

# Crystal structure and substrate specificity of the thermophilic serine:pyruvate aminotransferase from *Sulfolobus solfataricus*

Christopher Sayer,<sup>a</sup> Martin Bommer,<sup>b</sup> Michail Isupov,<sup>a</sup> John Ward<sup>b</sup> and Jennifer Littlechild<sup>a\*</sup>

<sup>a</sup>Henry Wellcome Building for Biocatalysis, Biosciences, College of Life and Environmental Sciences, University of Exeter, Stocker Road, Exeter EX4 4QD, England, and <sup>b</sup>Research Department of Structural and Molecular Biology, University College London, Darwin Building, Gower Street, London WC1E 6BT, England

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

The three-dimensional structure of the *Sulfolobus solfataricus* serine:pyruvate aminotransferase has been determined to 1.8 Å resolution. The structure of the protein is a homodimer that adopts the type I fold of pyridoxal 5'-phosphate (PLP)-dependent aminotransferases. The structure revealed the PLP cofactor covalently bound in the active site to the active-site lysine in the internal aldimine form. The structure of the *S. solfataricus* enzyme was also determined with an amino form of the cofactor pyridoxamine 5'-phosphate bound in the active site and in complex with gabaculine, an aminotransferase inhibitor. These structures showed the changes in the enzyme active site during the course of the catalytic reaction. A comparison of the structure of the *S. solfataricus* enzyme with that of the closely related alanine:glyoxylate aminotransferase has identified structural features that are proposed to be responsible for the differences in substrate specificity between the two enzymes. These results have been complemented by biochemical studies of the substrate specificity and thermostability of the *S. solfataricus* enzyme.

Received 22 December 2011  
Accepted 14 March 2012

**PDB References:** serine: pyruvate aminotransferase, holoenzyme, 3zrp; complex with PMP, 3zrq; complex with gabaculine, 3zrr.

## 1. Introduction

The aminotransferases (transaminases; EC 2.6.1.–) catalyse the transfer of an amino group from an amino acid to a keto acid (Mehta *et al.*, 1993). They use the cofactor pyridoxal 5'-phosphate (PLP), the biologically active form of vitamin B<sub>6</sub>, which is one of nature's most versatile cofactors (Braunstein & Shemyakin, 1953; Metzler *et al.*, 1954). The PLP normally covalently binds to an active-site lysine *via* a Schiff base (internal aldimine). The mechanism of aminotransferases is made up of two half-reactions. In the first half-reaction the donor substrate gives its amino group to the cofactor, resulting in a keto acid and enzyme-bound pyridoxamine 5'-phosphate (PMP). In the second half-reaction an amino group is transferred from PMP to an acceptor keto acid, producing an amino acid and restoring the PLP internal aldimine.

In most organisms, from bacteria (Umbarger *et al.*, 1963) to mammals (Ichiyama & Greenberg, 1957), serine biosynthesis proceeds *via* a phosphorylated pathway. The penultimate step is catalysed by a specialized phosphoserine:α-ketoglutarate aminotransferase (EC 2.6.1.52). A nonphosphorylated serine-degradation pathway also exists in animals (Walsh & Sallach, 1966) and plants (Liepman & Olsen, 2001) in which the amino group is added or removed by serine:pyruvate aminotransferase (EC 2.6.1.51; SPAT) or serine:glyoxylate aminotransferase (EC 2.6.1.45) (Sallach, 1956). Serine aminotransferases are also found in some bacteria (Izumi *et al.*, 1990). Proteins related to eukaryotic SPATs are found in a number of archaeal

genomes. The true physiological role of these proteins in archaea is unclear.

SPAT catalyses reversible amino-group transfer between serine and pyruvate to form alanine and hydroxypyruvate. In the first half-reaction of SPAT, serine is converted to hydroxypyruvate and the  $\alpha$ -amino group is transferred to the cofactor to generate PMP. In the second half-reaction pyruvate is aminated to form L-alanine and the PLP is regenerated. SPATs are the least characterized enzymes in group IV of the aminotransferases (Mehta *et al.*, 1993), which also contains alanine:glyoxylate aminotransferase and phosphoserine aminotransferase.

SPATs often have significant activity towards aromatic amino acids and methionine. Both rat and mouse liver SPATs prefer phenylalanine as a substrate and show only 33% and 17% relative activity towards serine (Noguchi, Minatogawa *et al.*, 1978). Human (Okuno *et al.*, 1980), dog, cat (Noguchi, Okuno *et al.*, 1978) and *Hyphomicrobium methylovorum* (Izumi *et al.*, 1990) SPATs are more specific for serine. Serine is a very poor substrate for aromatic amino-acid aminotransferases (Hayashi *et al.*, 1993) and this, together with the preference of SPATs for pyruvate over  $\alpha$ -ketoglutarate as an acceptor, distinguishes them from aromatic amino-acid aminotransferases.

The broad substrate range shown by SPATs makes them promising candidate enzymes for industrial biocatalysis, since reaction with a  $\beta$ -hydroxyl substrate is rare for aminotransferases. A two-enzyme pathway has been demonstrated for the enzymatic synthesis of the food-flavour precursor 6-deoxy-L-sorbose (Hecquet *et al.*, 1996) in which the carbon-carbon bond-forming enzyme transketolase (Takayama *et al.*, 1997) utilizes the hydroxypyruvate produced by spinach leaf serine:glyoxylate aminotransferase. As most bacteria do not have serine aminotransferases, the search for thermostable SPATs for biocatalytic applications has largely been confined to the archaea.

*Sulfolobus solfataricus* is a thermophilic archaeon isolated from hot volcanic springs and it grows optimally at 348 K and pH 3.0–4.5 (De Rosa *et al.*, 1975). Its internal pH is maintained at 6.5 by the efficient expulsion of protons (Moll & Schaefer, 1988). Its genome sequence (She *et al.*, 2001) shows a potential SPAT aminotransferase (gi:13815901), which is the subject of the present study and has 65% sequence identity to a protein from the related *S. tokodaii* and 49% sequence identity to the *Picrophilus torridus* DSM 9790 SPAT enzyme. The closest enzymes with known structures are the alanine:glyoxylate aminotransferases (AGATs) from yeast (Meyer *et al.*, 2005) and human (Zhang *et al.*, 2003), which share 36 and 30% sequence identity, respectively. These enzymes are involved in the degradation of glyoxylate, which is produced during the breakdown of sugars and amino acids (Rowell *et al.*, 1972).

This paper describes the characterization of the *S. solfataricus* SPAT protein and the structure of the recombinant enzyme determined in the holoenzyme (internal aldimine) form, in the PMP-bound form and as a complex with the inhibitor gabaculine. A structural comparison has been made with other group IV aminotransferase enzymes and this has

been used to explain the differences in substrate specificity between *S. solfataricus* SPAT and the closely related alanine:glyoxylate aminotransferase (EC 2.6.1.44; AGAT).

## 2. Methods

### 2.1. Expression and purification

The SPAT gene (NCBI gi:13815901) was amplified by PCR from *S. solfataricus* P2 genomic DNA and inserted into the pET-21a(+) vector (Novagen, Nottingham, England). The gene contains an internal *NdeI* site, which was removed by overlap extension PCR on the genomic DNA, leaving a new *NdeI* site at the N-terminus. The amplified product was cloned into pET-21a(+) between the *NdeI* and *XhoI* sites. An N-terminal hexahistidine tag containing the sequence MGH-HHHHH was then added by replacing the DNA between the *XbaI* and *NdeI* sites, forming the final expression plasmid pQR922. *Escherichia coli* BL21 (DE3) Rosetta2 [F<sup>−</sup> *ompT* *hdsS*<sub>B</sub>(r<sub>B</sub><sup>−</sup> m<sub>B</sub><sup>−</sup>) *gal dcm* (DE3) pRARE2\* (Cam<sup>R</sup>)] (Novagen) was used as the expression host.

