

# Structures of *Escherichia coli* Branched-Chain Amino Acid Aminotransferase and Its Complexes with 4-Methylvalerate and 2-Methylleucine: Induced Fit and Substrate Recognition of the Enzyme<sup>†,‡</sup>

Kengo Okada,<sup>§,||</sup> Ken Hirotsu,<sup>\*,§</sup> Hideyuki Hayashi,<sup>⊥</sup> and Hiroyuki Kagamiyama<sup>⊥</sup>

Department of Chemistry, Graduate School of Science, Osaka City University, Sugimoto, Sumiyoshi-ku, Osaka 558-8585, Japan, and Department of Biochemistry, Osaka Medical College, Takatsuki, Osaka 569-8686, Japan

Received February 23, 2001; Revised Manuscript Received April 26, 2001

**ABSTRACT:** The following three-dimensional structures of three forms of *Escherichia coli* branched-chain amino acid aminotransferase (eBCAT) have been determined by the X-ray diffraction method: the unliganded pyridoxal 5'-phosphate (PLP) form at a 2.1 Å resolution, and the two complexes with the substrate analogues, 4-methylvalerate (4-MeVA) as the Michaelis complex model and 2-methylleucine (2-MeLeu) as the external aldimine model at 2.4 Å resolution. The enzyme is a trimer of dimers, and each subunit consists of small and large domains, and the interdomain loop. The active site is formed by the residues at the domain interface and those from two loops of the other subunit of the dimer unit, and binds one PLP with its *re*-face directed toward the protein side. Upon binding of a substrate, Arg40 changes its side-chain direction to interact with the interdomain loop, and the loop, which is disordered in the unliganded form, shows its ordered structure on the active-site cavity, interacts with the hydrophobic side chain of the substrate, and shields it from the solvent region. The substrate binds to the active-site pocket with its  $\alpha$ -hydrogen toward the protein side, its side-chain on the side of O3 of PLP, and its  $\alpha$ -carboxylate on the side of the phosphate group of PLP. The hydrophobic side-chain of the substrate is recognized by Phe36, Trp126, Tyr129, Tyr164, Tyr31\*, and Val109\*. The  $\alpha$ -carboxylate of the substrate binds to the unique site constructed by three polar groups (two main-chain NH groups of the  $\beta$ -turn at Thr257 and Ala258 and the hydroxy group of Tyr95) which are activated by the access of Arg40 to the main-chain C=O group of the  $\beta$ -turn and the coordination of Arg97 to the hydroxy group. Since Arg40 is the only residue that significantly changes its side-chain conformation and directly interacts with the interdomain loop and the  $\beta$ -turn, the residue plays important roles in the induced fit of the interdomain loop and the  $\alpha$ -carboxylate recognition of the substrate.

Branched-chain amino acid aminotransferase (BCAT<sup>1</sup>) catalyzes a reversible transamination reaction between the L-branched-chain amino and  $\alpha$ -keto acids (1). The eBCAT is a homo hexamer, which is an assembly of three dimer units around a 3-fold axis (2). Each subunit consists of 308 amino acid residues with a molecular weight of 31 500 and contains a bound pyridoxal 5'-phosphate (PLP) as a cofactor.

<sup>†</sup> This study was supported by a Research Grant from the Japan Society for the Promotion of Science (Research for the Future) and by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan (12020249 and 13480196).

<sup>‡</sup> Coordinates for eBCAT and its complexes with 4-MeVA and 2-MeLeu have been deposited in the RSCB Protein Data Bank as entries 1IIK, 1IIM, and 1IIL.

\* To whom correspondence should be addressed. E-mail: hirotsu@sci.osaka-cu.ac.jp. Fax: +81-6-6605-3131.

<sup>§</sup> Osaka City University.

<sup>⊥</sup> Present address: Department of Molecular Biology, Nara Institute of Science and Technology, Ikoma, Nara, 630-0101, Japan.

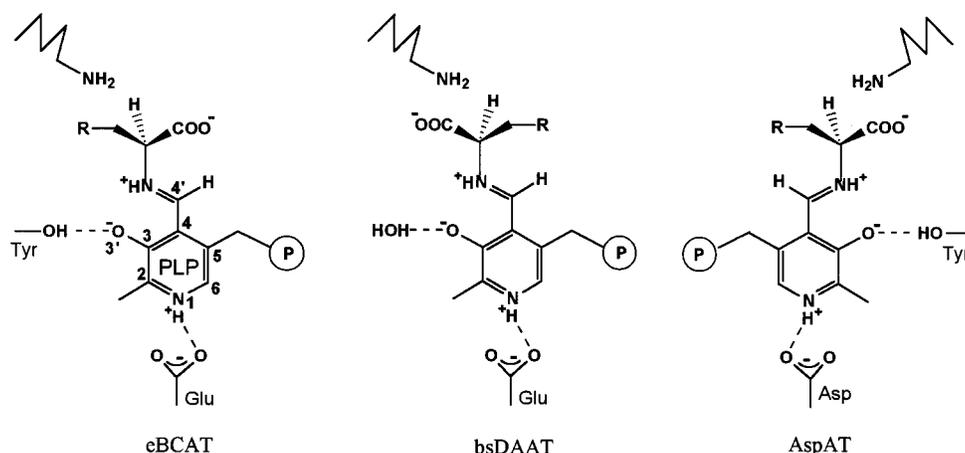
<sup>||</sup> Osaka Medical College.

<sup>1</sup> Abbreviations: BCAT, branched-chain amino acid aminotransferase; eBCAT, *Escherichia coli* BCAT; PLP, pyridoxal 5'-phosphate; AspAT, aspartate aminotransferase; bsDAAT, *Bacillus* sp. YM-1 D-amino acid aminotransferase; 4-MeVA, 4-methylvalerate; 2-MeLeu, 2-methylleucine; ASA, accessible surface area; rms, root-mean-square; Tyr31\*, the asterisk (\*) indicates a residue from another subunit of the dimer unit.

On the basis of the similarities in the sequence, secondary structure, and hydrophobicity profiles, many PLP-dependent enzymes are classified into the four currently known families with distinct fold types (3, 4). Aspartate aminotransferase (AspAT), which belongs to the fold type I and has been extensively studied (5–12), shuttles protons on the *si*-face of the planar  $\pi$ -system of the substrate-cofactor complex, while the fold type IV enzyme transfers protons on the *re*-face (13). The structure of *Bacillus* sp. YM-1 D-amino acid aminotransferase (bsDAAT) was determined as the first structure of the fold type IV enzyme confirming the *re*-face specificity of this enzyme (14). eBCAT and *Escherichia coli* 4-amino-4-deoxychorismate lyase (eADCL), classified as the fold type IV, were then subjected to X-ray crystallographic studies, and the *re*-face specificity of these enzymes were unambiguously demonstrated (2, 15).

The sequence identities of eBCAT with bsDAAT and eADCL are 28 and 25%, respectively, and the overall structures of the dimer units in these enzymes were quite similar (2, 15). The relative orientation of the substrate–PLP complexes in AspAT, bsDAAT, and eBCAT is viewed down from the solvent side (Scheme 1). The PLP rings of bsDAAT and eBCAT rotate 180° compared with that of

Scheme 1



AspAT of the fold type I. AspAT and bsDAAT have substrate–PLP configurations which are mirror images of each other. The orientation of a bound substrate in eBCAT is the same as that in AspAT, and that of PLP is the same as that in bsDAAT. The important residues and hydrogen bonds for the catalytic action are conserved in these enzymes. The catalytic lysine residue approaches the  $\alpha$ -hydrogen of the substrate, and glutamate and tyrosine (H<sub>2</sub>O in bsDAAT) form hydrogen bonds with the protonated N1 and O3 of PLP, respectively (Scheme 1).

Although the overall and active-site structures of the unliganded eBCAT have been determined (2), the mechanism for the substrate recognition and the behavior of the interdomain loop have not yet been uncovered. The eBCAT complexed with substrate analogues should help in understanding the role of the active-site residues, the mechanism of substrate recognition, and the function of the interdomain loop during the course of the catalytic process. Moreover, a comparison of the structure of eBCAT with those of currently known aminotransferases of the fold type I and bsDAAT belonging to the same fold type as eBCAT, provides new insight into the versatility and flexibility of the substrate recognition between the enzymes having different stereospecificities (2, 5–12, 14).

