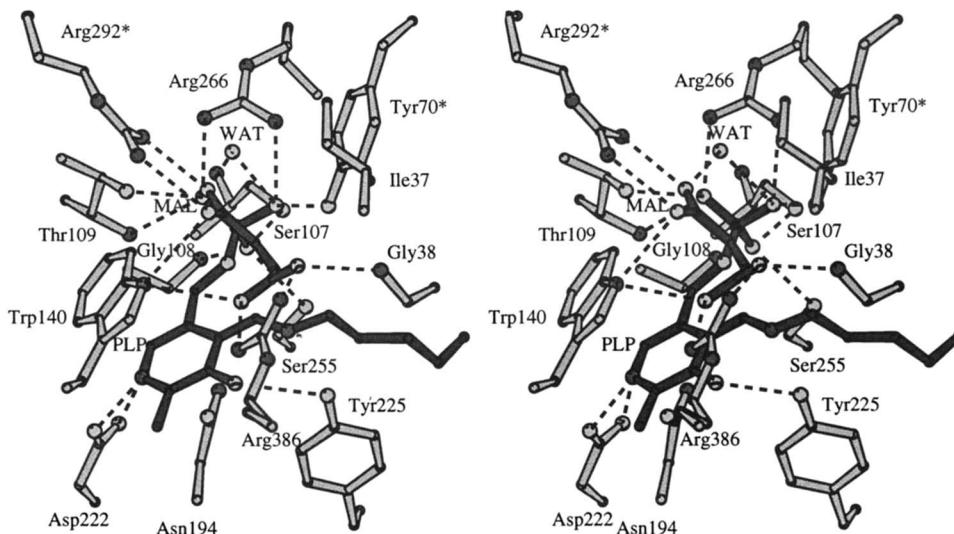
**Fig. 3.** One of the four active sites in the asymmetric unit is shown with an electron density map around the PLP cofactor, the maleate inhibitor, and nearby side chains. The difference electron density map was calculated with coefficients of  $|2f_o - f_c|$  and is drawn with a 1.5 sigma cutoff. The PLP cofactor and the maleate inhibitor are shown with dark grey bonds, and the active site residues are shown with light grey bonds.

especially those side chains forming interactions with the cofactor and inhibitor, are identical in all four proteins. In the yeast cytoplasmic enzyme, the pyridoxal phosphate cofactor forms a Schiff's base with the side chain of Lys258 (Fig. 4). Several other residues are in positions where they can make hydrogen bonds to the cofactor (Table 3). The amide nitrogen of Asn194 and the hydroxyl group of Tyr225 form hydrogen bonds with the 3' oxygen of the pyridine ring. The side-chain carboxylate of Asp222 can form a hydrogen bond or a salt bridge to the pyridine ring nitrogen. In addition, the indole ring side chain of Trp140 is in a position where it can form stacking interactions with the pyridine ring of the cofactor.

The phosphate group of the bound PLP is completely solvated by the protein. The amino acid side chains of residues Ser255, Thr109, Ser107, and Arg266 from one subunit and Tyr70 from the

other subunit form hydrogen bonds to the phosphate group of the cofactor. The backbone amide groups of Gly108 and Thr109 also make hydrogen bonds to the phosphate group. In addition, water548 forms a hydrogen bond to the phosphate group.

Each maleate inhibitor binds with an end-on symmetric geometry, forming hydrogen bonds with the conserved residues Arg386 and Arg292 from the other subunit (Fig. 4). The positions of these two arginines are largely responsible for the L-Asp aminotransferases' specificity for small dicarboxylic acid substrates. The maleate oxygen atoms also form hydrogen bonds with the amide nitrogen of Asn194, the backbone amide group of Gly38, the amide nitrogen of Trp140, and water548 in the active site. This water molecule thus bridges between the carboxylate groups of bound substrates or inhibitors and the phosphate group of the cofactor.



**Fig. 4.** One of the four active sites in the asymmetric unit is shown with possible hydrogen bonds between the PLP cofactor, the maleate inhibitor, and nearby side chains indicated by dashed lines. Hydrogen bonds were calculated with a 3.4 Å cutoff. The PLP cofactor and the maleate inhibitor are shown with dark grey bonds, and the active site residues are shown with light grey bonds.

