Crystal Structure of a D-Amino Acid Aminotransferase: How the Protein Controls Stereoselectivity[†,‡]

Shigetoshi Sugio,[§] Gregory A. Petsko,[∥] James M. Manning,[⊥] Kenji Soda,[Δ] and Dagmar Ringe*[,∥]

Departments of Biochemistry and Chemistry and Rosenstiel Basic Medical Sciences Research Center, Brandeis University, 415 South Street, Waltham, Massachusetts 02254

Received December 28, 1994; Revised Manuscript Received May 18, 1995[⊗]

ABSTRACT: The three-dimensional structure of D-amino acid aminotransferase (D-AAT) in the pyridoxamine phosphate form has been determined crystallographically. The fold of this pyridoxal phosphate (PLP)-containing enzyme is completely different from those of any of the other enzymes that utilize PLP as part of their mechanism and whose structures are known. However, there are some striking similarities between the active sites of D-AAT and the corresponding enzyme that transaminates L-amino acids, L-aspartate aminotransferase. These similarities represent convergent evolution to a common solution of the problem of enforcing transamination chemistry on the PLP cofactor. Implications of these similarities are discussed in terms of their possible roles in the stabilization of intermediates of a transamination reaction. In addition, sequence similarity between D-AAT and branched chain L-amino acid aminotransferase suggests that this latter enzyme will also have a fold similar to that of D-AAT.

Bacteria have, in addition to L-amino acid aminotransferases, an enzyme capable of transaminating only D-amino acids (Soda & Esaki, 1985). The enzyme is important for bacteria because it catalyzes the syntheses of D-glutamic acid and D-alanine, which are essential constituents of the bacterial cell wall, and a wide variety of other D-amino acids as well. As a consequence, D-amino acid aminotransferase (D-AAT)[1] is a target enzyme for the development of novel antimicrobial agents.

D-AAT from a thermophilic *Bacillus* species has been cloned, sequenced, and expressed in *Escherichia coli* (Masu, 1985; Tanizawa et al., 1989a). The enzyme is a dimer of overall molecular mass 65 000 Da, and each monomer requires one pyridoxal phosphate (PLP) molecule as a coenzyme. The enzyme is nonspecific and will accept a wide variety of D-amino acids as substrate. D-AAT has no identity in amino acid sequence with any of the well-studied aminotransferases that are specific for L-amino acids (Tanizawa et al., 1989b), although spectroscopic characterization of this protein indicates that the enzymatic mechanism is similar to that of L-aspartate aminotransferase (L-Asp-AT) (Martinez del Poso et al., 1989). Because it is much smaller than the known L-Asp-ATs and because of the lack of sequence identity, it was expected that the protein might fold in a new and unique way. There is small but significant sequence identity with branched chain L-amino acid amino-

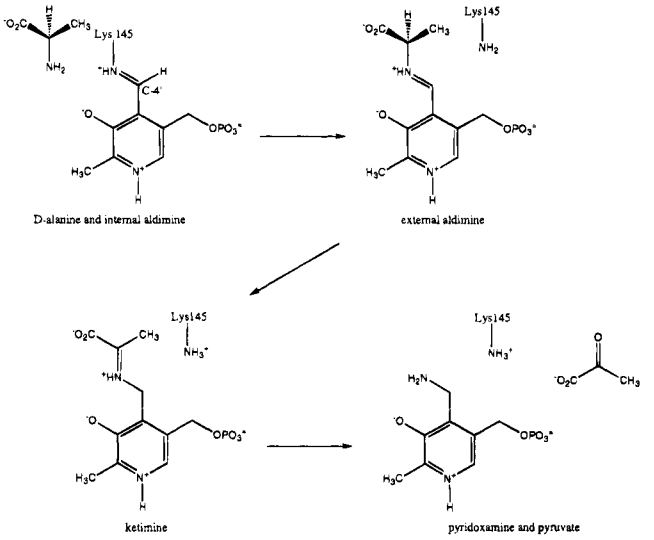

FIGURE 1: Reaction catalyzed by the D-amino acid aminotransferase. The cofactor is shown positioned as seen in the structure of D-AAT with the side facing solvent toward the viewer and the *re* side facing the protein. The incoming α-amino acid reacts with the internal aldimine between the active site lysine and pyridoxal phosphate (PLP) to give an external aldimine between the amino acid and PLP. The orientation of the amino acid must be such that the α-proton also faces the protein, with the side chain pointing toward the right. Rearrangement to the ketimine is followed by hydrolysis to the α-keto acid and pyridoxamine phosphate (PMP). Stabilization of these intermediates and the transition states leading to them promotes transamination over other possible reactions.

transferase (BC-AT) (28% with the enzyme from *E. coli*) and 4-amino-4-deoxychorismate lyase (23% with the enzyme from *E. coli*). It is therefore expected that the protein fold of these enzymes is similar to that of D-AAT.

The mechanism of transamination has been studied in great detail for the L-Asp-ATs. The reaction catalyzed by these enzymes proceeds by a ping-pong kinetic mechanism and consists of two half-reactions, each of which is comprised of three major steps (Figure 1) (Kirsch et al., 1984). The first step is the transaldimination in which the internal Schiff

[†] This work was supported by National Science Foundation Grant No. DMB9014924 and, in part, by a grant from the Lucille P. Markey Charitable Trust.

[‡] Atomic coordinates are available from the Brookhaven Protein Data Bank under entry code 1DAA.

* Address correspondence to this author at Brandeis University.

[§] Present address: The Green Cross Corp., Osaka, Japan.

[∥] Brandeis University.

[⊥] Rockefeller University, New York, NY.

[Δ] Kyoto University, Kyoto, Japan.

[⊗] Abstract published in *Advance ACS Abstracts,* July 1, 1995.

[1] Abbreviations: D-AAT, D-amino acid aminotransferase; L-Asp-AT, L-aspartate aminotransferase; BC-AT, branched chain aminotransferase; PMP, pyrdoxamine phosphate; PLP, pyridoxal phosphate.