We now report the X-ray crystallographic studies of the following three forms of eBCAT: the unliganded PLP form at 2.1 Å resolution, the complex form with 4-methylvalerate (4-MeVA) as the Michaelis complex model at 2.4 Å resolution, and the complex form with 2-methylleucine (2-MeLeu) as the external aldimine model at 2.4 Å resolution.

## MATERIALS AND METHODS

**Crystallization and Data Collection.** All chemicals were of the highest purity available from commercial sources. The expression of eBCAT, the purification of the expressed enzyme, and the crystallization of the purified enzyme have been reported elsewhere (1, 2). Briefly, the unliganded eBCAT in PLP form was concentrated to 35 mg/mL in 0.1 M HEPES buffer, pH 7.5, containing 10  $\mu$ M PLP and 1 mM sodium azide. The protein was crystallized by the hanging-drop vapor-diffusion method using 28% w/v PEG400 as a precipitant and 0.1 M MgCl<sub>2</sub> as an additive. Three microliters of protein solution was mixed with the same volume of the above crystallization solution and was kept at 20 °C over

0.5 mL of the crystallization solution to give yellow prismatic single crystals suitable for the X-ray diffraction experiment with dimensions of 0.2  $\times$  0.3  $\times$  0.6 mm<sup>3</sup> after several days. The preliminary X-ray diffraction experiment showed that the unliganded eBCAT was crystallized as a new form with the space group *C*222<sub>1</sub>, giving higher resolution X-ray data than the crystal form (space group *C*2) previously obtained and used for the structure determination. The crystals of eBCAT in the complex with 4-MeVA and eBCAT in the complex with 2-MeLeu were obtained by a cocrystallization method using the crystallization solutions containing 0.1 M of each substrate analogue and found to be isomorphous with the unliganded eBCAT. The complex crystals with a deep yellow color compared with the unliganded crystals appeared within 1 week.

All data were collected at 20 °C on a Rigaku RAXIS IIC imaging plate detector with graphite-monochromated CuK $\alpha$  radiation produced by a Rigaku rotating anode generator operating at 40 kV and 100 mA. The X-ray diffraction data for the unliganded eBCAT, eBCAT·2-MeLeu, and the eBCAT·4-MeVA were collected to 2.1, 2.4, and 2.4 Å resolution, respectively. All crystals belong to space group *C*222<sub>1</sub> had the following cell dimensions: *a* = 156.0, *b* = 100.9, and *c* = 141.5 Å for the unliganded eBCAT, *a* = 155.9, *b* = 100.7, and *c* = 141.5 Å for eBCAT·4-MeVA, and *a* = 156.1, *b* = 100.7, and *c* = 141.5 Å for eBCAT·2-MeLeu. There are three subunits in the asymmetric unit, and approximately 54% of the crystal volume is occupied by solvent. All data were processed and scaled using the programs, DENZO and SCALEPACK (16) (Table 1).

**Structure Determination.** The molecular replacement method was applied to solve the structure of the unliganded eBCAT in the orthorhombic form, using the structure of the unliganded eBCAT in the monoclinic form previously determined at a 2.5 Å resolution (2). The initial orientation matrix of the unliganded form was determined by the program AmoRe (17) using X-ray data from 12 to 4 Å resolution. Since the electron density map located at the boundary region between the protein molecules and bulk solvent was not clear and the initial *R*<sub>factor</sub> value indicated more than 0.35, the density modification including solvent flattening (18), histogram matching (19), and 3-fold non-crystallographic symmetry (NCS) averaging (20) were performed with the program DM (21) to give the significantly

Table 1: Data Collection and Refinement Statistics

	crystal	native	4-MeVA	2-MeLeu
diffraction data resolution (Å)	2.1	2.4	2.4	2.4
no. of reflections				
unique	62 258	42 569	42 170	42 170
observed	197 817	134 092	132 437	132 437
completeness (%)	95.3 (88.4) <sup>b</sup>	98.9 (98.5) <sup>b</sup>	98.6 (98.4) <sup>b</sup>	98.6 (98.4) <sup>b</sup>
$R_{\text{merge}}^a$ (%)	4.9 (19.3) <sup>b</sup>	5.2(16.5) <sup>b</sup>	5.7 (18.1) <sup>b</sup>	5.7 (18.1) <sup>b</sup>
refinement				
resolution limits (Å)	30–2.1	40–2.4	40–2.4	40–2.4
$R_{\text{factor}}$ (%)	18.9 (25.3) <sup>b</sup>	17.4 (23.4) <sup>b</sup>	18.2 (25.0) <sup>b</sup>	18.2 (25.0) <sup>b</sup>
$R_{\text{free}}$ (%)	23.1 (30.9) <sup>b</sup>	21.5 (27.6) <sup>b</sup>	22.5 (28.3) <sup>b</sup>	22.5 (28.3) <sup>b</sup>
deviations				
bond lengths (Å)	0.007	0.007	0.007	0.007
bond angles (deg)	1.2	1.2	1.3	1.3
mean $B$ factors				
main-chain atoms (Å <sup>2</sup> )	22.6	19.5	20.8	20.8
side-chain atoms (Å <sup>2</sup> )	24.7	21.1	22.6	22.6
cofactor atoms (Å <sup>2</sup> )	16.4	17.3	23.3	23.3
water atoms (Å <sup>2</sup> )	23.6	20.3	21.7	21.7

<sup>a</sup>  $R_{\text{merge}} = \sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_i I_{hkl,i}$ , where  $I$  = observed intensity and  $\langle I \rangle$  = average intensity for multiple measurements. <sup>b</sup> The values in the parentheses are for highest resolution shells (2.20–2.10 Å) in the native enzyme, (2.51–2.40 Å) in the 4-MeVA complex, and (2.51–2.40 Å) in the 2-MeLeu complex.

improved electron density map. The modeling of the polypeptide chains was done by the program O (22). The structure of the enzyme was refined by simulated annealing and energy minimization with 3-fold NCS restraints with the program X-PLOR (23, 24), using the X-ray data from 30 to 2.5 Å resolution. The topology of the entire structure of the enzyme was carefully checked and the correctness of the structure was verified by inspection of the electron density from the maps successively omitting 10 residues. The polypeptide chain was traced to the following part: Lys4–Trp126 and Ala134–Gln308 without ambiguity. However, the rest of the parts, Thr1–Lys3 and Gly127–Glu133, were not visible in any maps. In the active site, an additional electron density, which clearly matched the shape of the PLP molecule, was found around the catalytic lysine (Lys159) residue. The atomic models of the other two subunits in the asymmetric unit were obtained by taking advantage of the matrix of a noncrystallographic 3-fold rotation symmetry. At this stage of refinement, the resolution was increased to 2.1 Å, and the restraint on the 3-fold NCS was relaxed after several rounds of refinement and manual rebuilding. As a result,  $R_{\text{factor}}$  and  $R_{\text{free}}$  were reduced to 0.212 and 0.254, respectively. The ordered water molecules were added to the refined model based on the peak heights and distance criteria from both the  $2F_o - F_c$  and  $F_o - F_c$  maps. The water molecules that showed thermal factors higher than 50 Å<sup>2</sup> were removed from the list. Further model building and refinement cycles resulted in an  $R_{\text{factor}}$  of 0.189 and  $R_{\text{free}}$  of 0.231, using 61 843 reflections [ $F_o > 1\sigma(F_o)$ ] between 30.0 and 2.1 Å resolution (Table 1).

Since the  $R_{\text{diff}}$  values between the unliganded form and the two complexes (eBCAT•4-MeVA and eBCAT•2-MeLeu) were less than 0.05, the initial phases computed from the structure of the unliganded form except for the PLP cofactors and solvents were applied to the diffraction data between 40 and 2.4 Å for both complexes, resulting in an  $R_{\text{factor}}$  of 0.33 for eBCAT•4-MeVA and 0.35 for eBCAT•2-MeLeu. Difference Fourier maps of the active sites were calculated and the large electron densities were assigned to the PLP,

4-MeVA, and 2-MeLeu molecules. A Schiff base linkage between 2-MeLeu and PLP was clearly shown on an  $F_o - F_c$  map from the eBCAT•2-MeLeu data, whereas such a linkage was not found on that from the eBCAT•4-MeVA data, indicating that eBCAT•4-MeVA can be regarded as an enzyme–substrate complex model and eBCAT•2-MeLeu as an external aldimine model. Interestingly, the missing region (Gly127–Glu133) of the interdomain loop in the unliganded form was definitely located on an  $F_o - F_c$  map of eBCAT•4-MeVA or eBCAT•2-MeLeu. Refinements in the structures for eBCAT•4-MeVA and the eBCAT•2-MeLeu were done in a similar way as that for the unliganded one. Model building and refinement cycles resulted in an  $R_{\text{factor}}$  of 0.173 and  $R_{\text{free}}$  of 0.215 using 43 213 reflections [ $F_o > 1\sigma(F_o)$ ] between 40.0 and a 2.4 Å resolution for the eBCAT•4-MeVA complex and in an  $R_{\text{factor}}$  of 0.182 and  $R_{\text{free}}$  of 0.224 using 43 409 reflections [ $F_o > 1\sigma(F_o)$ ] between 40.0 and a 2.4 Å resolution for the eBCAT•2-MeLeu complex (Table 1).