**Table 3.** Interatomic distances in the active site<sup>a</sup>

N $\zeta$ PLP	O-3' PLP	2.7 (0.1)
N1 PLP	O81 D222	2.8 (0.3)
	O82 D222	2.7 (0.3)
O-3' PLP	N82 N194	2.9 (0.1)
	OH Y225	3.3 (0.5)
OP1 PLP	Oy S255	2.8 (0.0)
	N G108	2.7 (0.2)
	Oy S107	2.6 (0.3)
OP2 PLP	N $\eta$ 1 R266	3.0 (0.2)
	OH Y70	2.8 (0.1)
	O water	3.0 (0.2)
OP3 PLP	N T109	2.9 (0.1)
	N $\eta$ 2 R266	2.8 (0.1)
	Oy T109	2.5 (0.1)
O4B Mal	N $\eta$ 1 R292	3.0 (0.1)
	Nq1 W140	3.1 (0.3)
O4A Mal	N $\eta$ 1 R292	2.9 (0.3)
	N $\eta$ 2 R292	3.6 (0.2)
	O water	2.5 (0.2)
O1A Mal	N $\theta$ 1 W140	3.2 (0.3)
	N82 N194	3.4 (0.6)
	N $\eta$ 1 R386	3.1 (0.2)
O1B Mal	N G38	3.2 (0.3)
	N $\eta$ 2 R386	3.0 (0.2)

<sup>a</sup>Distances shown are the averages for the four active sites in the asymmetric unit. The standard deviations are indicated in parentheses.

#### Comparison to other L-Asp aminotransferases

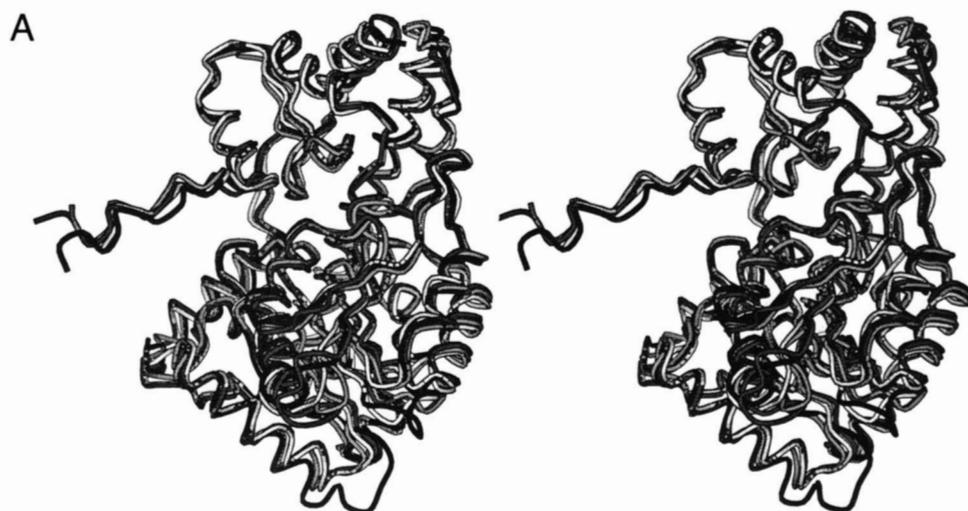
The yeast cytoplasmic aspartate aminotransferase structure has the same fold as the aspartate aminotransferases from *E. coli*, chicken

**Table 4.** RMSD in C $\alpha$  positions<sup>a</sup>

	Yeast cyt.	Chicken cyt.	Chicken mit.	<i>E. coli</i>
Yeast cyt.	—	0.9	0.9	1.3
Chicken cyt.	0.7	—	1.0	1.3
Chicken mit.	0.7	0.7	—	1.3
<i>E. coli</i>	1.4	1.3	1.3	—

<sup>a</sup>The four structures were superposed using a least-squares method in O with either the whole subunit (above the diagonal, using the alpha carbons from residues 5–63, 67–121, 133–152, 155–246, 250–272, and 283–404 of each structure) or the large domain (below the diagonal, using the alpha carbons from residues 49–63, 67–121, 133–152, 155–246, 250–272, and 283–325 of each structure).

