Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the branched-chain aminotransferase from *Streptococcus mutans*

Jing Ruan,[a,b] Jia Hu,[c] Aihong Yin,[c] Wenqi Wu,[c] Xuzhen Cong,[a,b] Xueting Feng[a,b] and Shentao Li[a]*

[a]Department of Immunology, School of Basic Medical Sciences, Capital Medical University, Beijing 100069, People's Republic of China, [b]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Capital Medical University, Beijing 100069, People's Republic of China, and [c]Medical Experiment and Test Centre, Capital Medical University, Beijing 100069, People's Republic of China

Correspondence e-mail: lishentao@sina.com

Received 20 February 2012
Accepted 25 April 2012

PDB Reference: SmIlvE, 4dqn.

The branched-chain amino-acid aminotransferase from *Streptococcus mutans* (SmIlvE) was recombinantly expressed in *Escherichia coli* with high yield. An effective purification protocol was established. A bioactivity assay indicated that SmIlvE had aminotransferase activity. The specific activity of SmIlvE towards amino-acid substrates was found to be as follows (in descending order): Ile > Leu > Val > Trp > Gly. The protein was crystallized using the hanging-drop vapour-diffusion method with PEG 3350 as the primary precipitant. The structure of SmIlvE was solved at 1.97 Å resolution by the molecular-replacement method. Comparison with structures of homologous proteins enabled the identification of conserved structural elements that might play a role in substrate binding. Further work is needed to confirm the interaction between SmIlvE and its substrates by determining the structures of their complexes.

1. Introduction

Streptococcus mutans is a facultatively aerobic Gram-positive organism. As a member of the human oral flora, it has been implicated as a primary cause of dental caries (Loesche, 1986; Hamada & Slade, 1980) and as a key pathogen of infective endocarditis (Ullman *et al.*, 1988). To persist in the oral cavity, *S. mutans* must be able to tolerate rapid environmental fluctuations and exposure to various toxic chemicals (Biswas & Biswas, 2011) and to adapt and respond to environmental stresses. Generally, adaptive and stress-response mechanisms such as tolerance to acids, starvation, oxygen, fluoride and expression of urease can be regarded as having evolved to assist *S. mutans* to survive the stresses that are common to its habitats (Zhai *et al.*, 2011; Borden, 2000). The virulence factors of *S. mutans*, such as carbohydrate metabolism, tolerance and production of acids, synthesis of cell polysaccharides and adhesion ability, contribute to dental caries.

The branched-chain amino acids (BCAAs) leucine, isoleucine and valine are essential amino acids in the mammalian diet, but are biosynthesized in bacteria (Tremblay & Blanchard, 2009). Catabolism of BCAAs is a major pathway in the metabolism of essential amino acids and as a source of various direct and indirect metabolites such as glutamate (Eden & Benvenisty, 1999). In humans excess BCAAs are degraded, and metabolic disorders of BCAAs result in excess BCAAs or their derivatives, which are toxic to the central nervous system.

Although the three BCAAs have different final metabolic products, their metabolism involves two common steps (Hutson *et al.*, 2005). The first step in the catabolism of BCAAs is a reversible transamination to their respective

α-keto acids (Eden & Benvenisty, 1999). This step is catalyzed by branched-chain aminotransferases (BCATs; EC 2.6.1.42), which catalyze the transfer of an α-amino group from BCAAs to α-keto acid acceptors using pyridoxal 5′-phosphate (PLP) as a coenzyme (Madsen *et al.*, 2002; Schadewaldt *et al.*, 1995). In the above process, leucine, isoleucine and valine lose their α-amino groups and are converted to the respective branched-chain α-keto acids ketoisocaproate (KIC), ketomethylvalerate (KMV) and ketoisovalerate (KIV) (Venos *et al.*, 2004). The second step is oxidative decarboxylation of the branched-chain α-keto acids (BCKAs), which is catalyzed by branched-chain α-keto acid dehydrogenase (Soemitro *et al.*, 1989).

Branched-chain aminotransferases play an important role in the metabolism of BCAAs and other related amino acids such as glutamic acid. It has been found that mammals have two forms of the enzyme, mitochondrial and cytosolic, while bacteria contain only one form of the enzyme. Members of the branched-chain aminotransferases are widely distributed in all three forms of life, but differ in homology. The enzyme catalyzes the initial step in the metabolism of BCAAs using pyridoxal 5′-phosphate as a cofactor; the products of BCAA transamination are used for the generation of metabolic energy and carbon (Massey *et al.*, 1976), regulation of the $NADH/NAD^+$ ratio (Yvon *et al.*, 2000), recycling of Glu from α-ketoglutarate (Lapujade *et al.*, 1998), production of siderophores (Kingsley *et al.*, 1996), biosynthesis of pantothenic acid (Brown & Williamson, 1982) and production of branched-chain fatty acids for cell-membrane synthesis.