*E. coli* cells harbouring the pET-21a vector containing the *S. solfataricus* SPAT gene were grown in LB medium containing 100  $\mu\text{g ml}^{-1}$  ampicillin at 303 K. When the OD<sub>600</sub> reached 0.8–1.0, isopropyl  $\beta$ -D-1-thiogalactopyranoside was added to 1 mM. After a 3 h induction at 303 K, the cells were harvested by centrifugation at 12 000g. The cell paste from 1 l culture was resuspended in 50 mM Tris-HCl pH 7.5 at a concentration of 10% (w/v). Sonication was carried out using a Soniprep 150 Sonicator (Sanyo). The lysed cell extract was incubated at 353 K for 20 min, followed by centrifugation at 12 000g to remove precipitated protein and cell debris.

The hexahistidine-tagged aminotransferase was purified on a HiLoad nickel column (Pharmacia, Uppsala, Sweden) using a linear gradient of 0–1 M imidazole in a buffer consisting of 50 mM Tris-HCl pH 7.5, 50  $\mu\text{M}$  PLP. The enzyme was further purified by gel filtration on a Superdex 200 gel-filtration column (Pharmacia, Uppsala, Sweden) using a buffer consisting of 50 mM Tris-HCl pH 7.5, 0.1 M NaCl, 50  $\mu\text{M}$  PLP.

### 2.2. Enzyme activity and stability assays

The hydroxypyruvate produced from the turnover of serine and either pyruvate, glyoxylate or  $\alpha$ -ketoglutarate was measured in a sequential assay. Initially, an aminotransferase reaction was carried out at 310 K in 50 mM Tris-HCl pH 7.3, 40  $\mu\text{M}$  PLP and 5 mM (substrate assay) or 10 mM (stability assay) of each substrate. A small sample (2–20  $\mu\text{l}$ ) was then added to 200  $\mu\text{l}$  of an assay solution consisting of 0.11 mg ml<sup>−1</sup> 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-tetrazolium (WST-1; Dojindo Laboratories, Kumamoto, Japan) and 0.1 M NaOH (final concentrations are given). The blue colour of the resulting WST-1 formazan was quantified by measuring its absorbance at 600 nm.

The catalytic pH optimum was determined by the above method using 50 mM sodium citrate (pH 2.5–6.6), potassium phosphate (pH 6.3–7.8) or sodium borate (pH 8.1–10.1) buffer in place of Tris-HCl.

For stability assays, the aminotransferase ( $0.4 \text{ mg ml}^{-1}$ ) was first incubated in 50 mM of the above citrate/phosphate/borate buffers, 100  $\mu\text{M}$  PLP, 2 mM pyruvate and 1 mg  $\text{ml}^{-1}$  bovine serum albumin at 323–262 K for 10 min before 1:4 dilution into 200 mM Tris–HCl buffer pH 7.3 for the aminotransferase assay, which was carried out at 310 K.

Turnover of all other amino-acid substrates was determined indirectly by measurement of the alanine produced from the amino acceptor pyruvate in a coupled assay. The reaction contained 100 mM potassium phosphate pH 8.0, 5 mM of the amino-acid substrate, 5 mM pyruvate, 40  $\mu\text{M}$  PLP, 15 mM NAD (pH adjusted to 8.0), 5  $\mu\text{M}$  1-methoxy-5-methylphenazine methylsulfate (Dojindo), 0.3 mM 2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-tetrazolium-5-carboxanilide (XTT; Biotium, Hayward, California, USA), 1 U  $\text{ml}^{-1}$  *Bacillus subtilis* alanine dehydrogenase (Sigma–Aldrich, Gillingham, England) and 0.01–0.1 mg  $\text{ml}^{-1}$  aminotransferase. The absorbance of the XTT-formazan produced was measured at 470 nm.

### 2.3. *S. solfataricus* SPAT crystallization

Initial crystallization was carried out with crystal screens from Molecular Dimensions by the microbatch method using an Oryx robot (Douglas Instruments). The volume of the droplet was 2  $\mu\text{l}$ , consisting of 1  $\mu\text{l}$  protein sample and 1  $\mu\text{l}$  screening condition. CocrySTALLIZATION of *S. solfataricus* SPAT was carried out in the presence of 10 mM gabaculine (5-amino-1,3-cyclohexadienyl carboxylic acid) and 20 mM phenylalanine, one of the amino-donor substrates. These cocrySTALLIZATIONS were performed using the same crystal screens.

The best holoenzyme crystals were grown using a final concentration of 5 mg  $\text{ml}^{-1}$  protein, 50 mM Tris–HCl buffer pH 8.5 and 15% PEG 400 at 293 K. The crystals of the PMP and gabaculine complexes were grown using different conditions of PEG in the crystal screens. The crystallization conditions of the crystal used to collect data for the gabaculine complex contained 200 mM calcium chloride.

### 2.4. Data collection and processing

Excess liquid was removed under silicone oil from crystals extracted directly from crystallization droplets for the holoenzyme and PMP complex before cooling. Glycerol was added to the microbatch droplet of the gabaculine-cocrySTALLIZED protein to 20% (v/v) before the crystal was frozen.

X-ray diffraction data were collected from the holoenzyme and the gabaculine–SPAT crystals on an in-house source at a temperature of 100 K using a MAR 345 (MAR Research) image plate mounted on a rotating-anode generator (Bruker AXS). The latter was operated at 100 mA and 35 kV (3.5 kW), producing Cu  $K\alpha$  radiation, which was collimated using XENOCs FOX2D CU25\_25P mirrors. Data for holoenzyme SPAT crystals were later re-collected on beamline 14.1 at the Daresbury Synchrotron at a wavelength of 1.488 Å. The PMP-complex data were collected on beamline I03 at the Diamond Synchrotron at a wavelength of 0.92 Å. In-house data were

processed using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 2001). Synchrotron data were processed with *MOSFLM* (Leslie, 1992).

### 2.5. Structure determination

The structure of the *S. solfataricus* SPAT holoenzyme was determined by the molecular-replacement method using the in-house *S. solfataricus* SPAT holoenzyme data and the *BALBES* program pipeline (Long *et al.*, 2008) running on the York Structural Biology Laboratory server. The best solution was found using the monomer of yeast AGAT (PDB entry 2bkw; Meyer *et al.*, 2005) as a model, which shares 37% sequence identity with *S. solfataricus* SPAT. The initial *R* factor was 55% and *R*<sub>free</sub> was 55% for the resulting molecular-replacement model and the model was refined to an *R* factor of 41% and an *R*<sub>free</sub> of 47%. The resulting model underwent cycles of isotropic *B*-factor crystallographic refinement using the program *REFMAC5* (Murshudov *et al.*, 2011) and manual model building performed in *Coot* (Emsley & Cowtan, 2004). Solvent molecules were added using *Coot*. The two subunits of each structure were rebuilt independently and refined using *REFMAC5* with local noncrystallographic symmetry (NCS) constraints. The van der Waals distances were restrained. The two complex structures were solved by molecular replacement using the coordinates of the holoenzyme structure.

The program *PyMOL* (DeLano, 2002) was used for preparation of the protein structure figures.

### 2.6. PDB accession codes

The atomic coordinates and structure factors have been deposited in the PDB with accession codes 3zrp, 3zrq and 3zrr for the holoenzyme, the PMP complex and the gabaculine complex, respectively.