0006-2960/95/0434-9661$09.00/0

base between a lysine residue and a PLP bound to the enzyme is replaced by a Schiff base between the substrate and the PLP, forming an external aldimine. The lysine residue is released by this process and has been shown to be the catalytic base involved in the next step (Toney & Kirsch, 1993). Both L-Asp-AT and D-AAT have such a lysine (Lys258 in L-Asp-AT and Lys145 in D-AAT), which presumably functions the same way in both enzymes. The second step is a 1,3-prototropic shift involving abstraction of a proton from the α-carbon of the substrate amino acid and reprotonation of the aldehydic carbon of the coenzyme to yield a ketimine intermediate. A number of residues are involved in the stabilization of this intermediate and/or the transition states leading up to it. Two of the most important are a tyrosine which hydrogen bonds with the phenolic oxygen of PLP and the Schiff base imine group of the internal/external aldimine and an aspartate which forms an ion pair with the pyridyl nitrogen of the cofactor. The roles of both of these residues have been probed by mutagenesis experiments in the L-Asp-ATs. Both the tyrosine and aspartic acid have been found to have a profound influence on the activity of the enzyme (Inoue et al., 1991; Goldberg et al., 1991, 1993; Onuffer & Kirsch, 1994). Therefore, it is interesting to see if their equivalents are found in the active site of D-AAT. In addition, the role of the catalytic lysine is crucial in the determination of the stereospecificity of this step (Almo et al., 1994). In the reaction catalyzed by L-Asp-AT, the lysine is facing the side of the substrate where the α-proton is located. The lysine returns the proton onto the *pro-S* side of the substrate α-carbon so that an L-amino acid results when the reaction goes in the reverse direction. This mechanism predicts that the D-amino acid must face in the opposite direction relative to the coenzyme in D-AAT in order to maintain the correct stereochemistry of the reaction. The final step in the first half-reaction is the hydrolysis of the ketimine to release the keto acid and the PMP form of the cofactor. The second half-reaction is the reversal of these steps with a different α-keto acid. This mechanism has been analyzed in structural terms based on the three-dimensional structures of L-Asp-AT inhibitor and intermediate forms (Kirsch et al., 1984; Malashkevich et al., 1993). A similar mechanism has been proposed for D-AAT.

Since the reactions catalyzed by the L-Asp-AT and D-AAT are identical and the same cofactor is used, one might anticipate that some residues of the active site would be the same for these two types of proteins, particularly those residues involved directly in the chemical transformation. In addition, substrate recognition might involve similar sites on both enzymes. Substrate is positioned properly in the L-Asp-AT active site by interaction of the α-carboxylate of the incoming amino acid with an arginine side chain. This feature of recognition could be similar on D-AAT. Side-chain recognition of the substrate, however, probably involves completely different sites since the substrate specificities of the two types of enzymes are very different. Finally, the stereospecificity of the reaction must be controlled by the orientation of the substrate relative to the PLP and the protein when it interacts with the active site. These orientations are likely to be opposite in the two enzymes. Comparison of the three-dimensional structures of L-Asp-AT and D-AAT will shed light on the uniqueness of the protein architecture required to bring about the identical chemical transformations on stereochemically different substrates. In addition, this comparison should indicate which protein residues constitute the minimum requirement to catalyze this chemical transformation and which of these residues are conserved in functional type if not in exact identity.

MATERIALS AND METHODS

Preparation of the Protein. Originally D-AAT was purified to homogeneity from *Bacillus sphaericus* (Soda et al., 1974), and its enzymatic and physicochemical properties were characterized (Yonaha et al., 1975). A thermostable variant of this enzyme was isolated from a novel thermophilic *Bacillus* species, YM-1 (Masu, 1985). The gene for the activity was sequenced, cloned into plasmid pICT113, and transformed into *E. coli*. The recombinant enzyme was overproduced and purified to homogeneity (Masu, 1985; Stoddard et al., 1987). The enzyme consists of two identical subunits, and each subunit is composed of 282 amino acid residues and a covalently bound pyridoxal-5′-phosphate (M_r 32 500).

Crystallization. The recombinant form of this protein was crystallized from an ammonium sulfate solution, and preliminary crystallographic work showed that this crystal form belongs to the hexagonal space group $P6_3$ with the unit cell parameters $a = b = 135$ Å, $c = 65$ Å, $\gamma = 120°$, and $Z = 12$ (Stoddard et al., 1987). This crystal form turned out to be unsuitable for crystal structure determination owing to the small size of the crystals, and a new crystal form was sought.

D-AAT also was crystallized from solutions containing 50 μM PLP, 1−2% (v/v) dioxane, 5 mM sodium azide, 18−22% (w/v) poly(ethylene glycol) 4000, and 25 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) at pH 5.5−6.5. Crystals were obtained in hanging drops containing 3 μL of protein diluted to 10 mg/mL with water plus 3 μL of the solution described above suspended over 0.5 mL of the reservoir solution. A few yellowish plate-shaped crystals above 1.0 mm × 1.5 mm × 0.05 mm were grown at room temperature within 1−2 weeks. The dioxane in the reservoir solution could be replaced by 2−4% (v/v) acetone, methanol, ethanol, or 2-propanol without any influence on crystal growth. Crystals were also grown from solutions containing poly(ethylene glycol) 8000 as a precipitant.

Preliminary precession camera work revealed that the crystals belong to monoclinic space group $P2_1$. The unit cell parameters are $a = 58.5$ Å, $b = 76.2$ Å, $c = 73.3$ Å, and $\beta = 103.9°$, which corresponds to a unit cell volume of 317 000 Å^3. The volume per unit mass (V_m), calculated with an assumption that there is one dimer (M_r 65 000) in the asymmetric unit, is 2.6 Å^3/Da, which is in the range of 1.6−3.6 Å^3/Da found in typical protein crystals (Matthews, 1968). These crystals were suitable for high-resolution crystal structure determination and were used to solve the three-dimensional structure of D-AAT.

Solution of the D-AAT Crystal Structure. Native crystals were soaked in reservoir solutions containing various concentrations of heavy metal reagents. These crystals were evaluated by measuring a subset of 120 reflections on a Rigaku AFC-5 four-circle diffractometer mounted on a Rigaku RU200 X-ray generator. The ethylmercuric chloride (EMC) derivative was obtained by soaking crystals in a 0.2 mM EMC solution for 6 h. No other heavy metal salt gave

Table 1: Data Collection and Refinement

	native	EMC derivative
Crystal Data		
space group	$P2_1$	$P2_1$
unit cell parameters		
a (Å)	58.5	58.4
b (Å)	76.2	76.5
c (Å)	73.3	73.0
β (deg)	103.9	103.9
Data Collection		
reflections, observed	106 545	126 180
reflections, unique	40 269	41 416
R_{sym}[a] (% on I)	6.3	8.0
resolution (Å)	∞−1.94	∞−2.39
$I/\sigma(I)$	>1.0	>1.0
completeness, overall	0.88	0.83
completeness, highest resolution	0.73	0.66
Phasing		
resolution (Å)		20.0−2.5
MFID[b]		0.25
phasing power[c]		3.87
mean figure of merit		0.64
Refinement		
resolution (Å)	10.0−1.94	
reflections [$I/\sigma(I) > 1$]	37 611	
R-factor[d]	0.184	
protein atoms	4466	
cofactor atoms	32	
water molecules	232	
B-factor model	individual	
restraints (rms observed)		
bond length (Å)	0.013	
bond angles (deg)	2.6	
improper angles (deg)	1.0	
dihedral angles (deg)	24.4	

[a] $R_{sym} = \sum\sum|F^2(i)/G(i) - \langle F^2(h)\rangle|/\sum\sum F^2(i)/G(i)$, where $F^2(i)$ is the ith integrated intensity, $\langle F^2(h)\rangle$ is the averaged integrated intensity with Miller indices $h = (hkl)$, and $G(i)$ is the inverse scale factor. [b] MFID (mean fractional isomorphous difference) $= \sum(||F_{PH}| - |F_P||)/\sum(|F_P|)$, where F_P and F_{PH} are the structure factor amplitudes for the native and heavy atom derivative, respectively. [c] Phasing power $= \langle F_H\rangle/\langle E\rangle$, where $\langle F_H\rangle$ is the rms of the heavy atom structure factor amplitude and $\langle E\rangle$ is the rms lack of closure error. [d] $R = \sum(F_o - F_c)/\sum F_o$.

an isomorphous derivative. Intensity data from native and EMC derivative crystals were collected on a San Diego Multiwire System Mark-II area detector (Xoung et al., 1985) mounted on a Rigaku RU-200H X-ray generator operated either at the University of California at San Diego or at Rutgers University. Data were collected at 20 °C with a scan width of 0.12°/frame and an exposure time of 30 s/frame. One crystal was used to collect either a native or derivative data set, and great care was taken to measure a complete set of anomalous data for the EMC derivative (Table 1). Each data set scaled internally with merging R-factors in the range of 6−8% on intensity. The EMC data scaled to the native data set with a mean fractional isomorphous difference of 0.25 on structure amplitudes.

