Appl Microbiol Biotechnol (2008) 79:775–784
DOI 10.1007/s00253-008-1487-1

BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Two groups of thermophilic amino acid aminotransferases exhibiting broad substrate specificities for the synthesis of phenylglycine derivatives

Daisuke Koma · Toshiya Sawai · Ryotaro Hara · Shigeaki Harayama · Kuniki Kino

Received: 19 December 2007 / Revised: 31 March 2008 / Accepted: 31 March 2008 / Published online: 15 May 2008

Abstract Thirty two thermophilic amino acid aminotransferases (AATs) were expressed in *Escherichia coli* as soluble and active proteins. Based on their primary structures, the 32 AATs were divided into four phylogenetic groups (classes I, II, IV, and V). The substrate specificities of these AATs were examined, and 12 AATs were found capable of synthesizing ring-substituted phenylglycine derivatives such as hydroxyl-, methoxy-, and fluorophenylglycines. Eleven out of the 12 AATs were enzymes belonging to two phylogenetic groups namely, one subgroup of the class I family and the class IV family. AATs in these two groups may thus be useful for the synthesis of a variety of ring-substituted phenylglycine derivatives.

Keywords Aminotransferase · Thermophile · Phenylglycine · Unnatural amino acid · Nonproteinogenic amino acid

Electronic supplementary material The online version of this article (doi:10.1007/s00253-008-1487-1) contains supplementary material, which is available to authorized users.

D. Koma · S. Harayama
NITE Biological Resource Center (NBRC),
Department of Biotechnology,
National Institute of Technology and Evaluation (NITE),
2-5-8 Kazusakamatari,
Kisarazu-shi, Chiba 292-0818, Japan

S. Harayama
e-mail: harayama@kc.chuo-u.ac.jp

T. Sawai · R. Hara · K. Kino
Laboratory of Applied Biochemistry,
Department of Applied Chemistry,
School of Science and Engineering, Waseda University,
3-4-1 Ohkubo,
Shinjuku-ku, Tokyo 169-8555, Japan

Present address:
D. Koma (✉)
Osaka Municipal Technical Research Institute,
1-6-50 Morinomiya,
Joto-ku, Osaka 536-8553, Japan
e-mail: koma@omtri.city.osaka.jp

Present address:
S. Harayama
Department of Biological Sciences,
Faculty of Science and Engineering, Chuo University,
1-13-27 Kasuga,
Bunkyo-ku, Tokyo 112-8551, Japan

Introduction

Amino acid aminotransferases (AATs) or transaminases are pyridoxal-5′-phosphate-dependent enzymes that catalyze the transfer of the amino group of an amino acid to an α-keto acid. Biologically, this reaction is important in the production of various amino acids, while industrially, AATs are useful for the synthesis of optically pure amines (Shin and Kim 1999) and nonproteinogenic amino acids (Ager et al. 2000; Ager and Fotheringham 2001) that have wide usages as chiral building blocks and intermediates for the synthesis of peptidemimetic pharmaceuticals (Pojitkov et al. 2000; Taylor et al. 1998).

Thermophilic enzymes are generally promising in a number of commercial applications (Sonnleitner and Fiechter 1983) largely due to their innate biochemical characteristics such as thermostability and tolerance against chemical regents. Thus, many thermophilic AATs such as aspartate aminotransferase (AspAT; Birolo et al. 1991; Cubellis et al. 1989; Marino et al. 1988; Okamoto et al.

1996), alanine aminotransferase (AlaAT; Word et al. 2000), and aromatic amino acid aminotransferase (AroAT; Andreotti et al. 1994; Andreotti et al. 1995; Matsui et al. 2000) have been purified and characterized. We also have been interested in using thermophilic AATs for the synthesis of nonproteinogenic amino acids and developed a new high-throughput method to evaluate the thermophilic AAT activities for the synthesis of nonproteinogenic amino acids, using the coupled reaction with an ω-AAT and an aldehyde dehydrogenase (Sawai et al. 2007). Furthermore, using this method, we found that an AAT encoded by the ORF ST1411 (GenBank accession no. BAB66478) from *Sulfolobus tokodaii* was able to synthesize some ring-substituted phenylglycines.