To date, the detailed crystal structure and characteristics of the branched-chain aminotransferase from *S. mutans* (SmIlvE) remain unknown. The PLP-dependent transaminases can be classified into five fold types, I–V, with BCATs belonging to type IV (Hirotsu *et al.*, 2005; Grishin *et al.*, 1995). To date, all observed type IV PLP transaminases consist of two domains with an interdomain loop (Goto *et al.*, 2003) and can be L- or D-amino-acid donor specific (Okada *et al.*, 2001). In the present work, we report the expression, purification, crystallization and crystal structure determination of SmIlvE. The information provided by our crystal structure will help to elucidate its function in BCAA metabolism and to predict the sites of interaction with PLP and BCAAs.

2. Materials and methods

2.1. Materials

Enzymes for recombinant DNA technology such as *Pfu* polymerase, T4 DNA ligase, *Bam*HI and *Xho*I were purchased from New England Biolabs. The PCR amplification kit was also obtained from New England Biolabs. DNA Quick Purify/Recover Kit and Plasmid Mini Kit were products of Omiga Co. *S. mutans* strain UA159 (ATCC 700610) was obtained from the American Type Culture Collection (ATCC).

2.2. Construction, expression and purification

The *S. mutans ilvE* gene, which encodes a mature protein of 341 amino acids, was PCR-amplified using genomic DNA of *S. mutans* strain UA159 (GenBank Accession No. AE014133) as a template. The PCR product was cloned into the prokaryotic expression vector pGEX-6P-1 with *Bam*HI and *Xho*I restriction-enzyme sites. The recombinant plasmid was transformed into *Escherichia coli* strain BL21 (DE3). Protein expression was induced by the addition of IPTG (0.1 m*M* final concentration) when the OD_{600} of the culture in LB medium reached approximately 0.8. The culture was allowed to grow for a further 12 h at 289 K. Bacterial cells were then harvested by centrifugation at 5000 rev min^{-1} for 10 min at 277 K. Following harvesting, the cell pellets were resuspended in ice-cold lysis buffer (1× PBS, 1 m*M* PMSF, 1 m*M* DTT) and homogenized by sonication on ice. Cell debris was removed completely by centrifugation of the lysate at 15 000 rev min^{-1} for 45 min at 277 K.

All of the following purification steps were performed at 289 K. The clear supernatant was applied onto a self-packaged GST-affinity column (3 ml glutathione-Sepharose 4B; GE Healthcare) pre-equilibrated with lysis buffer. The contaminant proteins were washed away with wash buffer (lysis buffer plus 200 m*M* NaCl). Cleavage of the GST tag was achieved with PreScission protease overnight at 277 K. The protein with an additional five-amino-acid tag (GPLGS) at its N-terminus was eluted with lysis buffer. The eluant was concentrated and desalted on a Sephadex G-25 (Pharmacia) column with buffer *A* (25 m*M* Tris–HCl pH 7.0). The desalted protein was further purified on a HiTrap Q HP (Pharmacia) column with buffer *A* and buffer *B* (25 m*M* Tris–HCl pH 7.0, 1 *M* NaCl). Further purification was achieved by gel-filtration chromatography on a Superdex 75 (Pharmacia) column with buffer *A*. The purified protein was concentrated using an Amicon Ultra Centrifugal Filter to about 15 mg ml^{-1} as measured using the Bradford protein assay (Bradford, 1976). The purified protein was analyzed by SDS–PAGE and by MALDI-TOF-TOF mass spectrometry.

2.3. MALDI-TOF-TOF mass-spectrometric analysis

Sample preparation for mass-spectrometric analysis was performed as described previously (Zhang *et al.*, 2010). MALDI-TOF-TOF analysis was performed on an ultrafleXtreme mass spectrometer (Bruker Daltonics, Bremen, Germany). 0.5 µl of digested sample was spotted onto an AnchorChip target (Bruker Daltonics). The sample droplet was dried at ambient temperature and coated with 0.5 µl matrix solution (5 mg ml^{-1} α-cyano-4-hydroxycinnamic acid in 50% acetonitrile and 0.1% TFA). The mass spectrometer was calibrated externally using the monoisotopic $[M + H]^+$ ions of peptide-calibration standards (Bruker Daltonics). Mass data acquisitions were piloted using *flexControl* v.3.3 (Bruker Daltonics). The positive MALDI-TOF and MS/MS LIFT spectra of selected ions were collected and peak lists were generated using the SNAP peak-detection algorithm, with a signal-to-noise threshold higher than 6 and with top-hat baseline subtraction in *flexAnalysis* v.3.3 (Bruker Daltonics). The MS spectra and MS/MS spectra were combined in *Biotools* (Bruker Daltonics) and searched for bacteria in the

NCBI nonredundant database using the *Mascot* search engine (Matrix Science; London, England) using the following parameters: monoisotopic; peptide mass tolerance ±100 p.p.m.; fragment mass tolerance ±0.5 Da; missed cleavages, 1.