cytoplasm, and chicken mitochondria. In order to measure how closely the positions of the secondary structure elements were conserved during evolution, the alpha carbon atoms of one subunit from each of the four structures were superposed using a least-squares algorithm in O (Jones et al., 1991), and the superposed alpha carbon traces are shown in Figure 5A. The superpositions were based on all residues except those found in loops containing insertions or deletions in any of the structures (residues 5–63, 67–121, 133–152, 155–246, 250–272, and 283–404 in each structure). All four structures superposed with pairwise RMSDs of alpha carbon positions of less than 1.4 Å, as shown in Table 4. Whether the superpositioning was based on just the large domain or on both the large and small domains together, the yeast cytoplasmic enzyme structure was more closely related to the chicken cytoplasmic and chicken mitochondrial structures than to the *E. coli* structure. In light of this result, it is not surprising that the chicken cytoplasmic enzyme structure worked well as a search model in molecular



**Fig. 5.** Comparison of the yeast cytoplasmic aspartate aminotransferase with the aspartate aminotransferases from chicken cytoplasm, chicken mitochondria, and *E. coli*. **A:** The alpha carbon tracings of one subunit of the *E. coli* enzyme (white), the chicken mitochondrial enzyme (light grey), and the chicken cytoplasmic enzyme (dark grey) are superposed on the alpha carbon tracing of one subunit of the yeast cytoplasmic enzyme (black). The superpositions were made with a least-squares minimization method using the alpha carbons of residues 5–63, 67–121, 133–152, 155–246, 250–272, and 283–404 of each structure. One area where the enzyme structures differ is in the loop at the bottom of the figure (residues 269e to 280) where there is an insertion of four residues in the yeast enzyme relative to the other three enzymes. **B:** The active site of the yeast cytoplasmic aspartate aminotransferase (black) is superposed on the active site of the chicken cytoplasmic aspartate aminotransferase (white). At each position where the amino acid sequence differs between the two enzymes, the residue for the yeast cytoplasmic enzyme is followed in parentheses by the residue for the chicken cytoplasmic enzyme. **C:** The 13 residues that are conserved in 26 aspartate aminotransferase sequences are shown in ball-and-stick diagrams on one subunit of an alpha-carbon trace of the yeast cytoplasmic L-Asp AT dimer. (Figure continues on facing page.)

replacement even though the two enzymes have less than 50% sequence identity.

To study how this fold was conserved, we can begin by asking where the conserved residues are found. There are 100 residues that are found in all four structures (about 25% of the residues). The conserved residues are scattered throughout the structure. There is no more conservation in the secondary structural elements (helices, strands) than in the protein as a whole, which is consistent with there being many amino acid sequences that are capable of folding into a helix or coil. The interfaces between the small and large domains in one subunit and between the two subunits in the dimer are slightly more conserved, with approximately 30% sequence identity. As might be expected, many of the substitutions are with similar residues.

One place where there is a high degree of strict conservation is in the active site. All of the residues that interact with the pyridoxal phosphate cofactor or the maleate inhibitor are strictly conserved in all four structures, as well as some of the residues that do not interact directly with the cofactor or inhibitor but help form the shape of the active site pocket. Although more than 50% of the

other residues in the protein vary from one species to another, the active site residues are not only the same side chains, but the positions of the side chains are found to superpose closely when the cofactors from two structures are superposed, as shown in Figure 5B.

In a comparison of 26 L-Asp AT sequences that included sequences from bacteria, archaea, and eukaryotic cytoplasm, mitochondria, and plastids (Winefield et al., 1995), 13 residues were found to be conserved throughout evolution. The corresponding yeast residues are shown on one subunit of the yeast enzyme in Figure 5C. Seven of the residues, Tyr70, Asn194, Asp222, Tyr225, Lys258, Arg 266, and Arg386 interact directly with the PLP cofactor or the maleate inhibitor as described above. The side chain of Phe360 holds the side chain of active site residue Arg386 in place through an arginine-aromatic stacking interaction. Gly75, Pro138, Pro195, Gly197, and Gly268 are important in maintaining the aminotransferase fold near key active site residues. Gly75 is located in a turn near active site residue Tyr70. Pro138 is in a turn just two residues from active site residue Trp140. Pro195 and Gly197 help form the turn next to active site residue Asn194. The