Different classifications of AATs have been proposed in literatures. For example, AATs have been divided into four subfamilies (I, II, III, and IV) based on their primary structures (Mehta et al. 1993). The distribution of the conserved domains (accession nos. PF00155, PF00202, PF01063, and PF00266 in Pfam; http://www.sanger.ac.uk/Software/Pfam/) and the motifs (accession nos. PDOC00098, PDOC00518, PDOC00519, PDOC00618, and PDOC00514 in PROSITE; http://kr.expasy.org/prosite/) further divided subfamily I into two classes (class I and class II) (Hwang et al. 2005). According to their three-dimensional structures, AATs have been grouped into two fold types (fold types I and IV; Schneider et al. 2000): Fold type I contains three subfamilies or four classes namely, subfamily I (class I/II), subfamily II (class III), and subfamily IV (class V), while fold-type IV contains a single subfamily or class namely, subfamily III (class IV). In these classifications, a single subfamily includes enzymes of different substrate specificities. In other words, it is difficult to deduce the substrate specificities of AATs directly from the subfamily classification. A possible exception would be subfamily III (class IV) AATs in which only two types of activities had so far been found: branched-chain amino acid (leucine, isoleucine, and valine) AATs (BCATs) and D-amino acid AATs.

Among sequence data originating from genome-sequencing projects, many open reading frames (ORFs) exhibiting significant similarities to known AATs have been discovered. Since the substrate specificities of the putative AATs could not be deduced from their primary structures, we characterized in this study the substrate specificities of 32 thermophilic AATs belonging to the three subfamilies (four classes) hoping that such study would clarify the structure/function relationships of thermophilic AATs by identifying several monophyletic subgroups of AATs sharing similar substrate specificities. We found that AATs in some phylogenetic clusters exhibited similar substrate specificities, while in other cases, there was no apparent correlation between the phylogenetic positions and substrate specificities.

Materials and methods

Cloning of AAT genes

The following AAT genes were amplified from the genomic deoxyribonucleic acids (DNAs) by polymerase chain reaction (PCR) using Robocycler (Stratagene): The APE0169 (accession no. BAA79080) and APE2248 (BAA81260) genes from *Aeropyrum pernix* K1 (NBRC 100138), the AF0933 (AAB90305), AF1417 (AAB89830), and AF2129 (AAB89121) genes from *Archaeoglobus fulgidus* DSM4304 (NBRC 100126), the MJ0955 (AAB98960), MJ0959 (AAB98961), and MJ1008 (AAB99010) genes from *Methanocaldococcus jannaschii* DSM2661 (NBRC 100440), the MTH1430 (AAB85907) gene from *Methanothermobacter thermautotrophicus* str. Delta H (NBRC 100330), the PH0207 (BAA29276), PH1308 (BAA30413), and PH1371 (BAA30477) genes from *Pyrococcus horikoshii* OT3 (NBRC 100139), the ST1217 (BAB66259) gene from *Sulfolobus tokodaii* 7 (NBRC 100140), the TTE0301 (AAM23597), TTE0933 (AAM24189), TTE1206 (AAM24436), TTE1368 (AAM24590), TTE2440 (AAM25573), and TTE2614 (AAM25734) genes from *Thermoanaerobacter tengcongensis* MB4 (NBRC 100824), the TVN0402 (BAB59546) gene from *Thermoplasma volcanium* GSS1 (NBRC15438), the TM1131 (AAD36207), TM1255 (AAD36330), TM1400 (AAD36471), and TM1698 (AAD36765) genes from *Thermotoga maritima* MSB8 (NBRC 100826), and the TTHA0046 (BAD69869), TTHA0124 (BAD69947), and TTHA0428 (BAD70251) genes from *Thermus thermophilus* HB8 (NBRC 101084). The genomic DNAs were prepared by using DNeasy Tissue Kit (QIAGEN). The primers used for PCR are summarized in Supplemental Table S1. All primers were designed in order that six histidine residues would be added to the C termini of the gene products. The PCR conditions with KOD and KOD plus polymerases (Toyobo) were as follows: 98°C (3 min) → [98°C (35 s) → 52°C (35 s) → 74°C (55 s)]×25 cycles → 74°C (2 min) for KOD polymerase, and 94°C (3 min) → [94°C (35 s) → 52°C (50 s) → 68°C (2 min)] ×25 cycle →68°C (3 min) for KOD plus polymerase. If the desired DNA band was not obtained on agarose gel electrophoresis after the first PCR reaction, a second PCR reaction was performed using 1 μl of 1/10-diluted first PCR product as a template DNA. Each PCR product was purified by GFX PCR DNA and Gel Band Purification Kit (GE healthcare) and then digested with appropriate restriction enzymes to introduce in pET21a(+) or pET21d(+) (Novagen). The recombinant plasmid was introduced into *Escherichia coli* JM109 (Takara) and screened. Finally, the host strain, *E. coli* Rosetta (DE3) (Novagen), was transformed with the recombinant plasmid.