Molecular-mass analysis was performed on the MALDI-TOF-TOF mass spectrometer in positive Lp mode with a salicylic acid (SA) matrix (Bruker Daltonics). Masses were calibrated to known peptide standards. A 0.5 µl aliquot of IlvE sample was spotted onto a MALDI sample plate, air-dried and coated with 0.5 µl matrix solution. After air-drying, the sample and matrix mixture was desalted with 1 µl 0.1% TFA on the sample plate. Mass data were acquired using *flexControl* v.3.3 (Bruker Daltonics). MS spectra were analyzed in *flexAnalysis* v.3.3 (Bruker Daltonics).

2.4. Bioactivity assay

SmIlvE activity was determined using a modification of the method of Collier & Kohlhaw (1972). Briefly, the reaction mixture contained 10 µg purified SmIlvE sample, 10 m*M* L-isoleucine (or the other L-amino acids), 10 m*M* 2-oxoglutarate, 10 m*M* PLP and 50 m*M* Tris–HCl buffer pH 8.0 in a final volume of 100 µl. Quantitative assays for SmIlvE activity with amino acids were performed at 276 K as described elsewhere (Gao *et al.*, 1997) using the colorimetric L-glutamic acid assay kit (Jiancheng, Nanjing, People's Republic of China). Specific activities were calculated as micromoles of L-glutamic acid formed per minute per microgram of protein using the molar extinction coefficient of NADH at 339 nm (6220 M^{-1} cm^{-1}).

2.5. Crystallization, data collection and processing

The purified protein was concentrated to about 15 mg ml^{-1}. The protein sample was centrifuged at 30 000*g* for 45 min to clarify the solution prior to crystal screening trials. Initial screening was performed at 291 K in 16-well plates by the hanging-drop vapour-diffusion method using sparse-matrix screening kits from Hampton Research (Crystal Screen, Crystal Screen 2, PEG/Ion, PEG/Ion 2, Crystal Screen Lite and Natrix), followed by refinement of the conditions by variation of the precipitant, pH, protein concentration and additives. Typically, 4 µl droplets prepared on siliconized cover slips by mixing 2 µl protein solution and 2 µl reservoir solution were equilibrated against 200 µl reservoir solution.

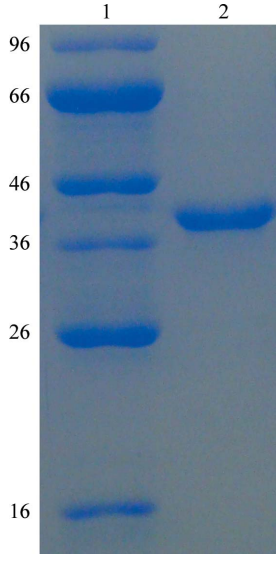

Figure 1
SDS–PAGE of purified SmIlvE. Lane 1, protein marker (labelled in kDa); lane 2, purified SmIlvE used for crystallization.

X-ray diffraction data sets were collected using synchrotron radiation with an ADSC Quantum 315 detector on beamline BL17U of Shanghai Synchrotron Radiation Facility (SSRF). Data were collected from a single crystal at a wavelength of 1.0000 Å to a maximum resolution of 1.97 Å. The crystal was immersed in a cryoprotectant solution (reservoir solution supplemented with 15% glycerol) for 5–10 s, picked up in a suitable nylon loop and flash-cooled in a nitrogen-gas stream at 100 K. The data were indexed and scaled using the *HKL-2000* program (Otwinowski & Minor, 1997).

2.6. Structure determination

The molecular-replacement method was employed to solve the phase problem using the *Phaser* program (McCoy *et al.*, 2007) with the crystal structure of *Mycobacterium tuberculosis* IlvE (MtIlvE; PDB entry 3ht5; 40% sequence identity; Tremblay & Blanchard, 2009) as the initial search model. Manual model building and refinement were performed with the programs *Coot* (Emsley & Cowtan, 2004) and *PHENIX* (Adams *et al.*, 2002). Solvent molecules were located from stereochemically reasonable peaks in the σ_A-weighted $2F_o - F_c$ difference Fourier electron-density map contoured at 1.2σ. Model geometry was verified using *MolProbity* (Chen *et al.*, 2010). Throughout refinement, the agreement between the model and the observed data was monitored by calculating R_{free} based on a subset containing 5% of all reflections. The stereochemical quality of the refined structures was checked by *MolProbity* (Chen *et al.*, 2010). Structural figures were drawn with *PyMOL* (DeLano, 2002).

3. Results and discussion

3.1. Expression and purification of SmIlvE

SmIlvE fused with an N-terminal GST tag was solubly expressed in *E. coli*. The fusion protein was cleaved with PreScission protease, producing the target protein with an additional five residues (GPLGS) at the N-terminus. After a series of purification steps, the purified SmIlvE was judged to be >95% pure on SDS–PAGE stained with Coomassie Brilliant Blue (Fig. 1). The precise sequence of the protein that was crystallized included residues 1–341 of SmIlvE and an additional five residues (GPLGS) at the N-terminus.