## 3. Results and discussion

### 3.1. Enzyme substrate specificity

The highest activity for the putative *S. solfataricus* SPAT was observed with methionine, the aromatic amino acids phenylalanine and tryptophan and the substituted aromatic amino acids bromophenylalanine and iodotyrosine. The amino-donor specificity is therefore similar to that of aromatic amino-acid aminotransferases (Hayashi *et al.*, 1993; Onuffer *et al.*, 1995). The dicarboxylic amino acids aspartic acid and glutamic acid were not found to be substrates of this SPAT, while their amide analogues asparagine and glutamine were. The amino-donor substrate specificity for *S. solfataricus* SPAT is shown in Fig. 1.

The *S. solfataricus* enzyme showed a clear preference for both pyruvate and glyoxylate as amino acceptors in the presence of the donor serine (data not shown). No activity was detected towards  $\alpha$ -ketoglutarate, which is the most frequently used amino-group acceptor of aminotransferases, including aromatic amino-acid aminotransferases. The activity towards serine, which was not observed for aromatic acid aminotransferases (Hayashi *et al.*, 1993), and the absence of activity

**Table 1**

Summary of data-processing and refinement statistics.

Values in parentheses are for the outer resolution shell. The Wilson *B* factor was estimated by *SFCHECK* (Vaguine *et al.*, 1999). Ramachandran plot analysis was performed by *PROCHECK* (Laskowski *et al.*, 1993).

Crystal	Holoenzyme	Gabaculine complex	PMP complex
Space group	<i>C</i> 2	<i>C</i> 2	<i>C</i> 2
Unit-cell parameters (Å, °)	<i>a</i> = 162.4, <i>b</i> = 55.0, <i>c</i> = 101.8, $\beta$ = 111.3	<i>a</i> = 160.8, <i>b</i> = 55.2, <i>c</i> = 102.3, $\beta$ = 111.5	<i>a</i> = 160.1, <i>b</i> = 55.1, <i>c</i> = 101.6, $\beta$ = 111.0
Wavelength (Å)	1.488	1.54	0.92
Resolution range (Å)	43–1.82 (1.92–1.82)	25–2.00 (2.03–2.00)	56.5–1.82 (1.92–1.82)
Completeness (%)	98.2 (94.4)	95.9 (91.9)	99.7 (99.9)
Multiplicity	2.8 (2.5)	2.8 (2.4)	4.1 (4.1)
$\langle I/\sigma(I) \rangle$	9.6 (2.0)	10.4 (2.0)	9.6 (2.0)
$R_{\text{merge}}^{\dagger}$ (%)	5.5 (32.3)	8.9 (51.3)	8.0 (63.3)
$R_{\text{cryst}}^{\ddagger}$ (%)	21.2	18.9	18.7
$R_{\text{free}}$ (5% of total data) (%)	25.6	23.1	23.3
R.m.s.d. bond lengths (Å)	0.012 [0.022]	0.012 [0.022]	0.012 [0.022]
R.m.s.d. bond angles (°)	1.25 [1.99]	1.30 [1.99]	1.25 [1.99]
Wilson <i>B</i> factor (Å <sup>2</sup> )	39.7	47.1	37.9
Average <i>B</i> factor (Å <sup>2</sup> )			
Protein	38.8	42.9	38.0
Solvent	45.4	48.6	45.5
Ligand	29.7	39.4	24.1
Ramachandran analysis	94.4	93.1	92.5
(% of residues in most favoured regions)			

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the intensity of reflection  $hkl$ ,  $\sum_{hkl}$  is the sum over all reflections and  $\sum_i$  is the sum over  $i$  measurements of the reflection.  $\ddagger R_{\text{cryst}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$ . § Target values are given in square brackets.

towards  $\alpha$ -ketoglutarate allows the classification of the *S. solfataricus* enzyme as a serine:pyruvate aminotransferase.

The specific activity of *S. solfataricus* SPAT towards serine was much lower than that towards aromatic amino acids. This is similar to rat and mouse liver serine aminotransferases (Noguchi *et al.*, 1978). Some activity was also observed towards homophenylalanine and kynurenine and this is similar to that found for human SPAT, which has been shown to also be responsible for kynurenine:glyoxylate aminotransferase activity (Okuno *et al.*, 1980).

The D-amino acids tested and phospho-L-serine were not substrates of *S. solfataricus* SPAT, which is in line with the role of the enzyme as an L-serine aminotransferase. DL-3-Phenyl-3-aminopropionate, the  $\beta$ -amino-acid analogue of phenylalanine, was not a substrate. An absence of activity towards  $\beta$ -alanine has been reported for rat liver serine aminotransferase (Oda *et al.*, 1989). An amino alcohol (phenylalaninol) and an amine (*sec*-butylamine) were also found not to be substrates of *S. solfataricus* SPAT.

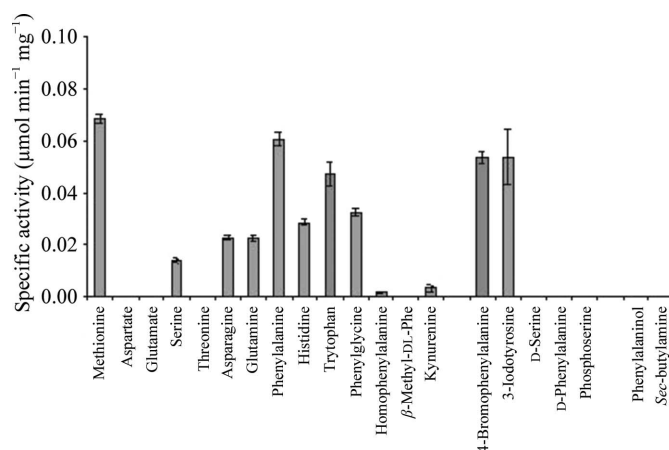
The optimal activity of *S. solfataricus* SPAT has been determined to be around pH 7.1 when assayed at 310 K using the serine:pyruvate reaction. At least 75% of the maximum activity is retained between pH 6.4 and 8.3. The enzyme was determined to be relatively thermostable. At physiological pH 6.5, *S. solfataricus* SPAT was stable for 10 min at 343 K, while half of the initial activity was lost at 353 K.

### 3.2. Quality of the models

The three *S. solfataricus* SPAT structures were refined to *R* factors of 21.2% to 1.82 Å resolution for the holoenzyme

structure, 18.9% to 2 Å resolution for the gabaculine complex and 18.7% to 1.82 Å resolution for the PMP complex using all data in the corresponding resolution ranges without a  $\sigma$  cutoff. These excluded 5.0% of the randomly distributed reflections assigned to calculate the  $R_{\text{free}}$  values, which were 25.6, 23.1 and 23.3%, respectively. The asymmetric unit of each structure contained a dimeric SPAT molecule. The electron density allowed the positioning of residues 6–382 in both subunits out of a total of 384 residues in the holoenzyme structure. Residues 5–383 were built for both subunits in the gabaculine-complex structure; in the PMP-complex structure the terminal Met1 and Arg384 were not built in both subunits. Each model also contained just over 500 water molecules, and the gabaculine-complex structure contained four calcium ions. The data-collection and model-refinement statistics of the three structures are listed in Table 1.