The binding sites for mercury atoms in the EMC derivative were located from an isomorphous difference Patterson map using the program HASSP (Terwilliger & Eisenberg, 1987). Although there were 6 mercury sites/asymmetric unit (Table 2), the Patterson map was extremely clear and easy to interpret. The six sites could be grouped into two similar clusters of three, related to one another by a noncrystallographic 2-fold rotation axis. The location of this axis corresponded well with that determined from the self-rotation function applied to the native data. An anomalous difference Patterson map showed peaks corresponding to those observed in the isomorphous difference Patterson map, indicating that the measured anomalous data would be useful in phase determination.

Table 2: Refined Heavy Atom Parameters for the Ethylmercury Chloride Derivative

atom	x[a]	y[a]	z[a]	occup[b]	B_{iso}[c]	binding site[d]
HG1	0.9243	0.0000	0.6104	0.594	10.79	A Cys142 Sγ
HG2	1.3089	−0.1580	0.5803	0.709	14.81	A Cys212 Sγ
HG3	0.8144	−0.2023	0.3626	0.611	13.91	A Cys164 Sγ
HG4	0.8377	0.1835	0.9423	0.557	18.42	B Cys164 Sγ
HG5	0.3430	0.0905	0.7529	0.740	19.27	B Cys212 Sγ
HG6	0.7317	0.0505	0.6446	0.571	8.18	B Cys142 Sγ

[a] Crystallographic fractional coordinate. [b] Fractional occupancy factor. [c] Isotropic temperature factor (Å^2). [d] A or B refers to either monomer in the asymmetric unit.

Electron Density Map Interpretation and Structure Refinement. The heavy atom-derived single isomorphous replacement/single anomalous scattering phases were improved by iterative cycles of solvent flattening (Wang, 1985) and averaging of the electron density (Bricogne, 1976) about the noncrystallographic 2-fold axis using an in-house program suite (AVGSYS; Sugio, unpublished results). In the initial averaged electron density map, a number of secondary structure elements, especially α-helices and β-sheets, were clear and easy to interpret. These were built as polyalanine using the molecular graphics program FRODO (Jones, 1978). A crystallographic restrained least-squares refinement with the program PROLSQ (Hendrickson & Konnert, 1980) was carried out to improve the fit of the partial model to the electron density. Phases were calculated for this partial atomic model and combined with the heavy atom-derived phases using the program COMBINE (Kabsch et al., 1990). These combined phases were used to produce a new electron density map which was solvent-flattened and averaged. Additional model building into this improved map gave a more complete set of protein coordinates, and the process of phase combination and map improvement was repeated three times until 68% of the polypeptide chain had been traced. At this stage, the sequence was aligned with the chain tracing by using the positions of the three mercury sites per monomer as markers for the location of the three cysteine residues. Side chains were added to the model to fit the electron density, which was mostly clear and unambiguous. Additional electron density which clearly matched the shape of the PLP molecule was found at the position of the essential lysine residue. The cofactor density is not contiguous with that of the K145 side chain, indicating that the cofactor may be in the form of pyridoxamine phosphate (PMP). Recent evidence (R. M. Maming, private communication) suggests that the enzyme is purified as a mixture of the PLP and PMP forms in different ratios. On the basis of the absence of contiguous density, we will assume the cofactor is in the PMP form.

After 90% of polypeptide chain had been traced, simulated annealing was done with an overall temperature factor by the program X-PLOR (Brunger et al., 1987). Conventional positional refinement followed by individual isotropic temperature factor refinement then was applied to the model against the intensity data in the range of 10.0−1.94 Å resolution. Ordered solvent molecules were added to the

model in locations where both the electron density and local interactions were indicated, by an in-house program (WATERHUNTER; Sugio, unpublished results). The interpretation around the active site was checked by a series of "omit maps". In this procedure, all the amino acid residues within 5.0 Å from any atom of the PLP at each active site were deleted from the atomic model. The rest of the structure was subjected to conventional crystallographic refinement in order to remove "memory" about the omitted residues from the model. A difference Fourier map was calculated to show unbiased electron density for the omitted region. The structure was then checked with the conventional difference Fourier technique, and manual model corrections were made if necessary. This entire procedure, from executing the refinement program to the manual correction, was repeated until no more significant features were observed on difference Fourier maps. The overall *R*-factor calculated from the final atomic coordinates, including 232 water molecules, is 0.184 for all data with intensities greater than their estimated standard deviation up to 1.9 Å resolution. The geometry of the final model is good, with rms deviations from ideal values of 0.013 Å for bond lengths and 2.6° for bond angles (Table 1).

Homology Search and Structure Comparison. The protein sequence database SWISS-PROT (Bairoch, 1992) was searched to find amino acid sequences similar to that of D-AAT. Each sequence picked from the database was then aligned to D-AAT by the program FASTA (Pearson & Lipman, 1987). Only 7 out of 308 residues are different between amino acid sequences of BC-AT from *E. coli* and *Salmonella typhimurium*, so that the triple-sequence alignment of the two BC-ATs and D-AAT was unambiguous (Figure 2). Structure comparison against a representative database with 360 structures, using a 30% sequence identity cutoff, using the program DALI, did not reveal any similarity to previously known structures with the same topology. The 3D packing of secondary structural elements, however, has similarities with typical α/β-domains. (Hobohm et al., 1993; Holm & Sander, 1993; Holm et al., 1992).