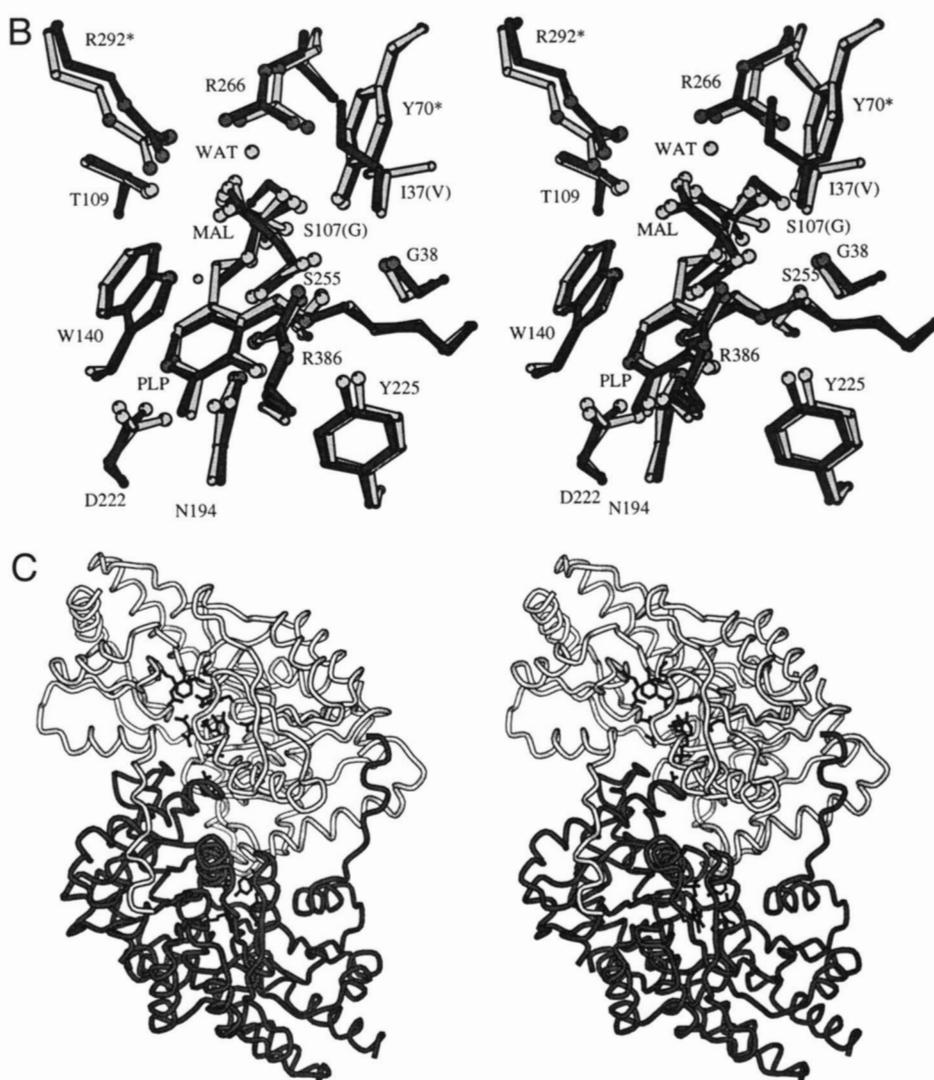


Fig. 5. *Continued.*

final conserved residue, Gly268, is located between a loop containing active site residue Lys258 and a loop that contains residues 107 and 109, which interact with the cofactor as described above. A glycine was likely maintained at this position because the close proximity of these three secondary structural elements does not leave any space for a side chain.

## Discussion

The structure of the *S. cerevisiae* cytoplasmic aspartate aminotransferase has been solved to 2.05 Å resolution. This protein shares the same three-dimensional fold and active site residues as other aspartate aminotransferases for which structures are now available (Ford et al., 1980; Arnone et al., 1985a, 1985b; Borisov et al., 1985; Harutyunyan et al., 1985; Smith et al., 1989; Kamitori et al., 1990; McPhalen et al., 1992; Jager et al., 1994; Okamoto et al., 1994; Malashkevich et al., 1995). Together with the three L-Asp aminotransferase structures available in the Protein Data Bank, we have a set of four enzyme structures that share less than 50% sequence identity but have three-dimensional folds that superpose very well. These four structures represent distant regions of the evolutionary tree. A detailed comparison of their structures provides a measure of tolerated substitutions that can arise during millions of years of evolution.