3.2. MALDI-TOF-TOF mass-spectrometric analysis

The molecular weight of the purified protein was 38.138 kDa as determined by MALDI-TOF-TOF (Fig. 2*a*), which was identical to its theoretical molecular weight. The MALDI-TOF-TOF MS spectrum of the digested protein is

Table 1
Aminotransferase activity of SmIlvE.

Substrate	Specific activity ± SD†
Isoleucine	104.60 ± 11.52
Leucine	86.84 ± 8.34
Valine	77.03 ± 5.85
Tryptophan	5.23 ± 1.15
Glycine	1.74 ± 0.92

† Specific activities are reported as micromoles of L-glutamic acid formed per minute per milligram of protein from four independent experiments conducted in duplicate. SD, sample standard deviation.

shown in Fig. 2(*b*). The MS spectra and the MS/MS spectra were analyzed in *Biotools* and were submitted to *Mascot*. Consequently, only SmIlvE was credibly obtained by *Mascot* searching. MALDI-TOF-TOF MS of the digested protein provided convincing evidence that the protein was SmIlvE.

3.3. Bioactivity assay

The bioactivity of SmIlvE in the catabolism of amino acids was determined using purified SmIlvE. The specific activity of SmIlvE towards amino-acid substrates was found to be as follows (in descending order): Ile > Leu > Val > Trp > Gly (Table 1). The relative activity of SmIlvE was analyzed with Ile arbitrarily defined as 100% relative activity (Fig. 3).

3.4. Crystallization and data collection

Thin hollow rhomboid crystals appeared after about a week from a condition consisting of 0.2 *M* sodium chloride, 100 m*M* HEPES pH 7.5, 25%(*w*/*v*) PEG 3350. The condition was further optimized by varying the precipitant, buffer pH, additives and protein concentration. Larger crystals with good diffraction quality were obtained from a reservoir solution comprising 80% of the reservoir solution above plus 20% 0.8 *M* succinic acid pH 7.0; these crystals were reproducible and were suitable for X-ray diffraction.

The crystal of SmIlvE had unit-cell parameters $a = b = 76.0$, $c = 119.49$ Å and belonged to space group $P3_121$. There is one SmIlvE molecule in the asymmetric unit, consistent with a Matthews coefficient of 2.6 Å^3 Da^{-1} and a solvent content of 53%. The diffraction pattern extended to a maximum of 1.97 Å resolution. Data-collection and processing statistics are summarized in Table 2.

Table 2
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Unit-cell parameters	
a (Å)	76.0
b (Å)	76.0
c (Å)	119.4
$\alpha = \beta$ (°)	90.0
γ (°)	120.0
Space group	$P3_121$
Wavelength (Å)	1.0000
Resolution (Å)	50.00–1.97 (2.01–1.97)
Total No. of reflections	371455 (13146)
No. of unique reflections	29016 (1565)
Completeness (%)	100.0 (100.0)
Average $I/\sigma(I)$	11.7 (4.0)
$R_{\rm merge}$† (%)	10.6 (51.0)
Refinement statistics	
No. of reflections used [$\sigma(F) > 0$]	29004
$R_{\rm work}$‡ (%)	20.4
$R_{\rm free}$‡ (%)	24.9
R.m.s.d. from ideal bond lengths (Å)	0.007
R.m.s.d. from ideal bond angles (°)	1.179
No. of atoms	2891
No. of protein atoms	2644
No. of ligand atoms	0
No. of solvent atoms	247
B value (Å^2)	31.2
Ramachandran plot, residues in	
Favoured regions (%)	98.2
Additional allowed regions (%)	1.4
Disallowed regions (%)	0.4

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean of the observations $I_i(hkl)$ of reflection hkl. ‡ $R_{\rm work} = \sum_{hkl} \big||F_{\rm obs}| - |F_{\rm calc}|\big| / \sum_{hkl} |F_{\rm obs}|$. $R_{\rm free}$ is the R factor for a selected subset (5%) of the reflections that were not included in prior refinement calculations.