RESULTS AND DISCUSSION

Structure Determination. The crystal structure of the thermostable D-AAT has now been solved and refined at 1.9 Å resolution. Solution of this crystal structure proved to be extremely difficult, for a number of reasons. First, although the monoclinic crystal form that we discovered was very suitable for high-resolution crystallographic study, it was not easy to grow. We finally discovered that chloride ions were essential for growth of this form and that the chloride needed to be present throughout the protein purification to facilitate the growth of monoclinic crystals. Ironically, no chloride ion could be identified in the final electron density maps. Second, it proved to be very difficult to prepare isomorphous heavy atom derivatives for phase determination. Standard heavy atom reagents like potassium tetrachloroplatinate and potassium uranylacetate failed to bind to the enzyme, as did over 30 other compounds that were tried. Our only success came through exploiting the reactivity of the three sulfhydryl residues in each D-AAT monomer. We found that small mercury-containing compounds such as EMC would bind irreversibly to the protein, presumably by reacting with the cysteine residues. However, we were unable to obtain a

```
                10        20        30        40        50
DAAT       GYTLWNDQIVKDSEVKIDKEDRGYQFGDGVYEVVKVYNGE----MFTVNEHIDRL
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_E  TTKKADYIWFNGEMVRWEDAKVHVMSHALHYGTSVFEGIRCYDSHKGPVVFRHREHMQRL
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_S  TTKKADYIWFNGEMVRWEDAKVHVMSHALHYGTSVFEGIRCYDSHKGPVVFRHREHMQRL

          60        70        80        90       100       110
DAAT    YASAEKIRITIPYTKDKFHQLLHELVEKNELNTGHIYFQVTRGTSPRAHQFPENTVKPV
        ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_E  HDSAKIYRFPVSQSIDELMEACRDVIRKNNLTSAYIRPLIFVGDVGMGVNPPAGYSTDV
          ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_S  RDSAKIYRFPVSQSIDELMEACRDVIRKNNLTSAYIRPLVFVGDVGMGVNPPGYTTDV

         120       130       140       150       160
DAAT    II-GYTKENPRPLENLEKGVKATFVEDIRWLR----CDIKSL-NLLGAVLAKQEAHEKGC
        ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_E  IIAAFPWGAYLGAEALEQGIDAMVSSWNRAAPNTIPTAAKAGGNYLSSLLVGSEARRHGY
        ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_S  IIAAFPWGAYLGAEALDQGIDAMVSSWNRAAPNTIPTAAKAGGNYLSSLLVGSEARRHGY

         170       180       190       200       210       220
DAAT    YEAI-LHRNNTVTEGSSSNVFGIKDGILYTHPANNMILKGITRDVVIACANEINMPVKEI
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_E  QEGIALDVNGYISEGAGENLFEVKDGVLFTPPFTSSALPGITRDAIIKLAKELGIEVREQ
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_S  QEGIALDVNGYISEGAGENLFEVKDGVLFTPPFTSSALPGITRDAIIKLAKELGIEVREQ

         230       240       250       260       270       280
DAAT    PFTTHEALKMDELFVTSTTSEITPVIEIDGKLIRDGKVGEWTRKLQKQFETKIPKPLHI
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_E  VLSRESLYLADEVFMSGTAAEITPVRSVDGIQVGEGRCGPVTKRIQQAFFGLFTGETEDK
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Ilve_S  VLSRESLYLADEVFMSGTAAEITPVRSVDGIQVGEGRCGPVTKRIQQAFFGLFTGETEDK

DAAT
Ilve_E  WGWLDQVNQ
Ilve_S  WGWLDPVNS
```

FIGURE 2: Sequence alignment of D-amino acid aminotransferase (D-AAT) and two branched chain amino acid aminotransferases (BC-ATs; Ilve_E, *E. coli*; Ilve_S, *S. typhimurium*). Sequence identity between D-AAT and the BC-ATs is approximately 28%. Alignment suggests a similar fold for the two types of enzymes. Only identities among all three proteins are marked.

second derivative because all other mercury compounds either bound to the same sites as EMC with identical occupancy or failed to bind to the protein at all. Therefore, we were forced to solve this structure, with over 65 000 Da of protein in the asymmetric unit, with a single heavy atom derivative. Although other protein structures have been solved with only one derivative, to our knowledge this is one of the largest low-symmetry structures for which that has been attempted. Finally, as will be shown below, the structure of D-AAT has proven to be unique, as was suspected from its lack of sequence identity with other proteins of known structure. The fold of the polypeptide chain does not belong to any of the existing families of protein motifs, not just for PLP-containing enzymes but for proteins in general. (L. Holm, private communication). Therefore, we could not make use of similarities with any existing protein structure to help interpret the electron density map.

Description of the Structure. D-AAT has a completely new protein fold (L. Holm, private communication). Although it contains the same elements, β-sheets and α-helices, that are found in other proteins, their relative arrangements here have not been observed in any other structure thus far. The structure of the monomer and the monomer−monomer relationships in the dimer bear no resemblance to L-Asp-AT (Almo et al., 1994), the β-subunit of tryptophan synthase (Hyde et al., 1988), phosphorylase b (Sprang et al., 1988), potato phosphorylase (Willnecker et al., 1991), tyrosine phenol lyase (Antson et al., 1993), ω-amino acid: pyruvate aminotransferase (Watanabe et al., 1991), or dialkylglycine decarboxylase (Toney et al., 1993), the other PLP-containing enzymes whose structures are now available. The monomer of D-AAT is comprised of two domains of very different