Our initial comparison described above indicates some of the characteristics of the protein structure that were conserved in order to preserve the three-dimensional fold. One of the driving forces for the conservation of the aspartate aminotransferase fold was the need to position the active site residues carefully. Each active site is made up of residues from both subunits and from both the large and small domains within a subunit. In order to bring all these residues together in the same way in all four aspartate aminotransferases, it appears that the structure of each of these parts of the protein was conserved.

## Methods

### *Protein crystallization and data collection*

Protein crystallization and data collection were described previously (Jeffery et al., 1998). *S. cerevisiae* cytoplasmic aspartate aminotransferase was prepared from wild-type yeast cells, and crystal conditions contained 5–15 mg/mL aspartate aminotransferase, 10 mM PMSF, 10 mM PLP, 10 mM sodium acetate buffer, pH 5.8, and 100 mM maleate in the protein solution and 0.2 M ammonium acetate, 0.1 M sodium acetate buffer, pH 4.6, and 20 to 22% PEG 4,000 in the reservoir solution. The best diffracting crystals were approximately 0.2 mm × 0.2 mm × 0.7 mm in size. The data collection statistics for the data set used in the structure solution are summarized in Table 2.

### *Molecular replacement*

The AMoRe program package (Navaza, 1994) was used to determine molecular replacement rotation and translation solutions for the chicken cytosolic L-Asp aminotransferase (PDB entry 2CST), which shares 46.7% sequence identity with the yeast cytosolic aspartate aminotransferase. Several different search models were tried in the molecular replacement. Some search models included a single subunit, and others included a dimer. In some of the

search models, amino acid side chains that differed between the two models were replaced with alanine, and, in some models, the loops that were likely to be different sizes due to amino acid insertions or deletions were left off. Since the space group was uncertain, all eight primitive orthorhombic standard and nonstandard space groups were tried. Once a set of successful parameters was found for the molecular replacement, it was found that any of the search models that included a dimer of the aminotransferase resulted in similar rotation and translation solutions. The program Transform from the ProtSys program package was used to apply the molecular replacement solutions to an initial model of the yeast enzyme.

### *Construction of initial L-Asp AT model*

An initial model of the yeast enzyme was made based on the structure of the chicken enzyme using the protein design and homology modeling utilities from the Quanta 4.0 Program package (Molecular Simulations, Inc., Burlington, MA). The protein sequences (from PDB entry 2CST and SwissProt entry aat\_yeast) were aligned using the Needleman-Wunsch protocol, and the amino acid residues in the chicken enzyme were replaced by the corresponding residues of the yeast enzyme. Insertions were left out of this initial model and were added later during refinement and manual rebuilding.

### *Refinement and manual rebuilding*

A test set of 10% of the structure factors was flagged for monitoring the free *R* factor (Brünger, 1992a; Kleywegt & Brünger, 1996) during the course of refinement. Rigid body Xplor refinement was used to improve the initial phases (Brünger et al., 1987; Brünger, 1992b; Brünger, 1996). In the first rounds of rigid body refinement, each of the four subunits was treated as a rigid body. Since other L-aspartate aminotransferases have been shown to have multiple conformations, additional rigid body refinement steps were performed in which fragments representing the N-terminal extended region, the small domain, and the large domain, were treated as separate rigid bodies (residues 2 to 89, 96 to 120, 134 to 151, 158 to 264, and 281 to 405).

Rigid body refinement was followed by rounds of positional refinement, *B* factor refinement, and manual rebuilding (Brünger et al., 1987; Brünger et al., 1990; Brünger, 1992b; Collaborative Computational Project, 1994; Brünger, 1996). Noncrystallographic symmetry matrices determined in O (Jones et al., 1991) for the four subunits in the asymmetric unit and averaged maps made using the R/Ave, Mapman and MaMa program packages (Jones, 1992; Kleywegt & Jones, 1993, 1994) were used in the first rounds of positional refinement and rebuilding. When the *R* factor reached 30% and did not continue to decrease, the four subunits were rebuilt individually using difference Fourier maps with coefficients of  $|2f_o - f_c|$  and  $|f_o - f_c|$ . Residues that were left out of the original model were added. In addition, the amino acid sequence determined through peptide sequencing (Cronin et al., 1991) differed from the amino acid sequence predicted from the DNA sequence determined in the Yeast Genome Project (file AAT2). Amino acid substitutions were made in the structure to match the protein sequence predicted from the DNA sequence when their presence was confirmed by the electron density map. Water molecules were added using a 2.0 sigma cutoff in difference Fourier maps with coefficients of  $|f_o - f_c|$ . Simulated annealing omit maps (Read,