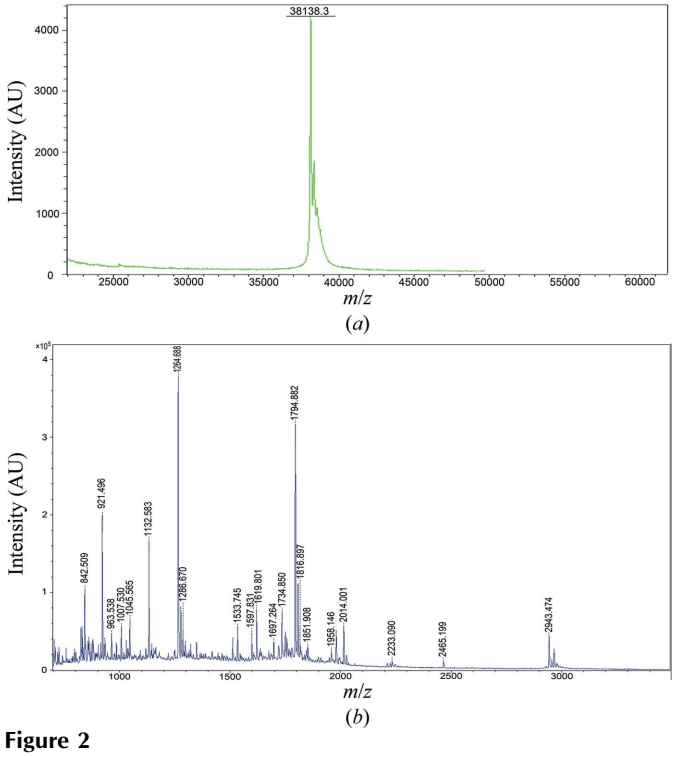

Figure 2
MALDI-TOF-TOF MS spectrum of SmIlvE. (a) Determination of molecular weight. (b) Peptide mass fingerprinting.

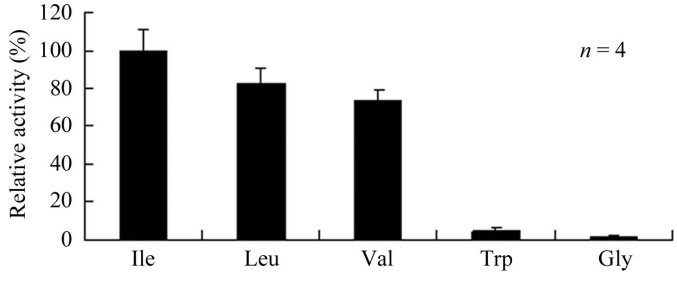

Figure 3
Relative activities of SmIlvE for Ile, Leu, Val, Trp and Gly. The activity of SmIlvE for Ile was arbitrarily defined as 100% relative activity.

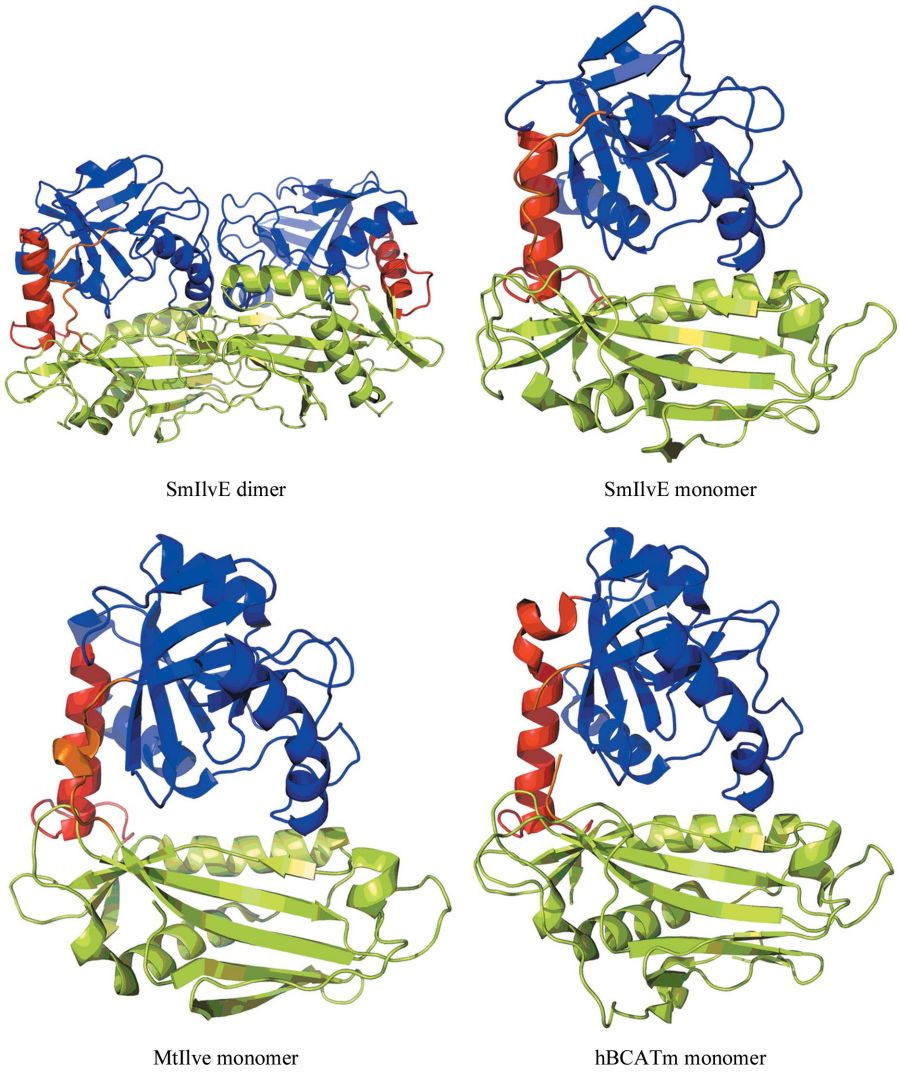

Figure 4
The two domains of SmIlvE and its two orthologues are shown in lime (domain 1) and blue (domain 2), with the interdomain loop and the interdomain helix coloured orange and red, respectively.