**Quality of the Structures.** The final model for the unliganded eBCAT comprises 3 × 298 residues, (three N-terminal residues and seven residues of the interdomain loop of each subunit are not visible), three PLPs, and 304 water molecules. The main-chain atoms of three subunits in the asymmetric unit were superimposed by least-squares fitting with an rms deviation of 0.12 Å and a maximum deviation of 0.40 Å, indicating that the three independent subunits have the same structure. The average thermal factors of the main-chain atoms (N, C $\alpha$ , C, and O) in the three subunits are 20.4, 23.5, and 23.8 Å<sup>2</sup>. The final models for two complexes comprise 3 × 304 residues, (three N-terminal residues and one C-terminal residue of each subunit are not visible), three PLPs, and 295 water molecules for eBCAT•4-MeVA and 304 water molecules for eBCAT•2-MeLeu. The main-chain atoms of the three subunits in the asymmetric unit were superimposed by least-squares fitting with an rms deviation of 0.04 and 0.06 Å with a maximum deviation of 0.20 and 0.53 Å for eBCAT•4-MeVA and eBCAT•2-MeLeu, respectively, implying that the overall structures of these subunits are quite similar. The average thermal factors of the main-chain atoms (N, C $\alpha$ , C, O) in the three subunits are 18.0, 20.7, and 20.7 Å<sup>2</sup> for eBCAT•4-MeVA, and 19.2, 21.7, and 21.7 Å<sup>2</sup> for eBCAT•2-MeLeu.

Analysis of the stereochemistry with PROCHECK (25) showed that all the main-chain atoms except for Glu260 for the unliganded form, Glu260 for the eBCAT•4-MeVA complex, and Phe208 and Glu260 for the eBCAT•2-MeLeu complex fall within the generously allowed regions of the Ramachandran plot for all structures. On the basis of the electron density maps, it is confirmed that the conformations of Phe208 and Glu260 are correct. We used the subunits with the lowest average thermal factors of the main-chain atoms as the reference standard for all three forms of the enzymes in the following description.

## RESULTS AND DISCUSSION

**Overall Structure.** The overall structure of the unliganded eBCAT in the orthorhombic form shown in Figure 1 (26) is the same as that in the monoclinic form (2), although the crystal packing interaction between the two forms is different. The shape of the hexameric molecule is a triangular prism

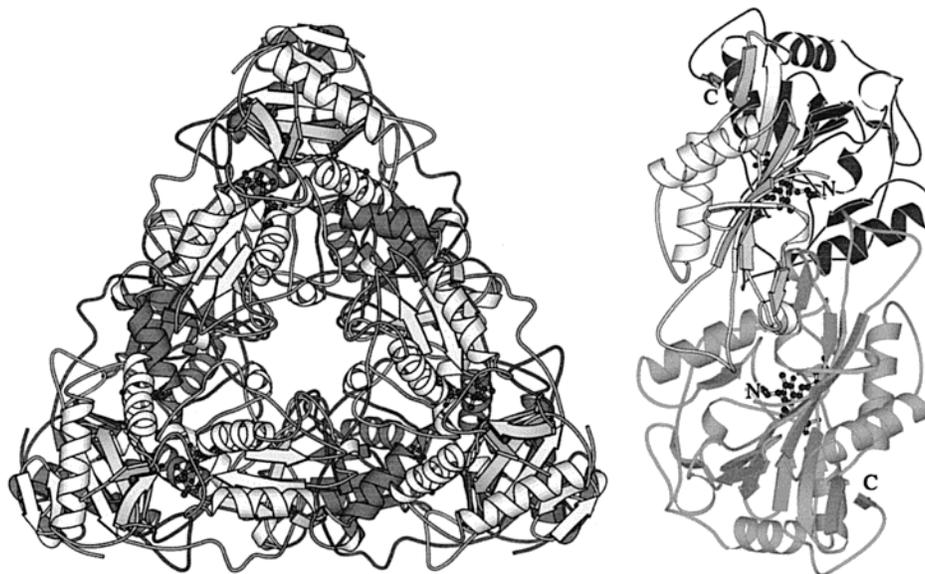


FIGURE 1: (a, left) C $\alpha$  tracing of the BCAT in the complex with 4-MeVA viewed down a molecular 3-fold axis. The dimer units located at the three corners of a triangular prism are shown by the lightly shaded and full ribbons. (b, right) The dimer unit viewed from a 2-fold axis. One subunit is represented by the shaded ribbon, and the small and large domains of the other subunit by lightly shaded and full ribbons, respectively. The cofactor, shown by the ball-and-stick model, is bound to the active-site pocket, which is located at the domain interface and the subunit interface.

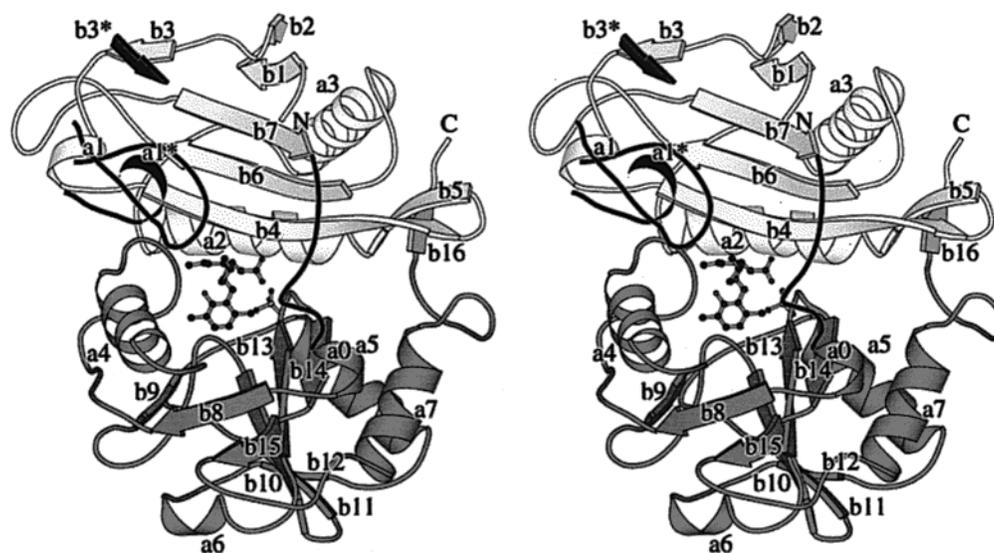


FIGURE 2: Stereoview of the subunit structure in eBCAT in the complex with 4-MeVA with secondary structure assignment.  $\alpha$ -Helices are denoted by a0–a7, and  $\beta$ -sheets by b1–b7 in the small domain and b8–b15 in the large domain. The shaded and open ribbons represent the large and the small domains, respectively. The interdomain loop and the loops from the other subunit of the dimer unit are shown by the full ribbons. The  $\beta$ -strand (b3\*) and  $\alpha$ -helix (a1\*) of the other subunit related by a molecular 2-fold axis are shown by the full ribbon. Figure prepared with MOLSCRIPT (26).

with  $D_3$  symmetry. Each of the three rectangular planes of the prism has two active sites, which remain 28 Å apart and face the solvent region. One subunit in the molecule interacts with the other five subunits and the surface areas of the subunit interfaces are 5482, 2242, 2002, 1823, and 910 Å<sup>2</sup> (27). The largest among these areas is found between the two subunits related by a 2-fold axis, indicating that the hexamer may be taken as the assembly of three dimer units (a dimer of dimers) around a 3-fold axis. Three of the remaining four subunit interfaces are distant from the active site, and the remaining one is located at the backside of the active-site cavity, suggesting that these four interfaces are not essential for the catalytic action. As was expected, the folding of this dimer unit is quite different from those of

the AspATs of fold type I (5–12), but is substantially the same as those of bsDAAT (14) and eADCL (15) of fold type IV as dimeric molecules. The dimer unit is essential for the enzyme function, because the active site is composed of the residues from both subunits of the dimer.