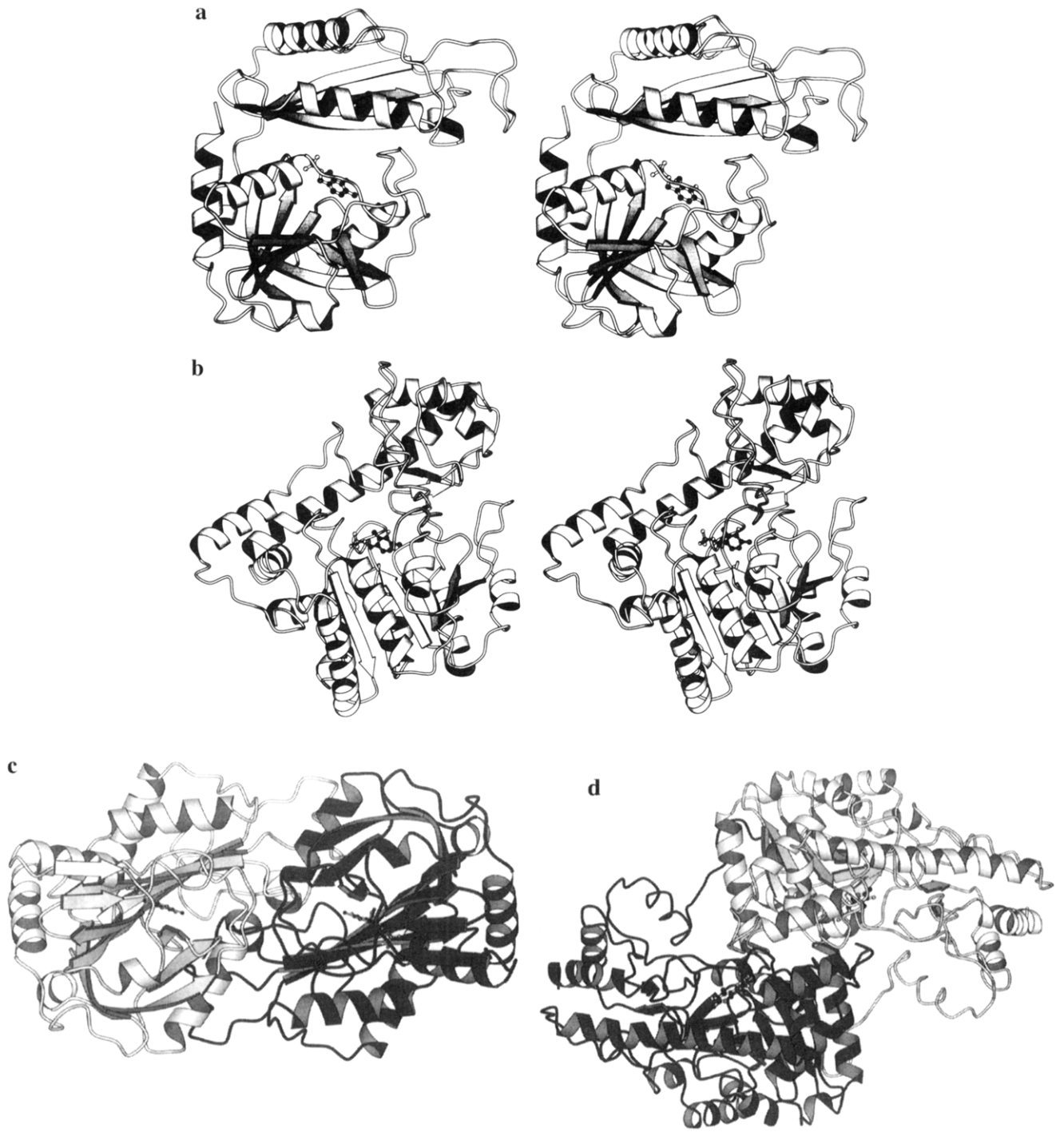

FIGURE 3: Ribbon stereodiagram of the fold of the (a) D-AAT and (b) L-Asp-AT monomers and dimers (c and d, respectively). The dimers are viewed down the 2-fold axis that describes their relative orientation to each other. The cofactor is included at the interface between domains and subunits.

secondary and tertiary structure (Figure 3a). The N-terminal domain, the "small" domain, consists of residues 1−120, while the C-terminal domain, the "large" domain, consists of residues 121−282. The N-terminal domain has a four-stranded antiparallel β-sheet consisting of strands 1, 4, 3, and 2, in that order. Helix 1, which has only 1.5 turns, together with a loop between strand 1 and helix 1, separates strands 1 and 2 well enough to allow the polypeptide chain to form a typical "Greek key" motif in this domain. Two long α-helices (helices 3 and 4) inserted between strands 2 and 3 cover one side of the N-terminal β-sheet such that they prevent solvent molecules from accessing the core of the β-sheet. The C-terminal domain, on the other hand, consists of two mixed β-sheets. One is the $\beta-\alpha-\beta$ motif formed by strand 5, helix 5, and strand 6, associated with strand 7, which is aligned in an antiparallel fashion to strand 6. The other is the four-stranded antiparallel β-sheet comprised of strands 9, 8, 11, and 12 with an additional strand, strand 10, lying parallel to strand 9 to make one edge of this mixed sheet. The two parallel strands (strands 9 and 10) connected by an α-helix (helix 6) also form a second $\beta-\alpha-\beta$ motif in the C-terminal domain. Strand 12 is followed by helix 8, which ends at residue 277. Electron density corresponding to the last five residues of either monomer was not seen in the current model, probably because they are structurally disordered.

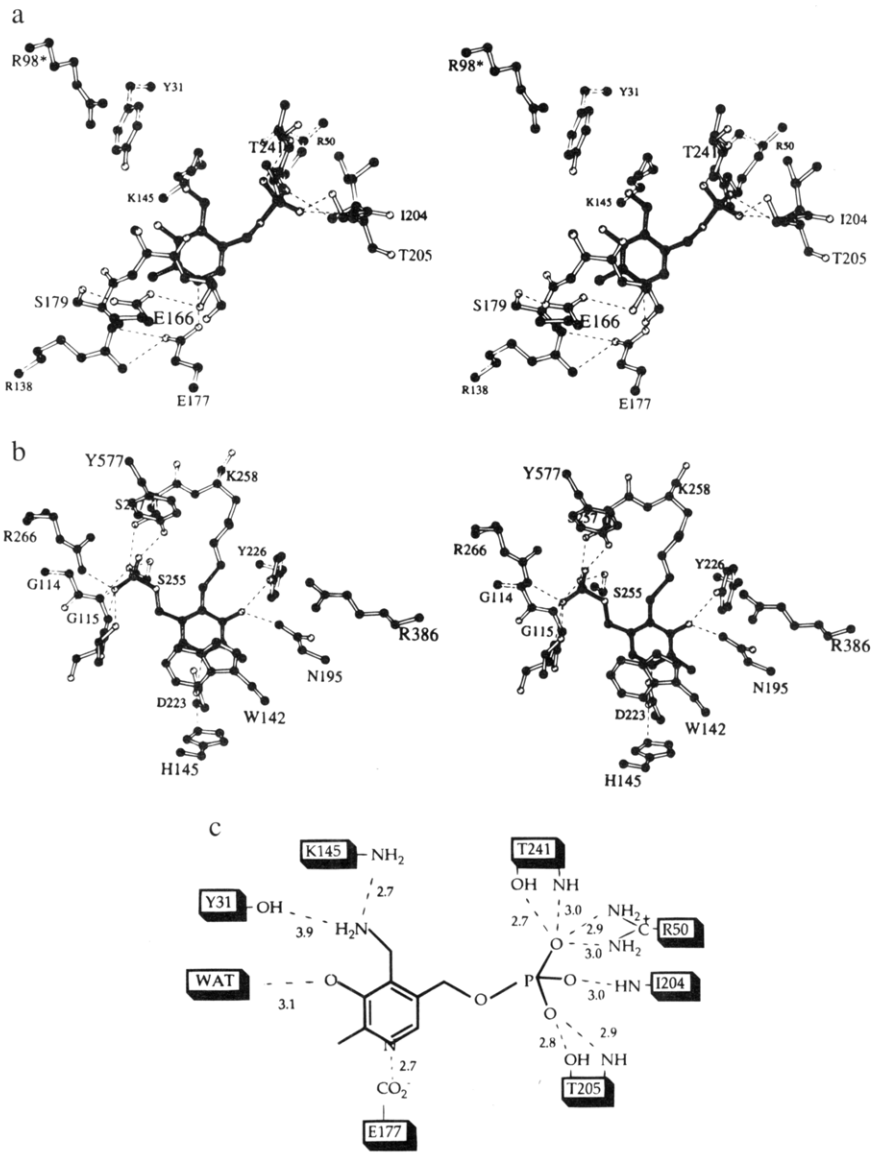

FIGURE 4: Stereoview of the active site regions of (a) D-AAT and (b) L-Asp-AT. Both views are from the direction of the solvent. Note that the two cofactors appear in opposite directions such that a proton at C4′ would be abstracted from the side of the cofactor facing the protein, i.e., the *re* face of the plane of the conjugated π-system of the cofactor−substrate imine (external Schiff base intermediate in D-AAT) and from the *si* face in L-Asp-AT. The incoming amino acid must interact with the cofactor in such a way that the α-proton is pointing toward the protein, specifically the catalytic lysine. This requires that the side chain of an α-amino acid interacting with L-Asp-AT is pointing toward the left in panel b and the side chain of a D-amino acid is pointing toward the right in panel a (as indicated schematically in Figure 1). (c) Schematic diagram of the active site region of D-AAT showing all interactions between PMP and the protein within hydrogen-bonding/salt bridge distance of the cofactor.