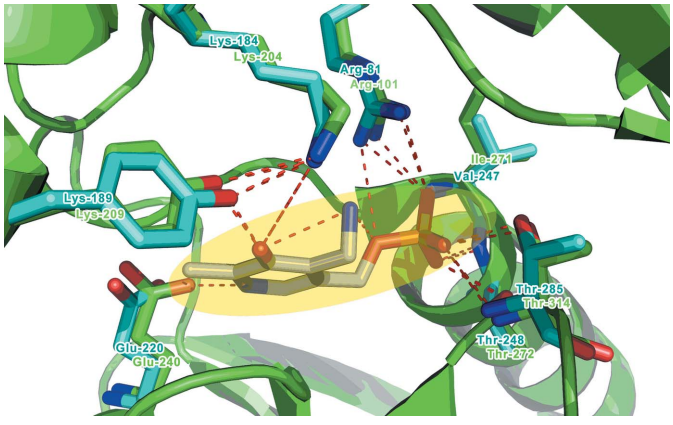

Figure 5
Key residues involved in interaction with PMP in the binding pocket of SmIlvE. MtIlvE is displayed in cyan and SmIlvE is shown in green. The PMP molecule is indicated.

3.5. Structure determination

The crystal structure of SmIlvE was solved by the molecular-replacement method with the crystal structure of MtIlvE as the initial search model in the program *Phaser* (McCoy *et al.*, 2007). The high rotation-function and translation-function *Z* scores (the rotation-function *Z* score is 8.1 and the translation-function *Z* score is 18.6) indicated the correct molecular-replacement solution.

The model was further improved by cycles of manual building and refinement using the programs *Coot* and *PHENIX*. The structure was subsequently refined to a final resolution of 1.97 Å with a working *R* value of 20.0% and a free *R* value of 24.9%. All residues in the polypeptide are well ordered, with the exception of the last two residues, which could not be built owing to a lack of interpretable electron density. The final model statistics are summarized in Table 2.

3.6. Structure description

The branched-chain amino-acid transaminases are known to function as dimers in solution. The gel-filtration results suggest that SmIlvE exists as a dimer, which is consistent with reports for other transaminases. Although there is only one SmIlvE molecule in the asymmetric unit, two SmIlvE molecules in adjacent asymmetric units form a functional dimer through the crystallographic twofold axis as indicated by the structural homologues (Fig. 4).

SmIlvE presents a globular overall structure with two separate domains (Fig. 5). The overall structure of SmIlvE belongs to the fold-type IV class of PLP-dependent enzymes (Yennawar *et al.*, 2002). The tertiary structure of SmIlvE has two domains arranged to form the active site, which is similar to other homologous proteins of the type IV class of PLP-dependent enzymes such as MtIlvE (Tremblay & Blanchard, 2009) and human mitochondrial branched-chain aminotransferase (hBCATm; PDB entry 1kt8; Yennawar *et al.*, 2002). Domain 1, which consists of residues 2–153 and 336–339, displays a Rossmann fold. It is composed of both N- and C-terminal residues and contains α-helices 1–3 and β-strands 1–7 and 18. The eight β-strands of this domain run in an up-and-down fashion to form a core antiparallel β-sheet with α-helices 1–3 on both sides.

Similar to the structure of MtIlvE, domain 2 of SmIlvE contains two β-sheets formed by β-strands 9–15 and by β-

strands 8, 16 and 17, which are surrounded by α-helices 4–6. The two domains are connected by α-helix 7 (residues 312–335) and a connecting loop (residues 154–165), corresponding to the interdomain motifs in MtIlvE. The interdomain helix is expected to provide a rigid framework for stabilization of the two interacting domains, while the interdomain loop is proposed to provide flexibility near the active-site entrance and to allow both substrate and product access into and out of the substrate-binding pocket (Tremblay & Blanchard, 2009; Yennawar *et al.*, 2002).

3.7. Comparison with structures of homologous proteins

Comparisons were made between the SmIlvE structure and homologous structures in the PDB. Two structures, MtIlvE and hBCATm (Tremblay & Blanchard, 2009; Yennawar *et al.*, 2002), were found by sequence alignment from the PDB using the program *BLAST*. Sequence identities are low, with only 40% identity between SmIlvE and MtIlvE and 31% between SmIlvE and hBCATm. In contrast, the sequence identity between hBCATm and MtIlvE is 39% (Fig. 6).