The subunit structure of eBCAT·4-MeVA is shown in Figure 2. The subunit is divided into the small domain (N-terminal to Tyr124 and Leu303 to Gln308) characterized by the  $\alpha/\beta$  structure, the interdomain loop (Pro125 to Glu136), and the large domain (Gln137 to Trp302) characterized by the pseudo barrel structure. The coenzyme PLP and 4-MeVA are bound to the cavity between the small and the large domains of one subunit, and two loops from the small domain of the other subunit approach the PLP. Thus, the active site

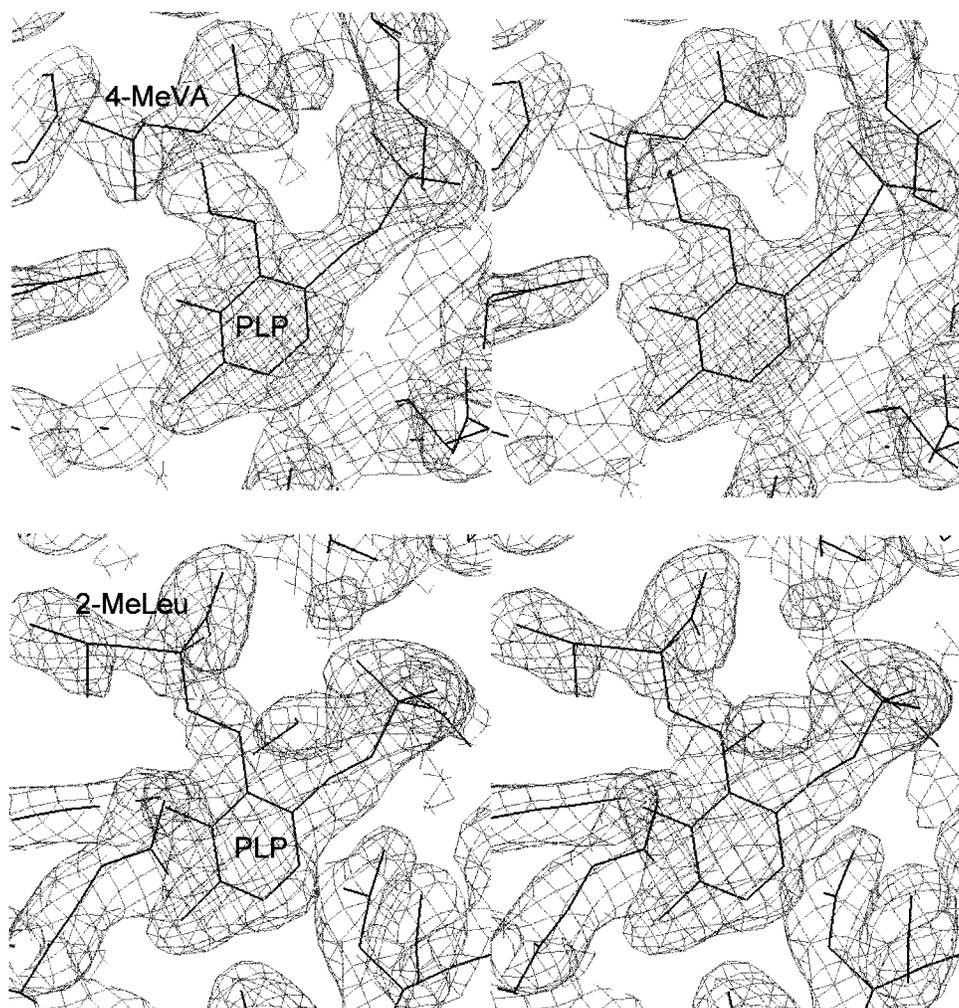


FIGURE 3: Stereoview of the  $2F_o - F_c$  electron density maps calculated using data between 40 and 2.4 Å resolution for PLP and residues close to PLP in eBCAT·4-MeVA (a, top) and eBCAT·2-MeLeu (b, bottom).

is comprised of residues from both domains of one subunit and the small domain of the other subunit. The large domain includes the catalytic residue, Lys159, to which the cofactor PLP molecule is bound through an aldimine double bond (internal aldimine bond).

**Active Site of the Unliganded eBCAT.** The stereo structure and hydrogen-bonding scheme of the active site are shown in Figures 4a and 5a, respectively. The residues comprising the active-site pocket are made up of three parts. The first part is the bottom of the active-site pocket, which consists of residues Phe36, Glu37, and Gly38 (from  $\beta$ -strand c), Arg59 (from  $\alpha$ -helix H2), Lys159 (from the loop between  $\beta$ -strand g and  $\alpha$ -helix H4), Glu193 (from  $\beta$ -strand i), Gly196 (from the loop between  $\beta$ -strands i and j), and Leu217 (from the loop between  $\beta$ -strand k and  $\alpha$ -helix H5). The second part forms the side or wall of the active-site pocket. The active-site wall consists of residues Arg40 (from  $\beta$ -strand c), Tyr95 and Arg97 (from  $\beta$ -strand e), Tyr164 (from  $\alpha$ -helix H4), Thr257 and Ala258 (from the  $\beta$ -turn connecting  $\beta$ -strands m and n), Tyr31\* (from  $\alpha$ -helix H1), and Val109\* (from the loop between  $\beta$ -strands e and f). The third part forms the lid of the active-site pocket. The lid is formed by Trp126 and Tyr129 (from the interdomain loop). Gly127–Tyr129 of the interdomain loop are not visible in the unliganded eBCAT as described above. Five water molecules (W1, W2, W3, W4, and W5) are located in the active site

and are involved in the hydrogen bond networks as shown in Figure 5a.

PLP is bound to the active-site pocket by noncovalent interactions with the residues forming the bottom of the active-site pocket except for Tyr164 and by forming an internal aldimine bond (Schiff base linkage) with the catalytic residue Lys159. The PLP has the same orientation as those found in bsDAAT and eADCL of fold type IV (14, 15) and rotates by 180° compared with the PLP-dependent enzymes of the other fold types (5–12, 28–31). The *re*-face of the PLP ring faces the protein side, and the catalytic residue Lys159 transfers protons on the *re*-face of PLP. Thus the *re*-face specificity of eBCAT was unambiguously established. The pyridine ring of PLP is sandwiched between the main-chain atoms of Gly196 from the side of the *si*-face and the side-chain of Leu217 from the side of the *re*-face. Similarly, PLP is sandwiched between Leu217 and Ser196 in bsDAAT and by Val217 and Ala196 in eADCL.

A five-link hydrogen bonding chain is formed by O3 of PLP, the hydroxyl group of Tyr164, W2, the main-chain C=O of Gly196, W3, and the main-chain NH of Ala259. Glu193 forms a salt bridge with the protonated nitrogen atom of the pyridine ring of PLP. The hydroxy group of Tyr31\*, the guanidino group of Arg97, the hydroxy group of Tyr95, and W1 make a hydrogen bond network in the form of a half