The dispersion precision indicator (Murshudov & Dodson, 1997) gives an overall estimate of the root-mean-square error in the coordinates of 0.14 Å for the holoenzyme structure, 0.13 Å for the PMP complex and 0.17 Å for the gabaculine complex for the well defined part of the structure. The overall *G*-factor is 0.1 for all three refined structures as calculated by the program *PROCHECK* (Laskowski *et al.*, 1993) and was used as a measure of the stereochemical quality of the model, which is better than average for the given resolutions. Pro10 is in the *cis* conformation in the PMP-complex structure. The holoenzyme structure and the gabaculine complex do not have any

**Figure 1**

The activity of *S. solfataricus* SPAT towards amino substrates (L-isomers unless stated otherwise). Pyruvate was used as the amino acceptor. The mean and standard deviation of at least five data points is shown. 5 mM (0.5 mM for homophenylalanine and 4-bromophenylalanine) of each amino-donor substrate and 5 mM pyruvate were used.

residues in the *cis* conformation. 92–94% of the nonglycine residues fall into the most favourable regions of the Ramachandran plot (Ramakrishnan & Ramachandran, 1965) in the three structures as defined by *PROCHECK*, with no residues in the disallowed or generously allowed regions. Approximately 41% of the amino acids are in  $\alpha$ -helices, 15% are in  $\beta$ -sheets and less than 2% are in  $3_{10}$ -helices.

### 3.3. Overall structure of *S. solfataricus* SPAT

The *S. solfataricus* SPAT monomer consists of two  $\alpha/\beta$  domains (Fig. 2): a large PLP-binding N-terminal domain (residues 2–262) and a smaller C-terminal domain (residues 263–381). The N-terminal domain folds into a typical  $\alpha/\beta/\alpha$  sandwich made up of a central seven-stranded  $\beta$ -sheet of mixed type with direction  $+-++++$  and topology  $+5x, +1x, -2x, -1x, -1x, -1$  in Richardson's classification (Richardson, 1981). The  $\beta$ -sheet is surrounded on both sides by ten  $\alpha$ -helices, with no helices between strands S7, S6 and S5. The loops between these strands are involved in cofactor binding. The smaller C-terminal domain consists of a four-stranded antiparallel  $\beta$ -sheet with  $+1, +2, -1$  topology and four  $\alpha$ -helices on the external side relative to the  $\beta$ -sheet.

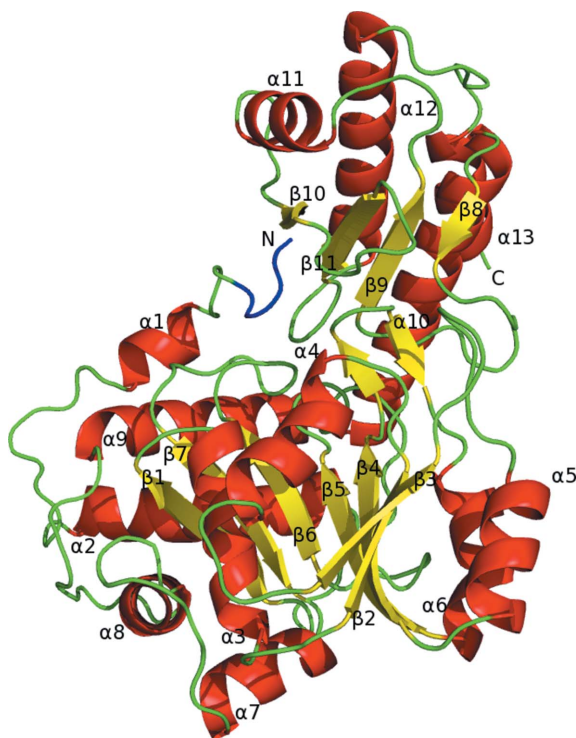
Size-exclusion chromatographic experiments showed elution of *S. solfataricus* SPAT at 90 kDa, suggesting the presence of dimers in solution, which is consistent with the oligomeric states of the closest structural homologues. The asymmetric unit of the crystal contained two tightly bound

subunits forming the functional homodimer. The formation of the dimer buries  $4017 \text{ \AA}^2$ , which amounts to 23% of the solvent-accessible surface area of each monomer being buried. Dimer formation results in the formation of 21 hydrogen bonds between the two monomers, mainly between the two N-terminal domains, and contacts between the N-terminal extension of one subunit and the N-terminal domain of the opposite subunit. A salt bridge is also formed between Asp222 and Arg98 of each monomer.

Molecular features that are known to contribute to thermostability include an increased number of hydrogen bonds and ion pairs, shortening of surface loops, tighter core packing, increased hydrophobicity and higher oligomeric states. It appears that both ion-pair interactions and hydrophobicity contribute to the thermostability of *S. solfataricus* SPAT. In addition to a number of ion pairs, there are three ion-pair networks close to the surface of the *S. solfataricus* SPAT protein, comprising Glu118–Lys115–Asn112–Arg293–Glu295, Arg270–Glu262–Arg267–His271 and Lys97–Glu93–Arg107–Asp90. The dimer interface of *S. solfataricus* SPAT is largely hydrophobic and is proposed to contribute to the ability of the enzyme to retain its quaternary structure at elevated temperatures.

### 3.4. The active site

The cofactor PLP in the holoenzyme structure is in the internal aldimine form, in which it is covalently bound to the active-site Lys189 *via* a Schiff base (Fig. 3*a*). The active-site cleft is formed by the two domains of one monomer and the large domain of the neighbouring monomer. Residues from both subunits are involved in cofactor binding, but the substrate site is formed mainly by residues of the two domains from one monomer. The cofactor is bound at the bottom of the active site, with its *re* side facing the solvent. The active-site Lys189 residue is located between strands  $\beta 6$  and  $\beta 7$  and binds on the *si* face of the cofactor, hiding the lysine from the solvent. The O atoms of the phosphate group interact with the main-chain amides of Gly63 and Thr64 and the side chains of Thr64 and Gln188. The cofactor is firmly anchored by hydrogen bonding from Thr243 and Tyr240 to the O atoms of the phosphate group. The carboxyl group of Asp163 is within hydrogen-bonding distance of the pyridine N atom. Asp163 is kept in place by interactions with the main-chain amide of residue Val165. The phenol hydroxyl group of the cofactor is within bonding distance of the side chains of Thr138 and Ser166. The pyridine ring of PLP is sandwiched between the side chains of Phe88, which stacks flat against the ring on the *re* side, and Val165 from the *si* side of the ring (Fig. 4). Ivanov & Karpeisky (1969) suggested that conditions for optimal operation at each of the sequential steps of the transamination reaction are provided by structural rearrangements occurring in a preceding step; the latter involves the reorientation of the pyridine ring of PLP. Therefore, an attempt was made to elucidate structures of *S. solfataricus* SPAT complexes with substrates and inhibitors in order to monitor the rearrangements in the active site.



**Figure 2**

Folding of the *S. solfataricus* SPAT subunit shown as a ribbon diagram coloured by secondary structure:  $\alpha$ -helices are coloured red,  $\beta$ -strands yellow and loops green. The N-terminal segment which changes conformation between different enzyme states is highlighted in blue. This figure and Figs 3–7 were prepared using PyMOL (DeLano, 2002).