SmIlvE is a member of the type IV class of PLP-dependent enzymes. This class of enzymes contains both eukaryotic and prokaryotic members. Structure comparisons were performed in order to reveal the similarities between the structures of SmIlvE and other orthologues. The results for superposition of all C^{α} atoms (r.m.s.d. values) are 1.07 Å for MtIlvE (281 topologically equivalent C^{α} atoms) and 1.29 Å for hBCATm (260 topologically equivalent C^{α} atoms). The greatest deviations occur in loop regions. Collectively, these values indicate that the type IV class of PLP-dependent enzymes share very similar folds.

Structural alignment of SmIlvE and MtIlvE revealed significant structural conservation in the active sites of type IV class PLP-dependent enzymes. In the structure of the MtIlvE–PMP complex the substrate PMP molecule forms an extensive network of hydrogen bonds with MtIlvE as well as with nearby water molecules. The key residues involved in the interaction with PMP in the binding pocket of MtIlvE are Arg101, Lys204, Tyr209, Glu240, Ile271, Thr272 and Thr314, while the corresponding residues in SmIlvE whose side chains occupy equivalent positions to those in MtIlvE are Arg81, Lys184, Tyr189, Glu220, Val247, Thr248 and Thr285. This is not surprising, as a sequence alignment between MtIlvE, human orthologues and orthologues from several other bacterial pathogens showed a significantly high conservation of these seven residues. Therefore, these residues may be involved in the binding of substrates.

```
SmIlvE  MT---------------VDLD----WKNLGF-EYHKLPFRYISYYKDGKWDDGKLTEDAT 40
MtIlvE  MTSGSLQFTVLRAVNPATDAQRESMLREPGFGKYHTDHMVSIDYAEGRGWHNARVIPYGP 60
hBCATm  -ASSSFKAADLQLEMTQKPHKKPGPGEPLVFGKTFTDHMLMVEWN-DKGWGQPRIQPFQN 58
         :                  .     .   * : ..  :  :.:  .  * : ::

SmIlvE  LHISESSPALHYGQEAFEGLKAYRTKDGSVQLFRPNMNAERLQRTADRLLMPQVPTDKFI 100
MtIlvE  IELDPSAIVLHYAQEVFEGLKAYRWADGSIVSFRADANAARLRSSARRLAIPELPDAVFI 120
hBCATm  LTLHPASSSLHYSLQLFEGMKAFKGKDQQVRLFRPWLNMDRMLRSAMRLCLPSFDKLELL 118
        : :  ::  ***. : ***:**::  * .:  **.  *  *:  :* ** :*..    ::

SmIlvE  DAAKQVVRANEEYVPPYGTGATLYLRPLLIGVGDVIGVHPADEYIFTIFAMPVGNYFKGG 160
MtIlvE  ESLRQLIAVDKAWVPGAGGEEALYLRPFIFATEPGLGVRPATQYRYLLIASPAGAYFKGG 180
hBCATm  ECIRRLIEVDKDWVPDAAG-TSLYVRPVLIGNEPSLGVSQPRRALLFVILCPVGAYFPGG 177
        :. :::: .:: :**  .   :**:**.::.    :**  . .    ::  *.* ** **

SmIlvE  -LAPTNFLIQDDYDRAAPHGTGAAKVGGNYAASLLPGKVAHERQFSDVIYLDPATHTKIE 219
MtIlvE  -IAPVSVWVSTEYVRACPGGTGAAKFGGNYAASLLAQAEAAENGCDQVVWLDAVERRYIE 239
hBCATm  SVTPVSLLADPAFIRAWVGGVGNYKLGGNYGPTVLVQQEALKRGCEQVLWLYGPDHQLTE 237
         ::*...  .  : **   *.*  *.****..::*    * :.  .:*::*    :   *

SmIlvE  EVGSANFFGITKDN---EFITP-LSPSILPSVTKYSLLYLAEHRFGMKAIEGDVCVDELD 275
MtIlvE  EMGGMNIFFVLGSGGSARLVTPELSGSLLPGITRDSLLQLAID-AGFAVEERRIDIDEWQ 298
hBCATm  VGTMNIFVYWTHEDGVLELVTPPLNGVILPGVVRQSLLDMAQTWGEFRVVERTITMKQLL 297
              :.    ..   .::** *.  :**.:.: *** :*     : . *  : :.:

SmIlvE  KFVEAG------ACGTAAVISPIGGVQHGDDFHVFYSETEVGPVTHKLYDELTGIQFGDV 329
MtIlvE  KKAAAGEITEVFACGTAAVITPVARVRHG-ASEFRIADGQPGEVTMALRDTLTGIQRGTF 357
hBCATm  RALEEGRVREVFGSGTACQVCPVHRILYKDRNLHIPTMENGPELILRFQKELKEIQYG-- 355
        :    *      ..***. : *:  : :        :  :   :   : . *. ** *

SmIlvE  KAPEGWIYKVDD 341
MtIlvE  ADTHGWMARLG- 368
hBCATm  IRAHEWMFPV-- 365
          .. *:  :
```

Figure 6
Sequence alignment of SmIlvE with MtIlvE and hBCATm. The sequence alignment was made using the program *ClustalW* (Thompson *et al.*, 1994).