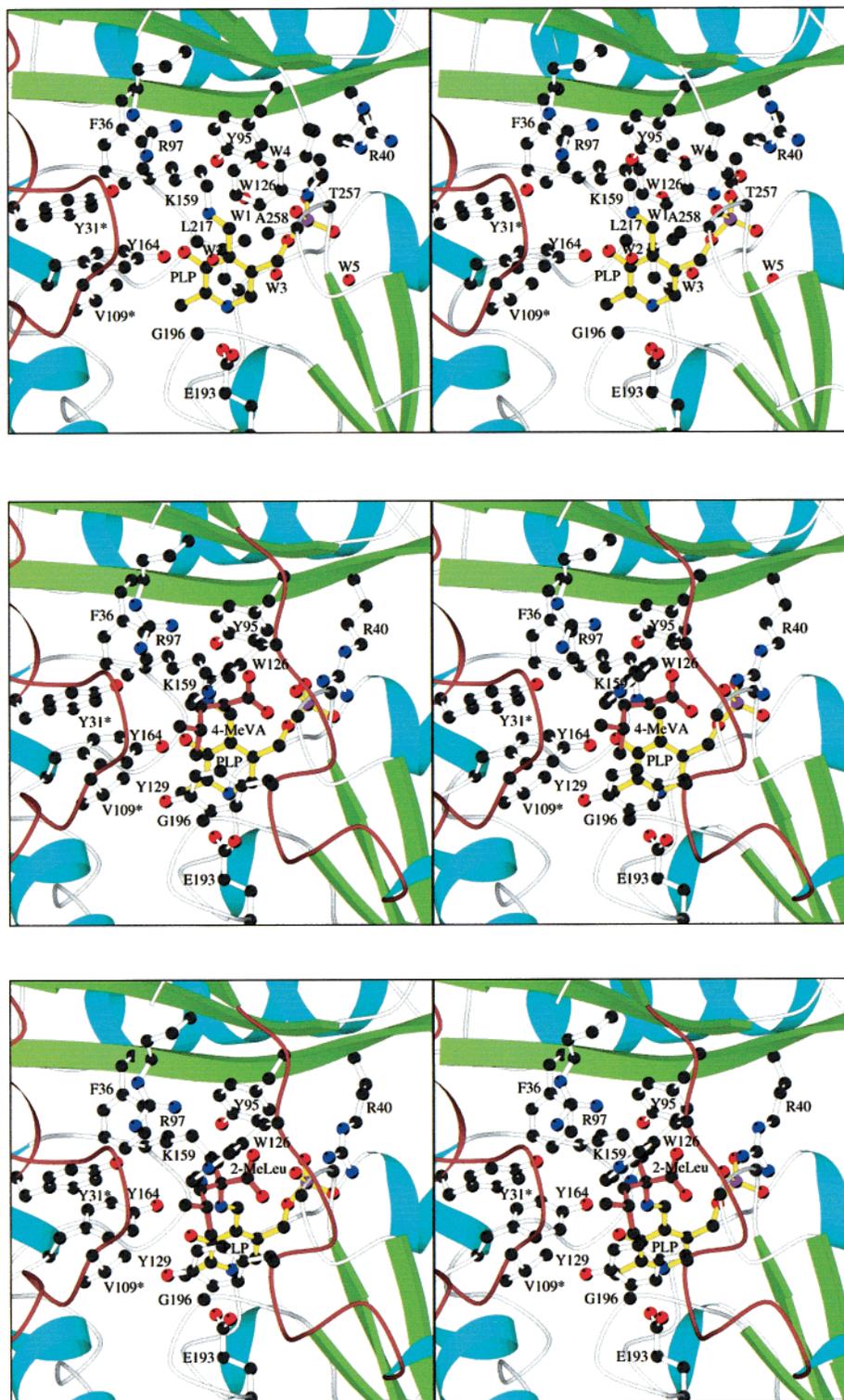


FIGURE 4: Stereoview of the active site in eBCAT. The front and the back of each figure are the solvent side (entrance of the active site) and the protein side (bottom of the active site), respectively. Figure prepared with MOLSCRIPT (26). (a, top) A close-up view of the active site of the unliganded eBCAT. Two loops (red) carrying Y31\* and V109\* of the small domain of the other subunit participate in the formation of the active site. The active-site residues and PLP (yellow) are shown by a ball-and-stick model. The interdomain loop are missing because the region is disordered. (b, middle) A close-up view of the active site of eBCAT in the complex with 4-MeVA. The missing interdomain loop in the unliganded form showed its ordered structure (red) on the active site and is displayed with W126 and Y129. The 4-MeVA is colored red with the cofactor yellow. Leu217, Thr257, Ala258, W3, W4, and W5 are omitted for clarity. (c, bottom) A close-up view of the active site of the eBCAT in the complex with 2-MeLeu. The 2-MeLeu is colored by red with the cofactor in yellow. Leu217, Thr257, Ala258, W3, W4, and W5 are omitted for clarity.

circle on the side of the *si*-face of the PLP plane. Arg40, which plays an important role in the substrate binding followed by the active-site closure, directs its side-chain

toward the outside of the active site. W1, W2, and W3 form a triangle with a side of about 4 Å and seem to be involved in weak interactions with one another.

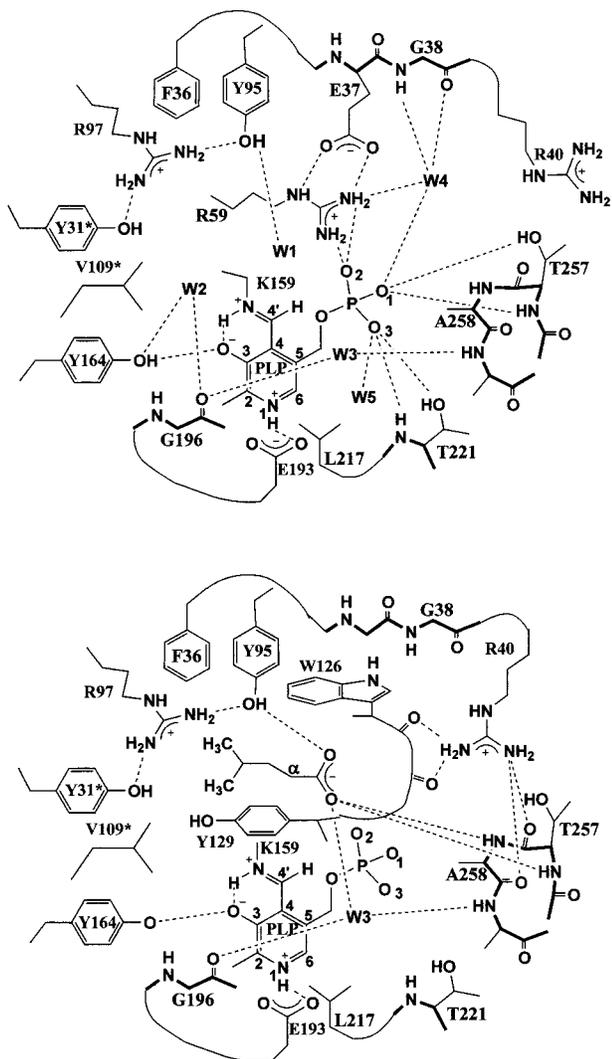


FIGURE 5: Schematic diagram showing hydrogen bond and salt bridge interactions of the active-site residues in the unliganded eBCAT (a, top) and eBCAT·4-MeVA (b, bottom). Putative interactions are shown by dotted lines if the acceptor and donor are less than 3.3 Å apart. Ile220 whose main-chain NH group interacts with OP2 of the phosphate in PLP is not shown. In eBCAT·4-MeVA, W3, W4, W5, and the hydrogen bonds associated with the phosphate group of PLP are omitted for clarity, since the hydrogen bond interactions around the phosphate group are the same between the unliganded eBCAT and eBCAT·4-MeVA.

Glu37, Gly38, Arg59, Ile220 (not shown in Figure 5), Thr221, Thr257, W4, and W5 are involved in the intricate hydrogen bonds to form a rigid base to fix the phosphate moiety of PLP (Figure 5a). The carboxylate of Glu37 interacts with the guanidino group of Arg59 by the side-on method (32). The guanidino group of Arg59 makes a salt bridge with OP2 of the phosphate moiety, which is further hydrogen bonded to the main-chain NH of Ile220. The main-chain NH and the hydroxy group of Thr257 form hydrogen bonds with OP1 of the phosphate. W4 connects OP1, the guanidino group of Arg59, and the main-chain NH and C=O groups of Gly38. The OP3 of the phosphate moiety is hydrogen bonded to the main-chain NH and hydroxy groups of Thr221 and W5. W5 is further hydrogen bonded to Arg222, Asn196, and Ser255. The negative charge of the phosphate group is largely compensated by the plus charge of Arg59 and the dipole of helix H5 (residues 220–231), whose N-terminus is close to the phosphate group.

The torsional angle of C3–C4–C4'–N in the PLP–Schiff base bond structure is  $-16^\circ$ , indicating that the internal aldimine bond (Schiff base, C4'=N) is essentially coplanar with the PLP ring to make a conjugated  $\pi$ -system of the PLP–Schiff base, although the C4'=N bond slightly deviates from the PLP plane toward the protein side. Thus, the nitrogen atom of the Schiff base is reasonably protonated to form an intramolecular hydrogen bond with O3 of PLP, suggesting the high  $pK_a$  value of the Schiff base. This is consistent with the result that eBCAT showed no detectable pH dependent spectral change in the range of pH 5.5–8.5 (1). In the unliganded PLP forms of bsDAAT and eADCL, the corresponding torsional angles are  $-50$  and  $-34^\circ$ , respectively, implying that the C4'=N significantly deviates from the PLP plane toward the protein side, and the coplanarity of the C4'=N bond with the PLP ring observed in eBCAT does not seem to be a common feature of the enzymes of the fold type IV.