Cloning of the following genes was described previously (Koma et al. 2006; Sawai et al. 2007): the MJ0001 (AAB97984) gene from *M. jannaschii* DSM2661, the PAE2315 (AAL64105) gene from *Pyrobaculum aerophilum* IM2 (NBRC 100827), the PTO0509 (AAT43094) gene from *Picrophilus torridus* DSM9790 (NBRC 100828), the ST1411 (BAB66478) gene from *S. tokodaii* 7, and the TTHA0411 (BAD70234) gene from *T. thermophilus* HB8.

Gene expression and protein purification

E. coli Rosetta (DE3) harboring the respective recombinant plasmid was cultivated in 100-ml Luria–Bertani (LB) medium until the absorbance at 660 nm reached 0.5. Subsequently, isopropyl-β-D-thiogalactopyranoside (IPTG) was added in the culture to the concentration of 1 mM, and the cells were continuously cultivated at 25°C or 46°C for 20 h.

The histidine-tagged recombinant AATs were purified using a Protino Ni-TED 2000 column (Macherey-Nagel) according to the method described previously (Sawai et al. 2007). The enzyme was concentrated by ultracentrifugation using Amicon Ultra-15 MW30,000 (Millipore), and the buffer of the enzyme solution was exchanged to a 20-mM phosphate buffer (pH 7.5) containing 50 μM pyridoxal-5′-phosphate (PLP) by desalting chromatography using a PD-10 column (GE Healthcare). The protein concentration was determined using the Coomassie protein assay reagent (PIERCE) and adjusted to 0.1–1 mg/ml. The purity of the proteins was examined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). The immunoblot analysis of histidine-tagged AATs using monoclonal anti-6× His antibody (HI192, Nacalai Tesque) was carried out as described previously (Koma et al. 2006). The proteins thus obtained were kept at 4°C until they were used for assays.

Enzyme assays

The substrate specificity for amino acids was investigated by measuring the amount of glutamic acid produced from 2-oxoglutaric acid when each amino acid was supplied as an amino group donor. The reaction mixture contained 40 mM 2-oxoglutarate, 40 mM amino acid (except phenylalanine [30 mM] and tryptophan [20 mM]), 0.5 mM PLP, and an appropriate amount of an AAT in a total of 200 μl of 100 mM phosphate buffer (pH 7.5). The enzyme concentration in the reaction mixture was adjusted between 1.25 and 25 μg/ml so that the maximum concentration of glutamic acid produced in the reaction was below 4 mM. The reaction was initiated by the addition of the enzyme solution (10 μl) and performed at 70°C (AATs from *P. torridus*, *T. thermophilus*, *T. tengcongensis*, and *T. volcanium*) or 80°C (AATs from *A. fulgidus*, *A. pernix*, *M. jannaschii*, *M. thermautotrophicus*, *P. aerophilum*, *P. horikoshii*, and *S. tokodaii*) for 5 min. The reaction was terminated by the addition of 50 μl trichloroacetic acid solution (30% *v*/*v*); the amount of glutamic acid produced was then measured by high-performance liquid chromatography (HPLC; Koma et al. 2006).

For nonproteinogenic amino acid synthesis, the activities of AATs in 100 mM phosphate buffer (pH 7.5) at 70°C were continuously monitored by performing a coupling reaction with a thermophilic ω-AAT and an aldehyde dehydrogenase (Sawai et al. 2007). The keto acids illustrated in Fig. 1 were used as the substrates. These compounds were chemically synthesized according to the method described previously (Creary 1987; Scheffer and Wang 2001).