1986; Hodel et al., 1992) were used throughout the structure to make final adjustments to the positions of the residues and inhibitors without model bias. After many rounds of Xplor refinement and manual rebuilding, the model fit the density well, and the *R* factors stopped decreasing. The final *R* factor was 23.1% for all data where  $F > 1\sigma$  from 10–2.05 Å. The free *R* factor was 29.9%. The final *R* factors are somewhat higher than in most reported structures. Several reasons why final *R* factors can be high are described in Kleywegt and Brunger (1996). In the case of the yeast cytosolic aspartate aminotransferase structure, the main reason the *R* factors remained high was that the high resolution diffraction data were relatively weak. The *R* factors for the lower resolution data were much lower. For a 10–3.0 Å resolution shell, the *R*<sub>free</sub> was 25.4% and the *R*<sub>working</sub> was 18.2%. For a 3–2.5 Å resolution shell, the *R*<sub>free</sub> was 31.3% and the *R*<sub>working</sub> was 23.4%. For a 2.5–2.05 Å resolution shell, the *R*<sub>free</sub> was 36.6% and the *R*<sub>working</sub> was 30.7%. The final model was found to have good geometry as determined by PROCHECK (Laskowski et al., 1993) and Xplor (Brunger et al., 1987; Brunger, 1992b, 1996) and indicated in Table 2. The final model has been deposited in the Protein Data Bank as entry number 1yaa.

### Acknowledgments

This work was supported by a Cystic Fibrosis Foundation Postdoctoral Fellowship to CJJ and a grant from the National Science Foundation. It was also supported (in part) by a grant from the Lucille P. Markey Trust. We are grateful to Dr. Andrea Hadfield for helpful discussions and Ezra Peisach and Dan Peisach for suggestions about methods to prepare the figures containing protein structures.