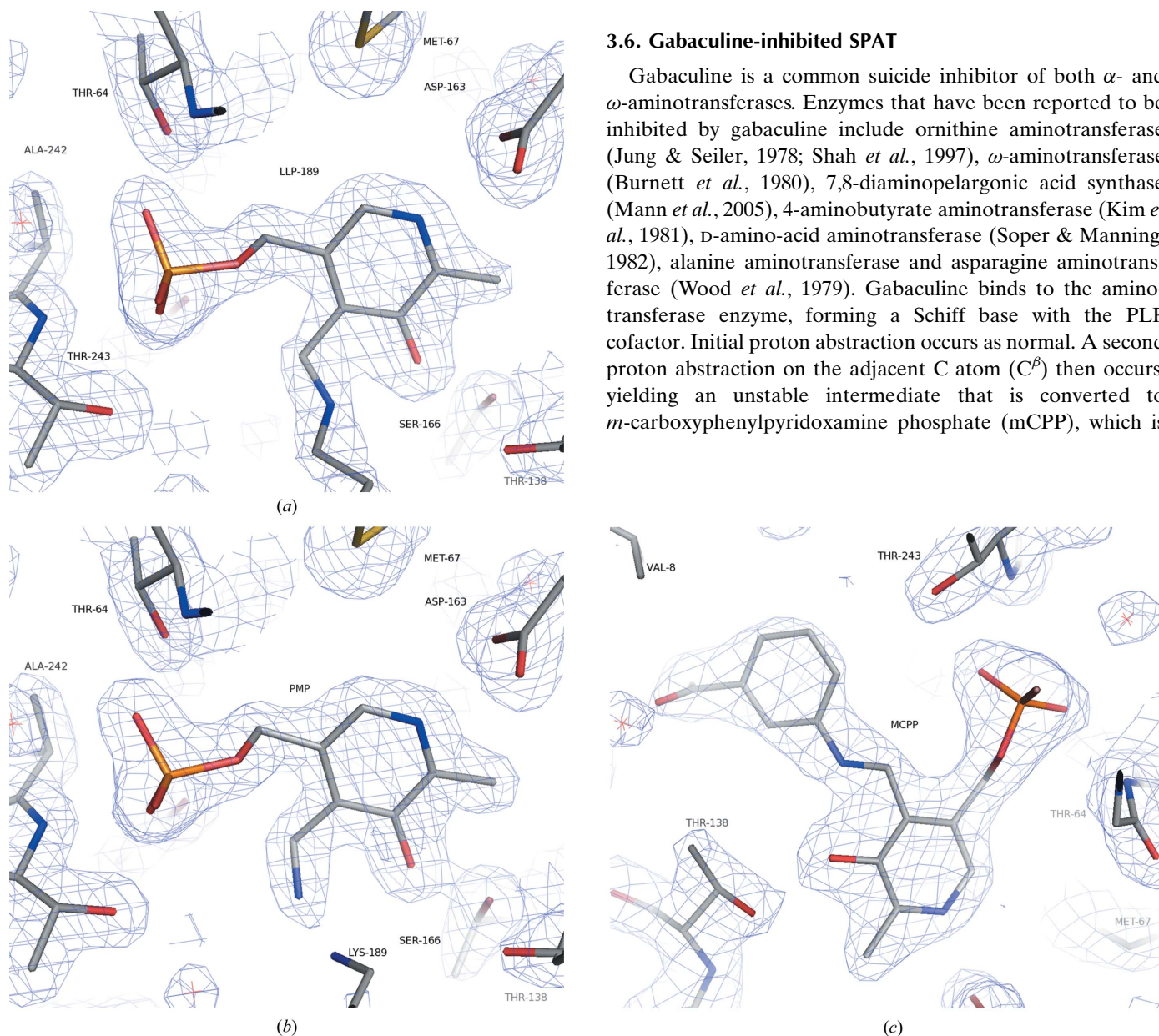
### 3.5. PMP complex

The structure of the PMP-bound form of the enzyme was obtained in an attempt to cocrystallize the enzyme–substrate complex. Cocrystallization of the protein with 20 mM of the donor substrate phenylalanine failed to produce interpretable density for the substrate. However, the Schiff base between the active-site Lys189 and PLP appeared to be broken in the electron density (Fig. 3*b*). This would imply that the enzyme had reacted with the donor substrate in the crystallization droplet and was trapped in the PMP state owing to the excess of the donor substrate. A subunit of the holoenzyme structure can be superimposed with a subunit of the PMP complex with

an r.m.s.d. of 0.35 Å for 370 C $\alpha$  atoms. No significant structural differences were observed in the active sites of the two structures except for a small displacement of the active-site Lys189 and the phosphate of PLP. No rotation of the pyridine ring of PLP was observed. This differs from chicken mitochondrial aspartate aminotransferase, for which a rotation of the cofactor pyridine ring of several degrees was observed between the internal aldimine and PMP-bound forms of the enzyme (McPhalen *et al.*, 1992). However, one would expect a close orientation of the pyridine ring at the start of each of the two transamination half-reactions. The N-termini in the holoenzyme and the PMP-complex structure adopt a similar conformation, although they are much better defined in the PMP-bound structure.

### 3.6. Gabaculine-inhibited SPAT

Gabaculine is a common suicide inhibitor of both  $\alpha$ - and  $\omega$ -aminotransferases. Enzymes that have been reported to be inhibited by gabaculine include ornithine aminotransferase (Jung & Seiler, 1978; Shah *et al.*, 1997),  $\omega$ -aminotransferase (Burnett *et al.*, 1980), 7,8-diaminopelargonic acid synthase (Mann *et al.*, 2005), 4-aminobutyrate aminotransferase (Kim *et al.*, 1981), D-amino-acid aminotransferase (Soper & Manning, 1982), alanine aminotransferase and asparagine aminotransferase (Wood *et al.*, 1979). Gabaculine binds to the aminotransferase enzyme, forming a Schiff base with the PLP cofactor. Initial proton abstraction occurs as normal. A second proton abstraction on the adjacent C atom (C $^{\beta}$ ) then occurs, yielding an unstable intermediate that is converted to *m*-carboxyphenylpyridoxamine phosphate (mCPP), which is