4. Conclusions

We have expressed the branched-chain aminotransferase from *S. mutans* (SmIlvE) in *E. coli* and purified it to a purity of >95%. Bioactivity assays indicated that SmIlvE has aminotransferase activity. The specific activity of SmIlvE towards amino-acid substrates was found to be (in descending order) Ile > Leu > Val > Trp > Gly. We also determined the crystal structure of SmIlvE to a maximum resolution of 1.97 Å. Comparison with structures of homologous proteins enabled the identification of conserved structural elements that might play a role in substrate binding. The three-dimensional structure of SmIlvE will help us to further understand its specific functions in caries and may provide a valid target for the rational design of anti-*S. mutans* drugs. Further work is needed to confirm the interaction between SmIlvE and its substrate by determining the structure of their complex.

The diffraction data sets were collected at Shanghai Synchrotron Radiation Facility (SSRF). This research was supported by the Basic Clinic Cooperation Project of Capital Medical University (11JL55). We are grateful to Professor Zhiyong Lou of Tsinghua University for his technical

assistance with data collection and processing, valuable comments and critical discussion.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.

Biswas, S. & Biswas, I. (2011). *Antimicrob. Agents Chemother.* **55**, 1460–1469.

Borden, G. H. W. (2000). *Microb. Ecol. Health Dis.* **12**, 138–148.

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.

Brown, G. M. & Williamson, J. M. (1982). *Adv. Enzymol. Relat. Areas Mol. Biol.* **53**, 345–381.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Collier, R. H. & Kohlhaw, G. (1972). *J. Bacteriol.* **112**, 365–371.

DeLano, W. (2002). *PyMOL*. http://www.pymol.org.

Eden, A. & Benvenisty, N. (1999). *FEBS Lett.* **457**, 255–261.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Gao, S., Oh, D. H., Broadbent, J. R., Johnson, M. E., Weimer, B. C. & Steele, J. L. (1997). *Lait*, **77**, 371–381.

Goto, M., Miyahara, I., Hayashi, H., Kagamiyama, H. & Hirotsu, K. (2003). *Biochemistry*, **42**, 3725–3733.

Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. (1995). *Protein Sci.* **4**, 1291–1304.

Hamada, S. & Slade, H. D. (1980). *Microbiol. Rev.* **44**, 331–384.

Hirotsu, K., Goto, M., Okamoto, A. & Miyahara, I. (2005). *Chem. Rec.* **5**, 160–172.

Hutson, S. M., Sweatt, A. J. & Lanoue, K. F. (2005). *J. Nutr.* **135**, 1557S–1564S.

Kingsley, R., Rabsch, W., Roberts, M., Reissbrodt, R. & Williams, P. H. (1996). *FEMS Microbiol. Lett.* **140**, 65–70.

Lapujade, P., Cocaign-Bousquet, M. & Loubiere, P. (1998). *Appl. Environ. Microbiol.* **64**, 2485–2489.

Loesche, W. J. (1986). *Microbiol. Rev.* **50**, 353–380.

Madsen, S. M., Beck, H. C., Ravn, P., Vrang, A., Hansen, A. M. & Israelsen, H. (2002). *Appl. Environ. Microbiol.* **68**, 4007–4014.

Massey, L. K., Sokatch, J. R. & Conrad, R. S. (1976). *Bacteriol. Rev.* **40**, 42–54.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Okada, K., Hirotsu, K., Hayashi, H. & Kagamiyama, H. (2001). *Biochemistry*, **40**, 7453–7463.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Schadewaldt, P., Hummel, W., Wendel, U. & Adelmeyer, F. (1995). *Anal. Biochem.* **230**, 199–204.

Soemitro, S., Block, K. P., Crowell, P. L. & Harper, A. E. (1989). *J. Nutr.* **119**, 1203–1212.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Tremblay, L. W. & Blanchard, J. S. (2009). *Acta Cryst.* F**65**, 1071–1077.

Ullman, R. F., Miller, S. J., Strampfer, M. J. & Cunha, B. A. (1988). *Heart Lung*, **17**, 209–212.

Venos, E. S., Knodel, M. H., Radford, C. L. & Berger, B. J. (2004). *BMC Microbiol.* **4**, 39.

Yennawar, N. H., Conway, M. E., Yennawar, H. P., Farber, G. K. & Hutson, S. M. (2002). *Biochemistry*, **41**, 11592–11601.

Yvon, M., Chambellon, E., Bolotin, A. & Roudot-Algaron, F. (2000). *Appl. Environ. Microbiol.* **66**, 571–577.

Zhai, F., Liu, X., Ruan, J., Li, J., Liu, Z., Hu, Y. & Li, S. (2011). *Acta Cryst.* F**67**, 287–290.

Zhang, L., Xiang, H., Gao, J., Hu, J., Miao, S., Wang, L., Deng, X. & Li, S. (2010). *Protein Expr. Purif.* **69**, 204–208.