Results

Production and purification of recombinant AATs

We could successfully overexpress most of the AAT genes as soluble proteins in *E. coli* Rosetta (DE3) by culturing the cells in LB medium for 20 h at 25°C after adding IPTG, while the genes for APE2248, PAE2315, and TTHA0411 were successfully overexpressed as soluble proteins by culturing the cells for 20 h at 46°C but not at 20 or 37°C (Koma et al. 2006).

Most of the purified proteins exhibited homogeneity or contained a few negligible bands on SDS-PAGE by Coomassie staining. The purified proteins encoded by the ORFs MJ1008, TTHA0124, and TTE0933, however, did not exhibit homogeneity and contained several distinct large-size bands (>100 kDa) on SDS-PAGE (Supplemental Fig. S1). These large-size bands were also stained with a specific antibody raised against the histidine tag, suggesting that the bands represent aggregates of desired proteins and/or of both desired proteins and other undesired compounds. These AATs were treated at 80°C for 15 min to inactivate any *E. coli* proteins. Since these AAT preparations might contain soluble aggregates, the specific activities could be underestimated.

Phylogenetic analysis

The phylogenetic tree of the 32 AATs was constructed together with 13 previously characterized AATs (Fig. 2). These AATs were grouped into four clusters that corresponded to four different Pfam families: class I (26 AATs), class II (three AATs), class IV (six AATs), and class V (ten AATs). The class I AATs were further divided into two subclusters: one contained AroAT from *Pyrococcus furio-*