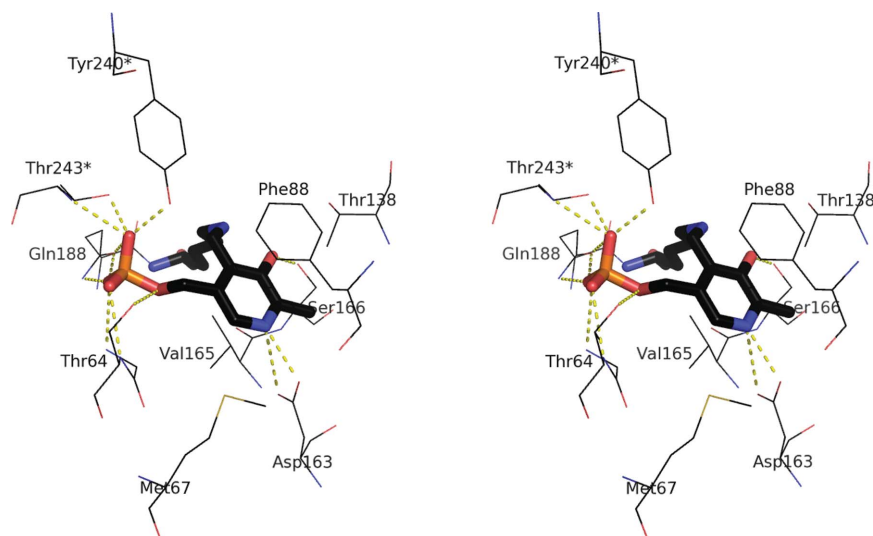


**Figure 3**

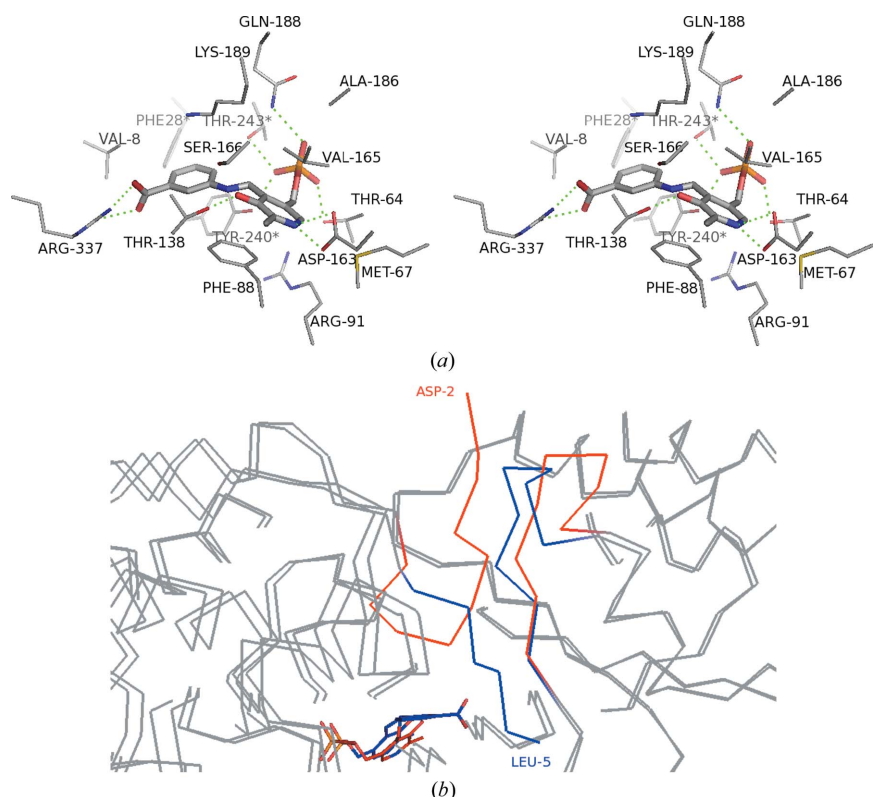
2F<sub>o</sub> - F<sub>c</sub> electron-density maps of the *S. solfataricus* SPAT active site are shown contoured at 1 $\sigma$ . (a) Electron-density map of the holoenzyme calculated at 1.82 Å resolution. The Lys189–PLP Schiff base is labelled LLP-189. (b) Electron-density map of the PMP-bound form calculated at 1.82 Å resolution, showing that the covalent link between the cofactor and Lys189 is broken. (c) Electron-density map of the gabaculine complex calculated at 2 Å resolution. The nonhydrolysable PLP–gabaculine complex is shown as an mCPP molecule.

extremely stable, resulting in an irreversible aromatic modification of the cofactor (Rando, 1977; Shah *et al.*, 1997; Fu &

Silverman, 1999). Alanine racemase and tryptophanase, which do not catalyse the second proton abstraction at  $C^\beta$  of their



**Figure 4**  
The cofactor–protein interactions in the active site of the *S. solfataricus* SPAT holoenzyme. The stereo diagram shows the residues within 4.5 Å of the cofactor. The Lys189–PLP Schiff base is highlighted as a stick model. Residues from the neighbouring subunit are indicated with an asterisk.



**Figure 5**  
Binding of the mCPP complex to *S. solfataricus* SPAT. (a) A stereo diagram showing the interactions of the gabaculine–PLP complex molecule mCPP bound in the *S. solfataricus* SPAT active site. Residues within 4.5 Å of the inhibitor are shown. (b) Rearrangement of the N-terminal residues between the PMP structure and the gabaculine complex. Both molecules are shown as a grey  $C^\alpha$  trace, except for the N-termini,  $\beta$ -sheets of the small domain and the moving loop 318–325, which are shown in red for the PMP structure and in blue for the gabaculine complex. The PMP and mCPP molecules are shown as stick models in the same colours.

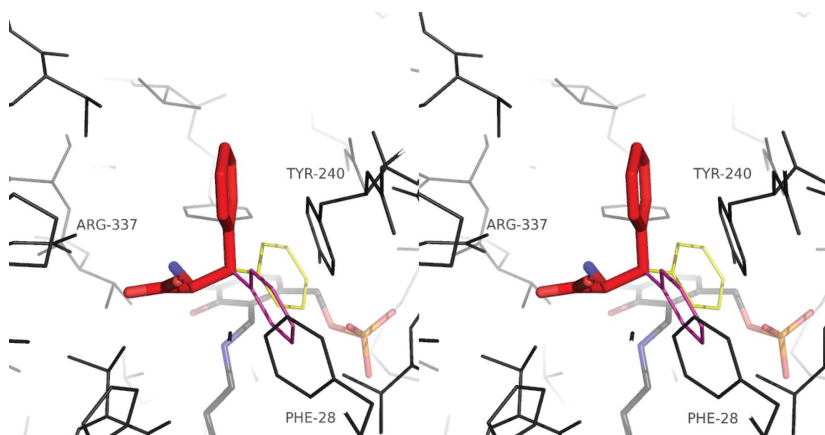
substrates, have been shown to be insensitive towards gabaculine (Soper & Manning, 1982). All previously reported aminotransferases have been shown to be irreversibly inhibited by gabaculine owing to the nature of the mCPP ligand, which is tightly bound to the active site of the enzyme. Release of the mCPP ligand is only achieved through denaturation of the protein (Shah *et al.*, 1997).

*S. solfataricus* SPAT was prepared in a slight molar excess of gabaculine and was shown to be completely inhibited after a 1 h incubation. The experiment was repeated in the presence of excess PLP and the enzyme still exhibited no activity, suggesting that the inhibitor ligand cannot leave the active site, which is consistent with previously reported gabaculine aminotransferase inhibition (Shah *et al.*, 1997; Fu & Silverman, 1999).

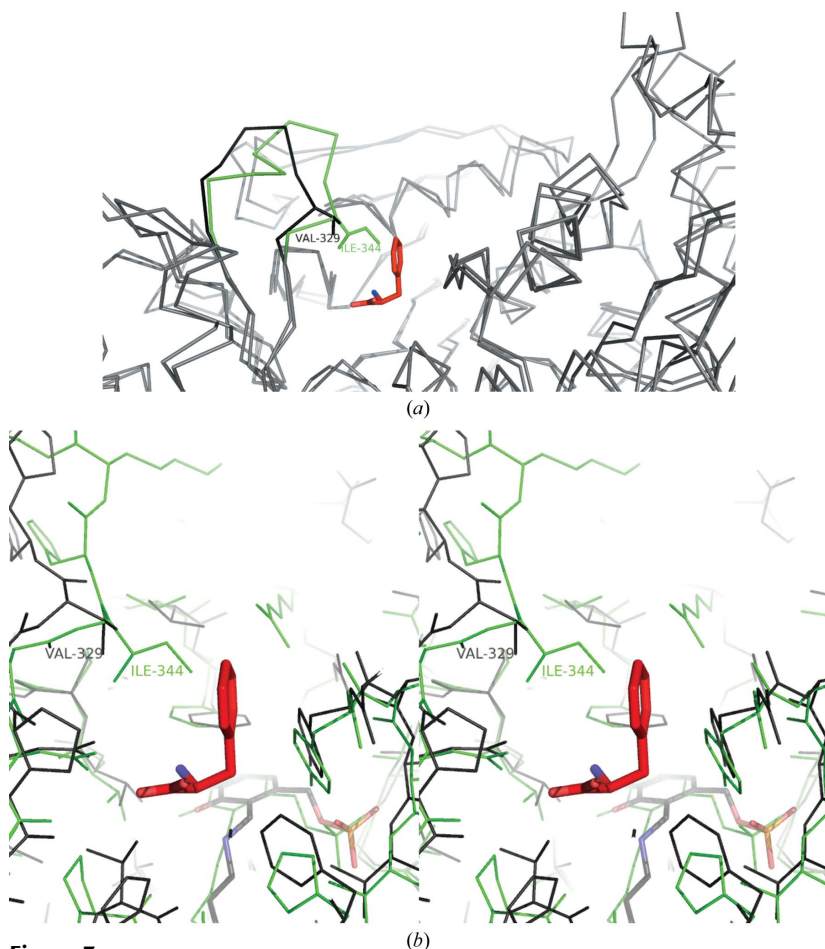
The electron density in the enzyme active site allowed modelling of the cofactor covalently bound to gabaculine as the irreversible mCPP ligand. A  $2F_o - F_c$  electron-density map for the mCPP ligand and the active-site lysine is shown in Fig. 3(c). The carboxyl group of the gabaculine molecule makes hydrogen bonds to Arg337 with a length of 3.2 Å. The gabaculine then displaces the internal lysine–PLP Schiff base to form the external aldimine complex with cofactor, which spontaneously rearranges to form the nonhydrolysable adduct mCPP. The gabaculine part of the mCPP molecule is bound in a hydrophobic pocket formed by residues Val8, Phe88 and Ala326 from the PLP-binding subunit and Phe28 and Tyr240 from the neighbouring subunit (Fig. 5a). The aromatic ring of the gabaculine molecule stacks against Tyr240 at an approximate angle of 45° and a distance of 3.7 Å. These interactions are strong enough to render the enzyme inactive and trap the mCPP ligand in the active site. This is consistent with other aminotransferase enzymes inhibited by gabaculine, such as ornithine aminotransferase (PDB entry 1gbn; Shah *et al.*, 1997), in which Tyr85 and Phe177 stack on either side of the gabaculine ring, which is reported to provide a structural basis for a dead-end irreversible intermediate.