### References

- Arnone A, Christen P, Jansonius JN, Metzler DE. 1985a. Mechanism of aminotransferase action. C. Hypothetical mechanism of action of aspartate aminotransferases. In: Christen P, Metzler DE, eds. *Transaminases*. New York: John Wiley & Sons. pp 326–362.
- Arnone A, Rogers PH, Hyde CC, Briley PD, Metzler CM, Metzler DE. 1985b. Pig cytosolic aspartate aminotransferase: The structures of the internal aldimine, external aldimine, and ketimine and of the b subform. In: Christen P, Metzler DE, eds. *Transaminases*. New York: John Wiley & Sons. pp 138–155.
- Borisov VV, Vorisova SN, Kachalova GS, Sosfenov NI, Vainstein BK. 1985. X-ray studies of chicken cytosolic aspartate aminotransferase. In: Christen P, Metzler DE, eds. *Transaminases*. New York: John Wiley & Sons. pp 155–164.
- Brunger AT. 1992a. The free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355: 472–474.
- Brunger AT. 1992b. *XPLOR version 3.1 manual*. New Haven, Connecticut: Yale University Press.
- Brunger AT. 1996. X-plor version 3.851. *X-PLOR (online)*. New Haven, Connecticut: Yale University.
- Brunger AT, Kruckowski A, Erickson J. 1990. Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Cryst A* 46: 585–593.
- Brunger AT, Kuriyan J, Karplus M. 1987. Crystallographic R-factor refinement by molecular dynamics. *Science* 235: 458–460.
- Collaborative Computational Project, Number 4. 1998. The CCP4 Suite: Programs for protein crystallography. *Acta Cryst D* 50: 760–763.
- Cronin VB, Maras B, Barra D, Doonan S. 1991. The amino acid sequence of the aspartate aminotransferase from baker's yeast (*Saccharomyces cerevisiae*). *Biochem J* 277: 335–340.
- Ford GC, Eichele G, Jansonius JN. 1980. Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. *Proc Nat Acad Sci USA* 77: 2559–2563.
- Harutyunyan EG, Malashkevich VN, Kochkina VM, Torchinsky YM. 1985. Three-dimensional structure of the complex of chicken cytosolic aspartate aminotransferase with 2-oxoglutarate. In: Christen P, Metzler DE, eds. *Transaminases*. New York: John Wiley & Sons. pp 164–173.
- Hodel A, Kim S-H, Brunger AT. 1992. Model bias in macromolecular crystal structures. *Acta Cryst A* 48: 851–859.
- Jager J, Moser M, Sauder U, Jansonius JN. 1994. Crystal structures of *Escherichia coli* aspartate aminotransferase in two conformations: Comparison of an unliganded open and two liganded closed forms. *J Mol Biol* 239: 285–305.
- Jeffery CJ, Barry T, Doonan S, Petsko GA, Ringe D. 1998. Crystallization and preliminary X-ray diffraction analysis of aspartate aminotransferase from *Saccharomyces cerevisiae*. *Acta Cryst D*. In press.
- Jones TA. 1992. A set of averaging programs. In: *Molecular replacement (CCP4)*. London, UK: SERC Daresbury Laboratory. pp 92–105.
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for the building of protein models in electron density maps and the locations of errors in these maps. *Acta Crystallogr D* 47: 916–924.
- Kamitori S, Okamoto A, Hirotsu K, Higuchi T, Kuramitsu S, Kagamiyama H, Matsuurra Y, Katsumi Y. 1990. Three-dimensional structures of aspartate aminotransferase from *Escherichia coli* and its mutant enzyme at 2.5 Å resolution. *J Biochem* 108: 175–184.
- Kirsch JF, Eichele G, Ford GC, Vincent MG, Jansonius JN, Gehring H, Christen P. 1984. Mechanism of action of aspartate aminotransferase proposed on the basis of its spatial structure. *J Mol Biol* 174: 497–525.
- Kleywegt GJ, Brunger AT. 1996. Checking your imagination: Applications of the free *R* value. *Structure* 4: 897–904.
- Kleywegt GJ, Jones TA. 1993. Masks made easy. *ESF/CCP4 Newsletter*. London, UK: SERC Daresbury Laboratory.
- Kleywegt GJ, Jones TA. 1994. Halloween . . . Masks and bones. In: *From first map to final model (CCP4)*. London, UK: SERC Daresbury Laboratory. pp 59–66.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24: 946–950.
- Laskowski R, MacArthur M, Moss D, Thornton J. 1993. Procheck: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26: 91–97.
- Malashkevich VN, Strokopytov BV, Borisov VV, Dauter Z, Wilson KS, Torchinsky YM. 1995. Crystal structure of the closed form of chicken cytosolic aspartate aminotransferase at 1.9 Å resolution. *J Mol Biol* 247: 111–124.
- McPhalen C, Vincent MG, Jansonius JN. 1992. X-ray structure refinement and comparison of three forms of mitochondrial aspartate aminotransferase. *J Mol Biol* 225: 495–517.
- Morris AL, MacArthur MW, Hutchinson EG, Thornton JM. 1992. Stereochemical quality of protein structure coordinates. *Proteins Struct Funct Genet* 12: 345–364.
- Navaza J. 1994. AMoRe: An automated package for molecular replacement. *Acta Cryst A* 50: 157–163.
- Okamoto A, Higuchi T, Hirotsu K, Kuramitsu S, Kagamiyama H. 1994. X-ray crystallographic study of pyridoxal 5'-phosphate-type aspartate aminotransferases from *Escherichia coli* in open and closed form. *J Biochem Tokyo* 116: 95–107.
- Read RJ. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst A* 42: 140–149.
- Smith DL, Almo SC, Toney MD, Ringe D. 1989. 2.8-Å-resolution crystal structure of an active-site mutant of aspartate aminotransferase from *Escherichia coli*. *Biochemistry* 28: 8161–8167.
- Winefield CS, Farnden KJF, Reynolds PHS, Marshall CJ. 1995. Evolutionary analysis of aspartate aminotransferases. *J Mol Evol* 40: 455–463.