The N-terminal domain is packed against the C-terminal domain on one side to make a hydrophobic domain interface at which the cofactor, here PMP, is bound (Figure 3a). Three α-helices (helices 6−8) also shield this hydrophobic region on the other side of the C-terminal domain. The PMP is mainly located at the interface of those two domains of the same monomer. An extended loop comes from the other monomer of the dimer to provide residues (Arg98 and His100) to the PLP-binding region and active site of the monomer. Thus, residues from both domains and both subunits contribute to each of the two symmetrical active sites (Figure 4a,c). Since D-AAT is a two-domain protein with its active site between the domains, the possibility of domain movement during catalysis must be considered. No such movement has yet been demonstrated, although the structure would allow it: there is only a single connecting loop of polypeptide chain between the two domains. Solution scattering measurements or crystal structures of the enzyme in the form of an external aldimine with an α-methyl D-amino acid, or complexed with a specific inhibitor, would be required to show such a conformational change.

Although the environments of the two subunits in the crystallographic asymmetric unit are not equivalent and those subunits were treated independently during the entire course of the refinement, the rms deviation of the atomic positions after the best superposition was only 0.62 Å for all the main chain atoms (N, Cα, C, O) and β-carbons, which means that the structures of the two subunits in the dimer are similar at the level often found for identical molecules in different crystal forms (rms = 0.5−1.0 Å).

Recognition of the Coenzyme. Although the structures of L-Asp-AT and D-AAT do not resemble one another at all, it is surprising and exciting to observe that many of the design features of the L-Asp-AT catalytic site, both for molecular

recognition and for transamination catalysis, are replicated in the D-AAT active site. We conclude that L-Asp-AT and D-AAT are striking examples of convergent evolution to a similar chemical mechanism. This unexpected result is the most important finding to have emerged from our studies of D-AAT, and it implies that our strategy of understanding the modulation of PLP reactivity by enzymes through comparison of the structures of apparently unrelated PLP-containing enzymes will be highly productive.

The electron density map at the active site of D-AAT clearly shows the position and orientation of the PMP molecule with the *re* side facing the protein. The phosphate group of PMP is tightly held in place by nearly a dozen interactions (Figure 4a,c). The three phosphate oxygen atoms form hydrogen bonds with side-chain hydroxyl groups of Thr205 and Thr241 and main chain amides of Ile204, Thr205, and Thr241, whereas the ester oxygen (O5′) does not participate in any nonbonded interactions at all. An ion pair is also found between the guanidino group of Arg50 and the PMP phosphate group. The N-terminus of the α-helix defined by residues 203−216 points directly at the PMP phosphate group. Thus the positive end of the helix dipole interacts with the negative charge of the phosphate moiety. Such a helix dipole−phosphate interaction is also observed in the L-Asp-ATs (Figure 4b).

The nitrogen atom (N1) in the pyridoxyl ring of PMP is within ion-pairing distance of the carboxylate of Glu177, an interaction that is similar to the ion pair with Asp223 in L-Asp-AT (residue numbering according to the *E. coli* sequence). This glutamate is also hydrogen bonded to an arginine, Arg138, in a manner reminiscent of the hydrogen bond between Asp223 and His145 in L-Asp-AT. It has been observed in several crystal structures of L-Asp-AT complexed with substrates or inhibitors that N1 acts as a pivot when the pyridoxyl ring moves by rotating around torsion angles at the P−O5′, O5′−C5′, and C5′−C5 bonds, while the phosphate group does not change its position, like an anchor (Schumacher & Ringe, 1994). Even though we have not examined any other forms of the D-AAT structure it is expected that the same movement can happen to the bound PLP molecule, because the structural features of the interactions between the enzyme and the cofactor are strikingly similar in both enzymes.

Tyr31 in D-AAT is found in a similar position to Tyr226 in L-Asp-AT, but it is not clear whether the roles of these tyrosine residues are the same. In the structure of L-Asp-AT, this tyrosine hydrogen bonds to the phenolic oxygen atom (O3′) and the imine nitrogen atom of PLP and plays an important role in the stabilization of intermediates during the reaction. In the structure of D-AAT, on the other hand, this tyrosine does not interact directly with the phenolic oxygen of PMP (3.9 Å from O3′) but does make a hydrogen bond with a water molecule, which in turns forms a hydrogen bond to O3′. A similar water molecule is observed in the structure of ω-amino acid:pyruvate aminotransferase (Watanabe et al., 1991). However, since the current crystal structure does not have an intact Schiff base between the enzyme and the cofactor, an alternative interpretation may be that the formation of this covalent bond brings the tyrosine close enough to the phenolic oxygen atom to make a direct interaction. The sequence comparison between D-AAT and BC-AT shows a phenylalanine in this position in the BC-AT sequence, which cannot provide a hydrogen bond from the side chain. However, the alignment between these sequences is not unambiguous in this region. Therefore, the water molecule in the D-AAT structure may participate in the stabilization of the cofactor, and a similar water molecule may be involved in BC-AT. It should be noted that a mutant in which a phenylalanine replaces the tyrosine has been studied in L-Asp-AT (Goldberg et al., 1991; Inoue et al., 1991; Almo et al, 1994), and the resulting enzyme is a competent transaminase, although with altered kinetic parameters.

There are several big differences in the architecture of the active site between D-AAT (Figure 4a) and L-Asp-AT (Figure 4b). The tryptophan residue (Trp142 in L-Asp-AT) that forms a parallel stacking interaction with the pyridoxyl ring in L-Asp-AT is absent in D-AAT. Instead, a unique loop of three serine residues, Ser179−Ser181, hydrogen bonded to the carboxylate group of Glu166 through the side-chain hydroxyl groups, forms a barrier on one side of the pyridoxyl ring. The side-chain amino group of Lys145, which is thought to make a Schiff base with the aldehyde group of the PLP in solution, does not form a covalent linkage with any other atoms in this crystal structure. This absence of internal aldimine formation was also observed in the series of crystal structures of ω-amino acid:pyruvate aminotransferase (Watanabe et al., 1991). In contrast, Lys258 in L-Asp-AT, which is equivalent to Lys145 in D-AAT, is covalently linked with the PLP molecule in most crystal structures so far determined.