A comparison of the gabaculine-bound and holoenzyme structures showed that they were very similar overall, with an r.m.s.d. of 0.4 Å calculated across 370 residues. A



**Figure 6**

A stereo diagram showing the possible binding of the substrate phenylalanine in the *S. solfataricus* SPAT active site. This is based on the structure of the gabaculine complex and the requirement for the scissile C $\alpha$ —H bond to be normal to the pyridine ring of PLP. The three energy-favourable rotamers of phenylalanine are shown as red, yellow and purple stick models. Arg337, which binds the carboxy group of the substrate, and Phe28 and Tyr240, which sterically prohibit binding of two of the rotamers (purple and yellow), are labelled.

**Figure 7**

Comparison of the active sites of SPAT and yeast AGAT. (a) Superimposition of the SPAT enzyme in the region of its active site with yeast AGAT shown as a C $\alpha$  trace with the modelled phenylalanine substrate shown in red. The differences in the loop regions are highlighted in black (SPAT) and green (AGAT). Ile344 in AGAT and Val329 in SPAT are shown as stick models and labelled. (b) A stereo diagram showing superposition of the active sites of the two enzymes as stick models with an allowed rotamer of phenylalanine (red) modelled in the SPAT (black) active site. Ile344 of AGAT and Val329 of SPAT are labelled.

significant difference in the location of the N-terminal region of the protein and an adjustment of the adjacent region formed by residues 318–325 located between  $\alpha$ -helix 11 and strand 3 of the small-domain  $\beta$ -sheet is observed between the two structures. In the PMP-bound and the holoenzyme structures the N-terminus folds back at residues 8–11 to form a two-stranded parallel  $\beta$ -sheet with small-domain residues 323–324. In the gabaculine-complex structure the N-terminal loop adopts an extended conformation antiparallel to the neighbouring strand 3 (residues 326–327) of the  $\beta$ -sheet of the small domain. It folds over the gabaculine ligand, ‘locking’ the complex in place and partially shielding the active site from the bulk solvent (Fig. 5b). It would appear that this rearrangement of the N-terminus is caused by binding of the mCPP carboxyl group to Arg337, leading to a displacement of Val8. This movement is accompanied by a change in the conformation of loop 318–325, with displacement of most C $\alpha$  positions by 3–5 Å towards the position occupied by the N-terminus in the holoenzyme and PMP-complex structures. Interestingly, the complexes of related proteins with their substrate analogues (PDB entries 2bkw, 1h0c and 1j04; Meyer *et al.*, 2005; Zhang *et al.*, 2003; Djordjevic *et al.*, 2010) are more similar to the SPAT holoenzyme structure than to the SPAT–gabaculine complex.

The bulky donor substrate phenylalanine was manually positioned in the *S. solfataricus* SPAT active site with the carboxy group of the substrate in the same position as in the gabaculine complex. This was orientated for catalysis according to the Dunathan hypothesis (Dunathan, 1966), positioning the cleaved C $\alpha$ —H bond of the amino acid normal to the plane of the PLP pyridine ring pointing towards the active-site Lys189. It appears that steric restrictions provided by the side chains of Phe88 and Tyr240 from the adjacent subunit allow only one single rotamer of the substrate side chain to fit into the active site (Fig. 6). This would explain the absence of activity towards substrates branching at C $\beta$  such as threonine and  $\beta$ -methyl-DL-phenylalanine (Fig. 1), since these substrates would sterically clash with aromatic residues within the active site. The substrate-binding pocket of SPAT is relatively hydrophobic, which does not favour the binding of the charged amino-acid side chains of aspartate and glutamate. This would explain the absence of activity of SPAT towards these dicarboxylic amino acids (Fig. 1) and the inability to use  $\alpha$ -ketoglutarate as an amino



acceptor. In conclusion, noncharged amino acids with no branching at C<sup>β</sup> should be good substrates for the *S. solfataricus* aminotransferase.

### 3.7. Structural comparisons

*S. solfataricus* SPAT shares 36% sequence identity with yeast AGAT (Meyer *et al.*, 2005; PDB entry 2bkv), with an r.m.s.d. of 1.33 Å calculated across 328 residues. The yeast AGAT enzyme has been reported to have high substrate specificity for alanine and glyoxylate, unlike SPAT, which exhibits a broad amino-acid substrate specificity. The limited substrate range of the AGAT has previously been attributed to nonproductive binding of substrates such as serine (Takada & Noguchi, 1985; Schlösser *et al.*, 2004; Meyer *et al.*, 2005). The structures of AGATs from human (Zhang *et al.*, 2003), *Anabaena* (Han *et al.*, 2005) and mosquito (Han *et al.*, 2006) have all been reported, and they share 30, 28 and 26% sequence identity, respectively, with *S. solfataricus* SPAT. The human and mosquito enzymes have been reported to be largely specific for catalysing alanine:glyoxylate transamination (Zhang *et al.*, 2003; Han *et al.*, 2006) and the structures of their active sites are similar to that of yeast AGAT.

The substrate-binding pocket of *S. solfataricus* SPAT is significantly different from the active sites of some other enzymes of the family IV aminotransferases, *e.g.* phosphoserine aminotransferase (19% sequence identity to SPAT; PDB entry 3ffr) and selenocysteine  $\beta$ -lyase NifS/CsdB (Lima, 2002; 17% sequence identity to SPAT; PDB entry 1kmj). These aminotransferases have several charged residues at the substrate-binding site; however, the SPAT substrate-binding pocket is mainly hydrophobic. The substrate-binding pockets of AGATs are remarkably similar to that of *S. solfataricus* SPAT.