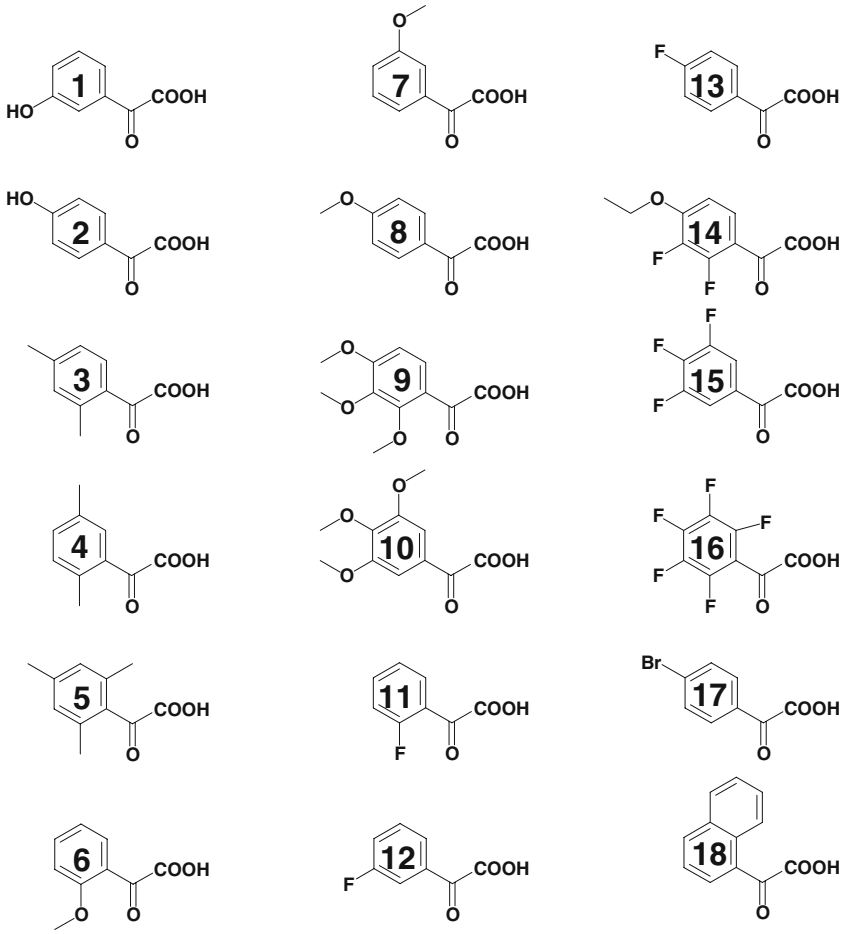

Fig. 1 The keto acids for the synthesis of phenylglycine derivatives. The name of these keto acids are as follows: *1*, 3-hydroxyphenylglyoxylic acid; *2*, 4-hydroxyphenylglyoxylic acid; *3*, 2,4-dihydroxyphenylglyoxylic acid; *4*, 2,5-dihydroxyphenylglyoxylic acid; *5*, 2,4,6-trihydroxyphenylglyoxylic acid; *6*, 2-methoxyphenylglyoxylic acid; *7*, 3-methoxyphenylglyoxylic acid; *8*, 4-methoxyphenylglyoxylic acid; *9*, 2,3,4-trimethoxyphenylglyoxylic acid; *10*, 3,4,5-trimethoxyphenylglyoxylic acid; *11*, 2-fluorophenylglyoxylic acid; *12*, 3-fluorophenylglyoxylic acid; *13*, 4-fluorophenylglyoxylic acid; *14*, 2,3-difluoro-4-ethoxyphenylglyoxylic acid, *15*, 3,4,5-trifluorophenylglyoxylic acid; *16*, 2,3,4,5,6-pentafluorophenylglyoxylic acid; *17*, 4-bromophenylglyoxylic acid; and *18*, 1-naphthylglyoxylic acid

sus, alanine–glyoxylate aminotransferase from *Thermus litoralis*, and eight AATs purified in this study, while the other contained AlaAT from *P. furiosus*, AspAT from *Sulfolobus solfataricus*, and 11 AATs purified in this study.

Substrate specificities for amino acids

The 32 AATs were further characterized by examining their substrate specificities for amino group donors (Tables 1 and 2). Three previously characterized thermophilic AATs belonging to the class I subgroup namely, the TTHA0046 AspAT (Okamoto et al. 1996), the APE2248 AlaAT (Koma et al. 2006), and the PH1371 AroAT (Matsui et al. 2000), were used as the standards for the comparison of the substrate specificities. Nineteen AATs belonging to the class I subgroup were divided into six groups according to their substrate specificities (Table 1). Four AATs encoded by the ORFs AF2129, TM1255, TTE1206, and TTE1368 (group 1) preferred aspartate similarly to TTHA0046 AspAT (Okamoto et al. 1996). On the other hand, the AAT encoded by the ORF TM1698 (group 2) was an alanine-specific AAT similarly to the thermophilic APE2248 AlaAT (Koma et al. 2006). The AATs encoded by the ORFs TTHA0411, MJ0001, and TTE0301 were active with phenylalanine, tryptophan, methionine, and leucine similarly to the thermophilic PH1371 AroAT (Matsui et al. 2000). Based on the relative activities with these amino acids, the TTHA0411 AAT and the PH1371 AroAT were grouped into group 3, while the MJ0001 and TTE0301 AATs were paired to form group 4. Seven AATs encoded by the ORFs APE0169, PAE2315, PH0207, PTO0509, ST1411, TM1131, and TVN0402 (group 5) used a broad range of hydrophobic amino acids namely, cysteine (except APE0169), methionine, alanine, phenylalanine, tryptophan, valine, leucine, and isoleucine, as amino group donors. Based on such substrate specificities, these enzymes are referred to as hydrophobic amino acid aminotransferases (HybATs) in this study. Alanine was the best substrate of the TTE2440 AAT (group 6) while methionine, leucine, and aromatic amino acids were also used as amino group donors by this enzyme. The substrate specificity of the TTE2440 AAT was not similar to those of any other AATs.