The most striking difference between L-Asp-AT and D-AAT is the relative arrangement of the lysine residues mentioned above with respect to the C4′ of the PLP. In all the reactions of PLP dependent aminotransferases specific for L-amino acids studied so far, a proton is added to or abstracted from the C4′ atom of the coenzyme−imine (external aldimine) intermediate on the *si* face. The proton transfer is facilitated by the active site lysine acting as a general base. The same strict stereospecificity is observed in other PLP-containing enzymes, such as decarboxylases, lyases, and racemases, including reactions in which transamination occurs as a side reaction. Three-dimensional structures of six such enzymes (L-Asp-AT, tryptophan synthase, tyrosine phenol lyase, ω-amino acid:pyruvate aminotransferase, and dialkylglycine decarboxylase) now available, show that the lysine forming the Schiff base approaches the coenzyme from the *si* face, the side facing the protein in each case, and this structural restriction causes the strict stereospecificity of proton exchange at the C4′ of PLP. In the case of L-AspAT, the amino acid forming the Schiff base is oriented such that the α-proton faces the protein and addition of a proton to the ketimine intermediate at the α-carbon results in the formation of an L-amino acid with the *S* configuration.

Recent hydrogen transfer experiments with D-AAT and BC-AT, however, show that the *pro-R* hydrogen is transferred at the C4′ position of the cofactor on the *re* face, which suggests a structural difference between these enzymes and the "ordinary" aminotransferases (Yoshimura et al., 1993). In the crystal structure of D-AAT, the coenzyme must therefore be bound to the enzyme such that the side chain of the catalytic lysine 145 extends toward the coenzyme on the *re* face, which is the right position from which to catalyze *pro-R* proton transfer at C4′. The crystal structure confirms this orientation and suggests by analogy that a D-amino acid

should be oriented as shown in Figure 1 with the α-proton pointing toward the protein.

Substrate Recognition

D-AAT vs L-Asp-AT. The structure of D-AAT clearly shows a cavity above the coenzyme-binding site. This region can provide the functionalities essential for recognition of D-amino acids as substrates, i.e., an imine of the PLP to form a Schiff base with the substrate, a "carboxylate trap" which must be positively charged, and a "side--chain pocket" which discriminates substrate species. It is mandatory for the aminotransferase reaction that the α-amino group of the substrate must be located close enough to make a covalant bond to C4′ of PLP in order to form an external aldimine complex. Under the above structural restraint, the α-carboxylate of the substrate would interact electrostatically with an arginine from the N-terminal domain of the other subunit of the dimer (Arg98*) (Figure 4a). No other candidates for the "carboxylate trap" can be found in the putative substrate-binding area. This feature is different from the interaction geometry observed in L-Asp-ATs, where the α-carboxylate interacts electrostatically with an arginine (Arg386) coming from the small domain of the same subunit (Figure 4b). When a D-amino acid is bound to the putative substrate recognition site of D-AAT by forming a Schiff base and this electrostatic interaction, the side chain of the substrate is automatically located in a region defined by a loop from Ser240 to Ser243. The side chains of Ser240, Thr242, and Ser243 in this loop face the same direction and form a "hole" together with the hydroxyl group of Ser180. This "hole" looks like the entrance of a pocket in which the side chain of the substrate can be caught and provides a surface where either nonpolar or polar side chains can interact. This side chain recognition mode is completely different from that in the L-Asp-ATs, which is dominated by the interaction of the charged substrate side chain with an arginine from the large domain of the other subunit (Arg292*). Given the recognition mechanism proposed above, D-AAT would bind a D-amino acid to form an external aldimine with the same orientation of α-amino and α-carboxyl groups relative to the orientation of the pyridoxyl phosphate ring as L-Asp-ATs bind L-amino acids, while the side-chain orientation is different due to the inversion of chirality at the α-carbon of the substrate. This difference as well as the side chain direction of the catalytic lysine (Lys145 in D-AAT and Lys258 in L-Asp-ATs) would directly determine the strict discrimination of L- and D-amino acids as substrates of these enzymes.

D-AAT vs BC-AT. D-AAT and BC-AT have 28% identical residues between their amino acid sequences. Therefore, it is expected that these enzymes have a similar molecular framework and use the same recognition machinery for the coenzyme. Particularly, a number of active site residues are conserved between these enzymes, such as K145 (active site lysine), E177 (interacts with pyridoxal nitrogen), I204, T205, T241, and R50 (interact with phosphate), suggesting that the cofactor is positioned in a similar orientation in BC-AT. This is confirmed by the observation that for BC-AT the *pro-R* hydrogen is transferred at the C4′ position of the cofactor on the *re* face. This leads to the conclusion that an L-amino acid must interact with BC-AT in the opposite orientation from the interaction between a D-amino acid and D-AAT. The obvious explanation for such altered substrate recognition is a different α-carboxylate recognition site. We have proposed R98 as the recognition residue in D-AAT. This residue is a methionine in BC-AT. Searching for a candidate on the other side of the active site leads to residues 35 and 88, both of which are Arg in BC-AT and neither of which have any interpretable interactions in the D-AAT. Only mutagenesis can distinguish between these two candidates.

Comparison with Results from Mutagenesis Experiments. The structure of D-AAT enables us to correlate the structural data with the results of previous mutagenesis experiments. In the first such experiment, one of the three tryptophan residues, Trp139, was changed to several other amino acids on the assumption (now shown by the D-AAT structure to be incorrect) that the parallel interaction between PLP and a tryptophan side chain, observed in all L-Asp-AT structures, would be important in D-AAT as well (Martinez del Poso et al., 1989). Variable changes in activity were observed. More striking were the differences in stability of the resulting proteins. The only mutant sufficiently stable for study was the W139F mutant. The mutant protein with an aspartate at this position (W139D) could not be isolated at all. It is now clear why these effects are observed. Trp139 lies at the interface between subunits, in a pocket formed by hydrophobic residues originating from both subunits. These include several isoleucines and leucines, the side chain of an arginine, and Trp139 from the other subunit. A smaller residue, such as alanine or proline, could result in improper packing of the subunits to each other. Charged residues, such as aspartate or histidine, would result in destabilization of this region of the interface. Presumably, the W139F mutant is the only stable protein because the phenylalanine side chain can substitute for indole in this nonpolar environment.

The second series of experiments aimed at our understanding the roles of the three cysteine residues in D-AAT (Merola et al., 1989). These residues, when mutated to glycine, resulted in proteins that had properties identical to the wild type enzyme in terms of specific activity and thermal stability. Two of these residues, C164 and C212, are at the surface of the protein and do not make any interactions with either the cofactor or the other subunit. The third cysteine, C142, lies close to the active site PLP and the subunit interface Trp139. However, these mutations do not seem to affect the subunit packing.

An interesting perturbation occurs when serine 146, adjacent to the active site Lys145, is changed to alanine (Merola et al., 1989). The mutation does not change the enzymatic activity of the enzyme but does change the SH titration behavior. In the wild type enzyme, only 1 SH/dimer is titrated by DTNB. However, 3−4 cysteine residues/dimer are titrated in the S146A mutant. In the presence of D-alanine, the mutant is titrated like the wild type enzyme. This result was interpreted to suggest that C142 is at or near the active site. The structure of D-AAT confirms this suspicion and explains the result in structural terms. Cys142 is indeed near the PLP cofactor, as is serine 146. When the serine side chain is shortened to alanine, the cysteine is exposed to the active site region and hence the solvent, thereby making the cysteine accessible to thiol reagents.