The phenylalanine model was built into the SPAT active site as previously described and was compared with the superimposed yeast AGAT enzyme structure. Most residues in the vicinity of the modelled substrate are conserved between the two proteins. The *S. solfataricus* SPAT enzyme has a significantly larger substrate-binding pocket as its loop region between strands 9 and 10 is two amino acids shorter. This gives more space for a bulkier substrate to bind compared with the AGAT enzyme, which is only active towards alanine and glycine. Ile344 in AGAT occupies the space where larger substrates would sit in the active site of the SPAT enzyme (Fig. 7). The C<sup>δ1</sup> atom of Ile344 is positioned 2.9 Å from the C<sup>γ</sup> atom of the modelled substrate, thus hindering the binding of any amino acid larger than alanine. In the *S. solfataricus* SPAT enzyme the corresponding Val329 is positioned further away, allowing the binding of larger amino acids; the closest atom of the Val329 side chain is positioned at least 5 Å away from the nearest atom of the modelled substrate.

This paper has described both biochemical and structural studies of a thermophilic *S. solfataricus* SPAT enzyme and has provided a greater insight into the structural rearrangements that occur within the active site of the enzyme during catalysis. It represents the first structure to be described for a thermo-

philic archaeal enzyme from this relatively unstudied group IV of aminotransferase enzymes. It has also provided a greater understanding of substrate specificity, which is of considerable interest in the utilization of this enzyme in commercial biocatalysis.

The authors would like to thank the beamline staff scientists at Daresbury Synchrotron and Diamond Synchrotron, Didcot, England. CS acknowledges a PhD GTA bursary from the University of Exeter. MB gratefully acknowledges PhD funding from the BBSRC. Funding for biocatalysis in JW's laboratory was from BBSRC and EPSRC GR/S62505/01. Funding of JL's laboratory has been supported by the Wellcome Trust, BBSRC and EPSRC.

### References

- Braunstein, A. E. & Shemyakin, M. M. (1953). *Biokhimiya*, **18**, 393–411.
- Burnett, G., Yonaha, K., Toyama, S., Soda, K. & Walsh, C. (1980). *J. Biol. Chem.* **255**, 428–432.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Djordjevic, S., Zhang, X., Bartlam, M., Ye, S., Rao, Z. & Danpure, C. J. (2010). *Acta Cryst.* **F66**, 233–236.
- Dunathan, H. C. (1966). *Proc. Natl Acad. Sci. USA*, **55**, 712–716.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Fu, M. & Silverman, R. B. (1999). *Bioorg. Med. Chem.* **7**, 1581–1590.
- Han, G. W. *et al.* (2005). *Proteins*, **58**, 971–975.
- Han, Q., Robinson, H., Gao, Y. G., Vogelaar, N., Wilson, S. R., Rizzi, M. & Li, J. (2006). *J. Biol. Chem.* **281**, 37175–37182.
- Hayashi, H., Inoue, K., Nagata, T., Kuramitsu, S. & Kagamiyama, H. (1993). *Biochemistry*, **32**, 12229–12239.
- Hecquet, L., Bolte, J. & Demuynck, C. (1996). *Tetrahedron*, **52**, 8223–8232.
- Ichiyama, A. & Greenberg, D. M. (1957). *J. Biol. Chem.* **224**, 331–340.
- Ivanov, V. I. & Karpeisky, M. Y. (1969). *Adv. Enzymol. Relat. Areas Mol. Biol.* **32**, 21–53.
- Izumi, Y., Yoshida, T. & Yamada, H. (1990). *Eur. J. Biochem.* **190**, 285–290.
- Jung, M. J. & Seiler, N. (1978). *J. Biol. Chem.* **253**, 7431–7439.
- Kim, D. S., Moses, U. & Churchich, J. E. (1981). *Eur. J. Biochem.* **118**, 303–308.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- Liepmann, A. H. & Olsen, L. J. (2001). *Plant J.* **25**, 487–498.
- Lima, C. D. (2002). *J. Mol. Biol.* **315**, 1199–1208.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* **D64**, 125–132.
- Mann, S., Marquet, A. & Ploux, O. (2005). *Biochem. Soc. Trans.* **33**, 802–805.
- McPhalen, C. A., Vincent, M. G. & Jansonius, J. N. (1992). *J. Mol. Biol.* **225**, 495–517.
- Mehta, P. K., Hale, T. I. & Christen, P. (1993). *Eur. J. Biochem.* **214**, 549–561.
- Metzler, D. E., Ikawa, M. & Snell, E. E. (1954). *Biochemistry*, **4**, 1518–1525.
- Meyer, P., Liger, D., Leulliot, N., Quevillon-Cheruel, S., Zhou, C.-Z., Borel, F., Ferrer, J.-L., Poupon, A., Janin, J. & van Tilbeurgh, H. (2005). *Biochimie*, **87**, 1041–1047.
- Moll, R. & Schaefer, G. (1988). *FEBS Lett.* **232**, 359–363.
- Murshudov, G. N. & Dodson, E. J. (1997). *CCP4 Newsl. Protein Crystallogr.* **33**, 31–39.

- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Noguchi, T., Minatogawa, Y., Takada, Y., Okuno, E. & Kido, R. (1978). *Biochem. J.* **170**, 173–175.
- Noguchi, T., Okuno, E., Takada, Y., Minatogawa, Y., Okai, K. & Kido, R. (1978). *Biochem. J.* **169**, 113–122.
- Oda, T., Miyajima, H., Suzuki, Y., Ito, T., Yokota, S., Hoshino, M. & Ichiyama, A. (1989). *J. Biochem.* **106**, 460–467.
- Okuno, E., Minatogawa, Y., Nakamura, M., Kamoda, N., Nakanishi, J., Makino, M. & Kido, R. (1980). *Biochem. J.* **189**, 581–590.
- Onuffer, J. J., Ton, B. T., Klement, I. & Kirsch, J. F. (1995). *Protein Sci.* **4**, 1743–1749.
- Otwinowski, Z. & Minor, W. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 226–235. Dordrecht: Kluwer Academic Publishers.
- Ramakrishnan, C. & Ramachandran, G. N. (1965). *Biophys. J.* **5**, 909–933.
- Rando, R. R. (1977). *Biochemistry*, **16**, 4604–4610.
- Richardson, J. S. (1981). *Adv. Protein Chem.* **34**, 167–339.
- Rosa, M. de, Gambacorta, A. & Bu'lock, J. D. (1975). *J. Gen. Microbiol.* **86**, 156–164.
- Rowell, E. V., Snell, K., Carnie, J. A. & Rowell, K. V. (1972). *Biochem. J.* **127**, 155–165.
- Sallach, H. J. (1956). *J. Biol. Chem.* **223**, 1101–1108.
- Schlösser, T., Gärtgens, C., Weber, U. & Stahmann, K.-P. (2004). *Yeast*, **21**, 63–73.
- Shah, S. A., Shen, B. W. & Brünger, A. T. (1997). *Structure*, **5**, 1067–1075.
- She, Q. *et al.* (2001). *Proc. Natl Acad. Sci. USA*, **98**, 7835–7840.
- Soper, T. S. & Manning, J. M. (1982). *J. Biol. Chem.* **257**, 13930–13936.
- Takada, Y. & Noguchi, T. (1985). *Biochem. J.* **231**, 157–163.
- Takayama, S., McGarvey, G. J. & Wong, C.-H. (1997). *Annu. Rev. Microbiol.* **51**, 285–310.
- Umbarger, H. E., Umbarger, M. A. & Siu, P. M. L. (1963). *J. Bacteriol.* **85**, 1431–1439.
- Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* **D55**, 191–205.
- Walsh, D. A. & Sallach, H. J. (1966). *J. Biol. Chem.* **241**, 4068–4076.
- Wood, J. D., Kurylo, E. & Tsui, D. (1979). *Neurosci. Lett.* **14**, 327–331.
- Zhang, X., Roe, S. M., Hou, Y., Bartlam, M., Rao, Z., Pearl, L. H. & Danpure, C. J. (2003). *J. Mol. Biol.* **331**, 643–652.