*Active Site of the eBCAT·4-MeVA as Michaelis Complex Model.* The  $2F_o - F_c$  electron density map for PLP and the residues close to PLP in eBCAT·4-MeVA is shown in Figure 3a. The stereo structure and hydrogen-bonding scheme of the active site are shown in Figures 4b and 5b, respectively. The 4-MeVA is the substrate analogue where the  $\alpha$ -amino group of L-leucine is replaced by a hydrogen atom. The binding of 4-MeVA liberates two water molecules (W1 and W2) from the substrate binding region of the active site. The interdomain loop showed its ordered structure as the lid of the active site, and shield 4-MeVA from the solvent region. Upon binding with 4-MeVA, the active-site residues except for Arg40 and the residues from the interdomain loop do not significantly change their positions and retain the interactions among them, since the corresponding residues of the active sites between the unliganded eBCAT and eBCAT·4-MeVA are superimposed within an rms deviation of 0.21 Å with a maximum displacement of 0.70 Å, when Arg40 is neglected.

The 4-MeVA as a substrate analogue is located on the *si*-face of PLP with one of the C $\alpha$  hydrogen atoms of 4-MeVA directed toward C4' of the Schiff base in PLP. The carboxylate of 4-MeVA corresponding to the  $\alpha$ -carboxylate of the substrate occupies the space formed by the liberated W1. One of the carboxylate oxygen atoms of 4-MeVA makes a hydrogen bond with the hydroxyl group of Tyr95, and the other is involved in three hydrogen bonds with the main-chain NH groups of Thr257 and Ala258 at the  $\beta$ -turn, and W3. The isopropyl group of 4-MeVA makes van der Waals contacts with Tyr164, Val109\*, Tyr31\* and Phe36, and Trp126 and Tyr129 of the interdomain loop. The conformation of the PLP–Schiff base  $\pi$ -system is essentially the same between the unliganded form and eBCAT·4-MeVA, although the pyridine ring of PLP rotates by  $6^\circ$  around the C3–C4 axis to make the C4' atom approach the protein side, and the C3–C4–C4'–N dihedral angle changes from  $-16^\circ$  in the unliganded form to  $4^\circ$  in eBCAT·4-MeVA.

*Active Site of the eBCAT·2-MeLeu Complex as External Aldimine Model.* The  $2F_o - F_c$  electron density map for PLP and the residues close to PLP in eBCAT·2-MeLeu is shown in Figure 3b. The stereo structure of the active site are shown in Figure 4c. The 2-MeLeu, where the  $\alpha$ -hydrogen of L-leucine is replaced by a methyl group, binds to the active-site pocket, displacing Lys159 to make a new Schiff base

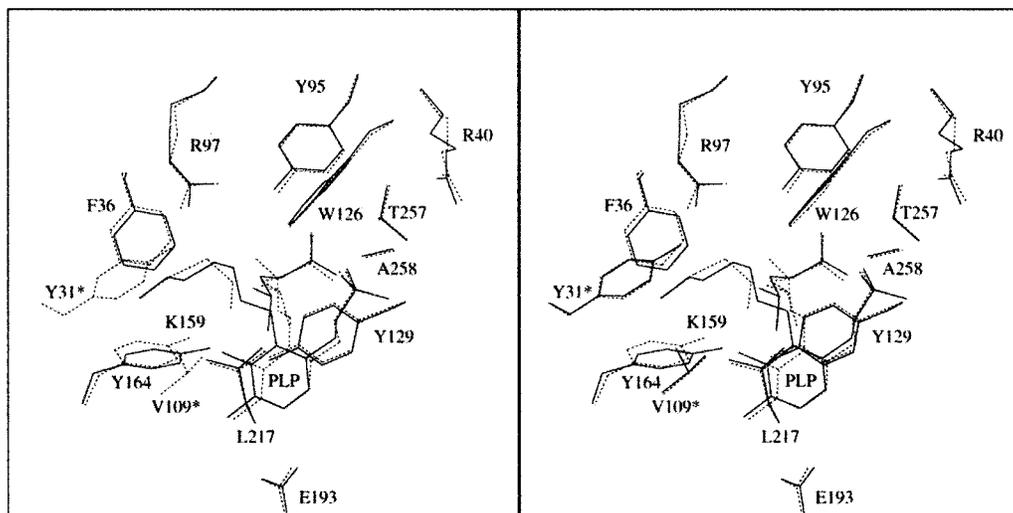


FIGURE 6: Superimposition of the active-site residues between eBCAT·4-MeVA and eBCAT·2-MeLeu by least-squares fitting of all atoms except for Lys159 and Tyr164. The active-site residues are well overlapped except for Tyr164. Figure prepared with MOLSCRIPT (26).

with PLP. The released amino group of Lys159, which is located on the *re*-face side of PLP, is hydrogen bonded to the hydroxyl group of Tyr164. The corresponding residues of the active sites between eBCAT·4-MeVA and eBCAT·2-MeLeu are superimposed within an rms deviation of 0.21 Å with a maximum displacement of 1.01 Å, when Lys159 and Tyr164 are neglected. Thus, the arrangement of the active-site residues in eBCAT·2-MeLeu is the same as that observed in eBCAT·4-MeVA except Lys159 and Tyr164 as is shown in Figure 6. All the water molecules (W3, W4, and W5) located in the active site of eBCAT·4-MeVA were also conserved in eBCAT·2-MeLeu.

The  $\alpha$ -carboxylate, C $\alpha$ , and C $\beta$  atoms of 2-MeLeu occupy the same positions as the corresponding ones of 4-MeVA. The isopropyl group of 2-MeLeu moves slightly toward the solvent side compared with that of 4-MeVA. The pyridine ring of PLP rotates by about 30° around the N1–C6 bond toward the solvent side. The Schiff base linkage (C4′=N) is significantly deviated from the pyridine ring plane of PLP with the torsional angle of N–C4′–C4–C3 = 30°. The rotation of the PLP ring and C4′=N around C4–C4′ produces an external aldimine bond with 2-MeLeu which is fixed at nearly the same position as that of 4-MeVA. The rotation of the PLP ring approaches O3 of PLP to the hydroxy group of Tyr164 within the distance of 1.7 Å to induce a fairly large movement of the side-chain of Tyr164 to avoid this access. The side-chain of Tyr164 and the O3 of PLP moves toward the protein side and the solvent side, respectively, to maintain the hydrogen bond between them. As a result, the hydroxy group of Tyr164 acts as an acceptor of the hydrogen bond with the released amino group of Lys159. In the aspartate and ornithine aminotransferases (5–12, 33, 34), Tyr70 and Thr322, respectively, which are on the side of the phosphate group of PLP and hydrogen-bonded to the phosphate group, interact with the released amino group of the catalytic lysine.

The eBCAT and bsDAAT enzymes belonging to fold type IV are divergently evolved from the same ancestor to have substrate specificities for the L- and D-amino acids, respectively. Although many of the active-site residues are replaced except for Lys159, Glu193, Glu37, Arg59, Leu217, and Thr257, not only the main-chain folding of the active site

(2, 14) but also the orientation of the cofactor and the cofactor–substrate complex are conserved between the two enzymes. The eBCAT and bsDAAT show the same type of cofactor rotation toward the solvent side upon binding of a substrate, since the cofactors are well duplicated in positions in the unliganded and the complexed forms of these enzymes (35). The  $\alpha$ -carbon,  $\alpha$ -carboxylate carbon and  $\beta$ -carbon atoms of 2-MeVal are overlapped on the  $\alpha$ -carbon,  $\beta$ -carbon, and  $\alpha$ -carboxylate carbon atoms of D-alanine, respectively. Thus, the C $\alpha$ -H bond of a substrate has the same position and direction in both enzymes, although the  $\alpha$ -carboxylate and the side-chains of the substrate are reversed. All the residues (Arg59, Ile220, Thr221, and Thr257) and water molecules (W4 and W5) directly interacting with the phosphate group of PLP are conserved in eBCAT and bsDAAT, resulting in the complete conservation of the hydrogen bond networks around the phosphate group of the cofactor.

**Induced Fit.** The interactions between the interdomain loop and the active-site residues are shown in Figure 7. It is well-known that AspATs from *Escherichia coli* and higher animals show the conformational change from the open form to the closed form upon binding of a substrate (5–12). The small domain moves to close the active site and shield the substrate from the solvent region and this domain movement is considered to play important roles in the catalytic actions of the enzyme. As described above, eBCAT does not exhibit any domain movement but shows active-site closure by the induce fit of the interdomain loop linking the small and large domains. In this respect, there exist open and closed forms in eBCAT. The interdomain loop is missing on the electron density map in the unliganded eBCAT, since the mobile loop will fluctuate on the active-site cavity, and be exposed to the solvent region. The substrate binding triggers the ordering of the interdomain loop, and the flexible loop is unambiguously located on the entrance of the active site in eBCAT·4-MeVA and eBCAT·2-MeLeu.