One of the major members of the class II subgroup is histidinol phosphate aminotransferase. Histidinol phosphate aminotransferase from *T. maritima* used not only histidinol phosphate but also aromatic amino acids as amino group donors (Fernandez et al. 2004). Similarly to the *T. maritima*

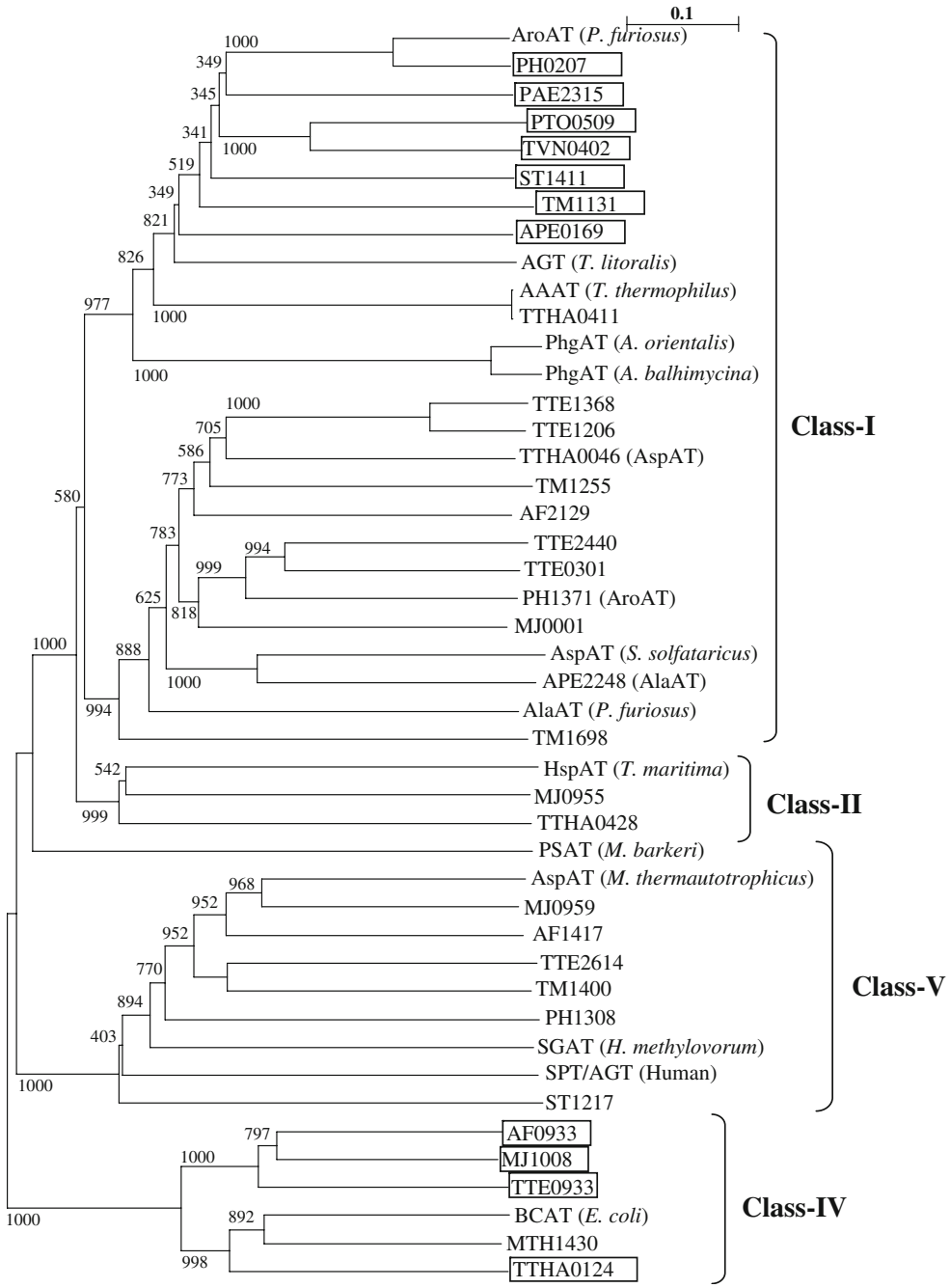

Fig. 2 The phylogenetic analysis of AATs inferred from the amino acid sequence alignment. The tree was constructed by the neighbor-joining method using ClustalX (Jeanmougin et al. 1998) based on the amino acid sequences of the AATs shown in this figure. *Numbers on nodes* indicate bootstrap values. The AATs capable of synthesizing ring-substituted phenylglycine derivatives are surrounded by *boxes*. The accession numbers of the AATs characterized in this study are described in "Materials and methods." The AATs characterized in previous studies and incorporated in the present phylogenetic analysis are as follows: α-aminoadipate aminotransferase (*AAAT*) from *T. thermophilus* HB27 (BAC76939), alanine–glyoxylate aminotransferase (*AGT*) from *T. litoralis* (BAB40321), AlaAT from *P. furiosus* (AAL81621), AroAT from *P. furiosus* (AAL80245), AspAT from *S. solfataricus* (AAK41178), AspAT from *M. thermautotrophicus* (*M. thermoformicicum*) (BAA05953), BCAT from *E. coli* (AAT48207), human serine–pyruvate/alanine–glyoxylate aminotransferase (*SPT/AGT*; AAA51680), histidinol phosphate aminotransferase (*HspAT*) from *T. maritima* (AAD36117), phenylglycine aminotransferase (*PhgAT*) from *A. balhimycina* (CAC48367), phenylglycine aminotransferase (*PhgAT*) from *A. orientalis* (CAA11790), phosphoserine aminotransferase (*PSAT*) from *M. barkeri* (AAZ70257), and serine–glyoxylate aminotransferase (*SGAT*) from *H. methylovorum* (BAA19919)