CONCLUSIONS

The three-dimensional structure of D-AAT is a completely new protein fold and bears no obvious resemblance to L-Asp-

AT or any other PLP-containing enzyme. Comparison of this structure with that of L-Asp-AT shows, however, that some features of the active site are similar to that of L-Asp-AT. A model of BC-AT constructed from the coordinates of D-AAT and sequence comparison provides some insight into the stereochemical recognition of the substrate amino acids of these two enzymes and makes some predictions about the structure of the active site of BC-AT.

ACKNOWLEDGMENT

We are grateful to Professor N.-h. Xuong for making the UCSD Research Resource for Protein Crystallography (NIH-RR01644) available to us for data collection with the expert assistance of Dr. Narendra Narayan. We thank Professor Ann M. Stock for the use of X-ray equipment at the Center for Advanced Biotechnology and Medicine and the University of Medicine and Dentistry of New Jersey. We thank Dan Peisach for help in making figures.

REFERENCES

Almo, S. C., Smith, D. L., Danishefsky, A. T., & Ringe, D. (1994) *Protein Eng. 7*, 405−412.

Antson, A. A., Denidkina, T. V., Gollnick, P., Dauter, Z., Von Tersch, R. L., Long, J., Berezhnoy, S. N., Phillips, R. S., Harutyunan, E. H., & Wilson, K. S. (1993) *Biochemistry, 32*, 4195−4206.

Bairoch, A. (1992) *Nucleic Acids Res. 20*, 2019−2022.

Bricogne, G. (1976) *Acta Crystallogr. A32*, 832−847.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.

Goldberg, J. M., Swanson, R. V., Goodman, H. S., & Kirsch, J. F. (1991) *Biochemistry 30*, 305−312.

Goldberg, J. N., Zheng, J., Deng, H., Chen, Y. Q., Callender, R., & Kirsch, J. F. (1993) *Biochemistry 32*, 8092−8097.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Eds., pp 13.01−13.23, Indian Academy of Sciences, Bangalore.

Hobohm. (1993) *Protein Sci. 1*, 409−417.

Holm, L., & Sander, C. (1992) *J. Mol. Biol. 223*, 123−138.

Holm, L., Ouzounis, C., Tuparev, G., Vriend, G., & Sander, C. (1992) *Protein Sci. 1*, 1691−1698.

Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., & Davies, D. R. (1988) *J. Biol. Chem. 263*, 7857−7871.

Ichihara, A. (1985) in *Transaminases* (Christen, P., & Metzler, D. E., Eds.) pp 430−438, J. Wiley & Sons, New York.

Iuoue, K., Kuramitsu, S., Okamoto, A., Hirotsu, K., Higuchi, T., Morino, Y., 7 Kagamiyama, H. (1991) *J. Biochem. 109*, 570−576.

Jones,T. A. (1978) *J. Appl. Crystallogr. 11*, 268−272.

Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F., & Holmes, K. C. (1990) *Nature 347*, 37−44.

Kirsch, J. F., Eichele, G., Ford, G. C., Vincent, M. G., Jansonius, J. N., Gehring, H., & Christen, P. (1984) *J. Mol. Biol. 174*, 497−525.

Malashkevich, V. N., Toney, M. D., & Jansonius, J. N. (1993) *Biochemistry, 32*, 13451−13462.

Martinez del Pozo, A., Merola, M., Ueno, H., Manning, J. M., Tanizawa, K., Nishimura, K., Asano, S., Tanaka, H., Soda, K., Ringe, D. & Petsko, G. A. (1989) *Biochemistry 28*, 510−516.

Masu, Y. (1985) M.Sc. Thesis, Institute for Chemical Research, Kyoto University, Kyoto, Japan.

Matthews, B. W. (1968) *J. Mol.Biol., 33*, 491−497.

Merola, M., Martinez del Pozo, A., Ueno, H., Recsei, P., Di Donato, A., Manning, J. M., Tanizawa, K., Masu, Y., Asano, S., Tanaka, H., Soda, K., Ringe, D., & Petsko, G. A. (1989) *Biochemistry 28*, 505−509.

Onuffer, J. J., & Kirsch, J. F. (1994) *Protein Eng. 7* , 413−424.

Pearson, W. R., & Lipman, D. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2444−2448.

Soda, K., & Esaki, S. (1985) in Transaminases (Christen, P., & Metzler, D. E., Eds.) pp 463−469, J. Wiley & Sons, NY.

Soda, K., Yonaha, K., Misono, H., & Osugi, M. (1974) *FEBS Lett. 46*, 359−363.

Sprang, S. R., Acharya, K. R., Goldsmith, E. J., Stuart, D. I., Varvill, K., Fletterick, R. J., Madsen, N. B., & Johnson, L. N. (1988) *Nature 336*, 215−221.

Stoddard, B., Howell, L., Asano, S., Soda, K., Tanizawa, K., Ringe, R., & Petsko, G.A. (1987) *J. Mol. Biol. 196*, 441−442.

Tanizawa, T., Masu, Y., Asano, S., Tanaka, H., & Soda, K. (1989a) *J. Biol. Chem. 264*, 2445−2449.

Tanizawa, T., Asano, S., Masu, Y., Kuramitsu, S., Kagamiyama, H., Tanaka, H., & Soda, K. (1989b) *J. Biol. Chem. 264*, 2450−2454.

Terwilliger, T. C., & Eisenberg, D. (1987) *Acta Crystallogr. A43*, 1−5.

Toney, M. D., & Kirsch, J. F. (1993) *Biochemistry 32*, 1471−1479.

Toney, M. D., Hohenester, E., Cowan, S. W., & Jansonius, J. N. (1993) *Science 261*, 756−759.

Wang, B. C. (1985) *Methods Enzymol. 115*, 90−112.

Watanabe, N., Yonaha, K., Sakabe, K., Sakabe, N., Aibara, S., & Morita, Y. (1991) in *Enzymes Dependent on Pyridoxal Phosphte and Other Carbonyl Compounds as Cofactors* (Fukui, T., Kagamiyama, H. Soda, K., & Wada, H., Eds.) pp 121−124 Pergamon Press, Oxford.

Willnecker, J., Jahnke, K., & Buchner, M. (1991) in *Enzymes Dependent on Cofactors* (Fukui, T., Kagamiyama, H., Soda, K., & Wada, H., Eds.) pp 393−396 Pergamon Press, Oxford.

Xoung, N.-h., Sullivan, D., Nielsen, C., & Hamlin, R. (1985) *Acta Crystallogr. B41*, 267−269.

Yonaha, K., Misono, H., & Soda, K. (1975) *FEBS Lett. 55*, 265−267.

Yoshimura, T., Nishimura, K., Ito, J., Esaki, N., Kagamiyama, H., Manning, J. M., & Soda, K. (1993) *J. Am. Chem. Soc. 115*, 3897−3900.

BI942981W