The conformation of the interdomain loop in eBCAT·4-MeVA is the same as that in eBCAT·2-MeLeu. The first three residues, Pro125, Trp126, and Gly127, are in the extended form of the main-chain. Ala128, Tyr129, Leu130, and Gly131 then makes a type I  $\beta$ -turn. Finally, Gly131,

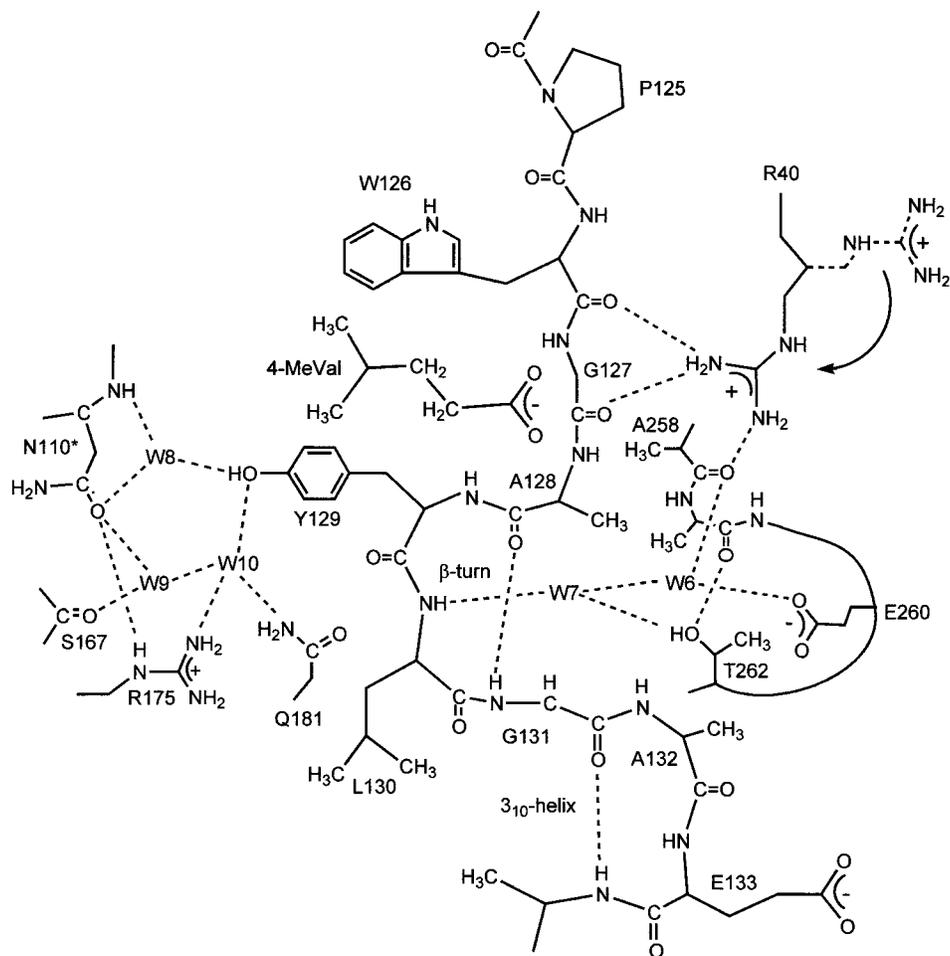


FIGURE 7: Schematic drawing of the interactions of the interdomain loop with the active-site residues, water molecules, and the substrate analogue 4-MeVA in eBCAT·4-MeVA. Putative interactions are shown by dotted lines if the acceptor and donor are less than 3.3 Å apart. All the interactions shown in this figure are conserved in eBCAT·2-MeVal.

Ala132, Glu133, and Ala134 make a further turn by the  $3_{10}$  helix followed by a short  $\alpha$ -helix. Gly131 is located at the junction between the  $\beta$ -turn and  $3_{10}$  helix, and Glu133 and Ala 134 are also members of a short  $\alpha$ -helix. Upon approach of the  $\alpha$ -carboxylate of a substrate, Arg40 changes its side-chain direction from the outside of the active site to the inside, and the guanidino group of Arg40 forms hydrogen bonds with the main-chain carbonyl groups of Thr257 and Ala258 at the  $\beta$ -turn (Figure 5b). Concurrently, the guanidino group is hydrogen bonded to the main-chain carbonyl groups of Trp126 and Gly127 of the interdomain loop, bridging the  $\beta$ -turn and the interdomain loop. The main-chain carbonyl group of Ala258 is involved in the one-dimensional hydrogen bond network starting from Ala258 and leading to the main-chain NH group of Leu130 through W6 and W7. W6 and W7 are further hydrogen bonded to the carboxylate group of Glu260 and to the hydroxy group of Thr262, respectively. The side-chain of Tyr129 continuously approaches to Asn110\*, Arg175, and Gln181 and interacts with these residues through water molecules (W8, W9, and W10). These water molecules (W6–W10) are introduced to eBCAT·4-MeVA and eBCAT·2-MeLeu and are involved in the intricate hydrogen bond network linking the interdomain loop and the active-site residues.

When the active-site structure of the unliganded form is compared with that of the liganded form, Arg40 is the only residue that significantly changes its location. In this respect,

the directional change of the side-chain in Arg40 might trigger the structural change to the closed form. The hydrogen bonds formed by the ordered interdomain loop and the access of Trp126 and Tyr129 to the substrate with van der Waals contact might play an important role in the stabilization of the closed form.

*Comparison of the Interdomain Loop between bsDAAT and eBCAT.* The residues (Trp or Val126, Gly127, Tyr129, and Gly131) of the interdomain loop in eBCAT are conserved in most of the BCAT enzymes, whose primary sequences have more than 30% sequence identities with eBCAT, suggesting that the interdomain loops undergo an induced fit and have similar interactions with the active-site residues and substrates in these BCATs. In bsDAAT, the interdomain loop does not change its conformation upon binding of the substrate. The loop is located on the verge of the active-site cavity both in the unliganded and the liganded forms. The primary sequence of the interdomain loop in eBCAT is compared with that in bsDAAT (15).

eBCAT: -Trp126-Gly127-Ala128-Tyr129-Leu130-Gly131-Ala132-Glu133-  
bsDAAT: -Asn118-Pro119-Arg120-Pro121- -Leu122-Glu123-

The loop in bsDAAT is characterized by two proline residues and the deletion of the residues corresponding to Leu130 and Gly131 in eBCAT, while that in eBCAT is characterized by two glycine residues (Gly127 and Gly131).

The shorter loop with two proline residues in bsDAAT makes the main-chain of the loop rigid. In addition, Arg120 forms salt bridges with two glutamates in the active-site region. In eBCAT, two glycine residues in the loop make the main-chain of the loop flexible. The flexibility of the loop and the moderate interactions of the loop with the active-site residues are important factors for the induced fit observed in eBCAT. The induced fit of the interdomain loop is mediated by the introduction of structure waters joining the loop and the active-site residues.

It is noteworthy that in the liganded bsDAAT, the *si*-face of the substrate–cofactor complex is exposed to the solvent region, although the *re*-face directed to the protein side is protected from the access of water molecules (35), indicating that the transamination reaction may proceed without the active-site closure. Presumably, the shielding of only the reaction side of the planar  $\pi$ -system of the substrate–cofactor complex is enough for the transamination reaction, and the active-site closure is employed for the strict recognition of the substrate. The wide spectrum of the substrate amino acids in bsDAAT and the narrow one in eBCAT and AspATs support this idea (14, 35). The closed forms observed in AspATs and eBCAT limit the shape and size of the active-site cavity to achieve the exact selection of the substrate.

**Substrate Recognition.** Upon the access of the substrate analogue 4-MeVA to the active site of eBCAT, its  $\alpha$ -carboxylate and side chain take the place of W1 and W2, respectively, liberating these two water molecules. Arg40, which directs its side chain toward the outside of the active site and forms hydrogen bonds with the main-chain carbonyl groups of Cys41 and Ala94, reminded us of Arg292 in AspAT (5–12). Arg292 changes its side-chain direction from the solvent side to the protein side to make a direct salt bridge with the  $\beta$ -carboxylate of the substrate aspartate. Thus, Arg40 of eBCAT was expected to make a salt bridge with the substrate  $\alpha$ -carboxylate, but the direct interaction was not observed between Arg40 and the  $\alpha$ -carboxylate of 4-MeVA. One of the  $\alpha$ -carboxylate oxygen atoms of 4-MeVA makes a hydrogen bond with the hydroxy group of Tyr95 and the other with the main-chain NH groups of Thr257 and Ala258 at the  $\beta$ -turn, and W3. Interestingly, Arg40 is indirectly involved in the recognition of the  $\alpha$ -carboxylate of the substrate. The side-chain of Arg40 loses hydrogen bonds with Cys41 and Ala94, approaches the one side of the  $\beta$ -turn at Thr257 and Ala258, and interacts with the main-chain C=O groups, thus polarizing the main-chain NH groups on the other side (Figure 5b). Additionally, Arg97 is involved in indirect recognition of the  $\alpha$ -carboxylate, which is mediated through Tyr95. The hydroxy group of Tyr95 is polarized by the hydrogen bond with the guanidino group of Arg97. The activated NH groups at the  $\beta$ -turn and the hydroxy group of Tyr95 form strong hydrogen bonds with the  $\alpha$ -carboxylate of 4-MeVA, and balance the negative charge of the  $\alpha$ -carboxylate of 4-MeVA. Consequently, Thr257 and Ala258 at the  $\beta$ -turn, Tyr95, W3, Arg40 and Arg97 form the unique recognition site for the  $\alpha$ -carboxylate of the substrate.

The isopropyl group of 4-MeVA corresponding to the side-chain of the substrate leucine binds to the hydrophobic cavity formed by Tyr164, Val109\*, Tyr31\*, and Phe36 in the unliganded enzyme. Upon binding of 4-MeVA, Trp126 and Tyr129 of the interdomain loop approach 4-MeVA to close the cavity and shield the isopropyl group from the solvent

region. The C $\alpha$  atom of 4-MeVA lies just above the C4' of PLP with the C2...C4' distance of 3.4 Å, and one of the C $\alpha$  hydrogen atoms corresponding to the substrate  $\alpha$ -amino group is directed toward the C4' of PLP, indicating that 4-MeVA and PLP are in a suitable geometry for the transaldimination reaction (from PLP-Lys159 to PLP-substrate) to proceed. The side-chain and  $\alpha$ -carboxylate of 2-MeLeu are recognized by the same mechanism as that observed in eBCAT•4-MeVA. The recognition sites for the  $\alpha$ -carboxylate and the isopropyl group of the substrate are crucial in determining the geometry of the bound substrate through the entire transamination reaction process.

**Conclusion.** Upon binding of a substrate, the interdomain loop shows its ordered structure, approaches the substrate, and shields it from the solvent region. The directional change in the side-chain in Arg40, which is induced by the long-range electrostatic interaction between the  $\alpha$ -carboxylate of the substrate and the guanidino group of Arg40, seems to trigger the disorder–order transition of the interdomain loop. In addition, the ordered structure of the loop may be stabilized by the dispersion force between the hydrophobic side-chain of the substrate and the interdomain loop, and the hydrogen bonds with water molecules. The enzyme prepares the unique binding site for the  $\alpha$ -carboxylate of the substrate. The  $\alpha$ -carboxylate interacts with three polar groups (two main-chain NH groups and one OH group) which are activated by two arginine residues. The hydrophobic side chain of the substrate binds the hydrophobic pocket constructed by the side chains of five aromatic residues and valine.

## REFERENCES

- Inoue, K., Kuramitsu, S., Aki, K., Watanabe, Y., Takagi, T., Nishigai, M., Ikai, A., and Kagamiyama, H. (1988) *J. Biochem.* 104, 777–784.
- Okada, K., Hirotsu, K., Sato, M., Hayashi, H., and Kagamiyama, H. (1997) *J. Biochem.* 121, 637–641.
- Grishin, N. V., Phillips, M. A., and Goldsmith, E. J. (1995) *Protein Sci.* 4, 1291–1304.
- Jansonius, J. N. (1998) *Curr. Opin. Struct. Biol.* 8, 759–769.
- Jansonius, J. N., and Vincent, M. G. (1987) *Structural basis for catalysis by aspartate aminotransferase in biological macromolecules and assemblies* (Jurnak, F. A., and McPherson, A., Eds.) pp 187–285, J. Wiley and Sons, New York.
- Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K., and Kagamiyama, H. (1998) *J. Mol. Biol.* 280, 443–461.
- McPhalen, C. A., Vincent, M. G., and Jansonius, J. N. (1992) *J. Mol. Biol.* 225, 495–517.
- Malashkevich, V. N., Strokopytov, B. V., Borisov, V. V., Dauter, Z., Wilson, K. S., and Torchinsky, Y. M. (1995) *J. Mol. Biol.* 247, 111–124.
- Okamoto, A., Higuchi, T., Hirotsu, K., Kuramitsu, S., and Kagamiyama, H. (1994) *J. Biochem.* 116, 95–107.
- Jäger, J., Moser, M., Sauder, U., and Jansonius, J. N. (1994) *J. Mol. Biol.* 239, 285–305.
- Miyahara, I., Hirotsu, K., Hayashi, H., and Kagamiyama, H. (1994) *J. Biochem.* 116, 1001–1012.
- Rhee, S., Silva, M. M., Hyde, C. C., Rogers, P. H., Metzler, C. M., Metzler, D. E., and Arnone, A. (1997) *J. Biol. Chem.* 272, 17293–17302.
- Yoshimura, T., Nishimura, K., Ito, J., Esaki, N., Kagamiyama, H., Manning, J. M., and Soda, K. (1993) *J. Am. Chem. Soc.* 115, 3897–3900.
- Sugio, S., Petsko, G. A., Manning, J. M., Soda, K., and Ringe, D. (1995) *Biochemistry*, 34, 9661–9669.

15. Nakai, T., Mizutani, H., Miyahara, I., Hirotsu, K., Takeda, S., Jhee, K.-H., Yoshimura, T., and Esaki, N., (2000) *J. Biochem.* 128, 29–38.
16. Otwinowski, Z. (1993) Data collection and processing. *Proceedings of the CCP4 Study Weekend*, pp 56–62, SERC Daresbury Laboratory, Warrington.
17. Navaza, J. (1994) *Acta Crystallogr., Sect. A* 50, 157–163.
18. Wang, B.-C. (1985) *Methods Enzymol.* 115, 90–112.
19. Zhang, K. Y. J., and Main, P. (1990) *Acta Crystallogr., Sect. A* 46, 377–381.
20. Bricogne, G. (1974) *Acta Crystallogr., Sect. A* 30, 395–405.
21. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr., Sect. D* 50, 760–763.
22. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A* 47, 110–119.
23. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science* 235, 458–460.
24. Brünger, A. T. (1991) *Annu. Rev. Phys. Chem.* 42, 197–223.
25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291.
26. Kraulis, J. (1991) *J. Appl. Crystallogr.* 24, 946–950.
27. Nicholls, A., Bharadawaj, R., and Honig, B. (1993) *Biophys. J.* 64, A166.
28. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Wilson Miles, E., and Davies, D. R. (1988) *J. Biol. Chem.* 263, 17857–17871.
29. Gallagher, D. T., Gilliland G. L., Xiao, G., Zondlo, J., Fisher, K. E., Chinchilla, D., and Eisenstein, E. (1998) *Structure* 6, 465–475.
30. Burkhard P, Rao GSJ, Hohenester, E., Schnackerz, K. D., Cook, P. F., and Jansonius, J. N. (1998) *J. Mol. Biol.* 283, 121–133.
31. Shaw, J. P., Petsko, G. A., and Ringe, D. (1997) *Biochemistry* 36, 1329–1342.
32. Mitchell, J. B., Thornton, J. M., Singh, J., and Price, S. L. (1992) *J. Mol. Biol.* 226, 251–262.
33. Shen, B. W., Hennig, M., Hohenester, E., Jansonius, J. N., and Schirmer, T. (1998) *J. Mol. Biol.* 277, 81–102.
34. Shah, S. A., Shen, B. W., and Brunger, A. T. (1997) *Structure* 5, 1067–1075.
35. Peisach, D., Chipman, D. M., Van Ophem, P. W., Manning, J. M., and Ringe, D. (1998) *Biochemistry* 37, 4958–4967.

BI010384L