Table 1 Amino acids utilization of class I aminotransferases

Amino acid	Relative activity (%)[a]																		
	Group 1 (AspAT)					Group 2 (AlaAT)		Group 3 (AroAT)		Group 4		Group 5 (HybAT)							Group 6
	TTHA 0046	AF 2129	TM 1255	TTE 1206	TTE 1368	APE 2248	TM 1698	PH 1371	TTHA 0411	MJ 0001	TTE 0301	APE 0169	PAE 2315	PH 0207	PTO 0509	ST 1411	TM 1131	TVN 0402	TTE 2440
Asp	100	100	100	100	100	0	0	0	0	0	0	0	11±1	0	0	1±0	1±0	6±0	0
Lys	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7±0	0	0	2±0
Arg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2±1	0	0	3±0
His	0	0	0	0	0	0	0	49±2	0	4±0	5±0	0	5±0	0	1±0	18±1	0	0	8±0
Asn	45±4	16±1	60±3	35±3	17±1	0	0	0	0	0	0	0	3±1	0	4±0	2±0	0	0	0
Gln	2±0	1±0	20±4	3±1	2±0	6±0	11±0	0	0	0	3±0	5±0	4±0	1±0	2±0	9±0	1±0	0	4±0
Ser	0	0	0	0	0	18±0	0	0	0	0	0	0	0	0	0	0	0	0	0
Thr	0	1±0	0	0	0	6±1	0	0	0	0	0	11±1	1±0	0	0	7±0	0	0	0
Cys	0	0	0	0	0	0	33±0	0	6±0	0	0	0	53±0	35±4	38±3	78±3	21±1	8±0	5±1
Met	0	0	0	0	0	0	0	40±0	98±4	68±2	46±5	100	97±2	83±5	96±8	100	100	90±5	31±1
Ala	0	0	0	0	0	100	100	2±0	1±0	0	0	40±2	70±1	28±1	51±6	63±3	16±1	29±1	100
Trp	0	0	0	0	0	0	0	26±2	23±6	60±1	48±2	22±0	48±1	11±0	10±1	57±3	7±0	11±0	26±1
Phe	0	0	0	0	0	0	0	100	100	58±4	60±0	89±9	100	100	45±4	64±3	95±4	85±3	19±0
Val	0	0	0	0	0	0	0	0	5±2	0	0	39±1	38±0	42±1	48±0	58±0	53±2	58±0	3±0
Leu	0	0	0	0	0	0	0	20±0	27±5	100	100	66±4	95±2	72±7	100	67±1	54±2	100	60±1
Ile	0	0	0	0	0	0	0	0	0	0	0	37±1	36±3	21±4	24±2	37±0	54±3	23±0	0

[a] The values indicate the average of three independent experiments and means ± standard deviation.

Table 2 Amino acids utilization of class-II, III and IV aminotransferases

Amino acid	Relative activity (%)[a]												
	Class II (group 7)		Class IV (group 8, BCAT)					Class V					
	MJ 0955	TTHA 0428	AF 0933	MJ 1008	MTH 1430	TTHA 0124	TTE 0933	MJ 0959	TM 1400	TTE 2614	AF 1417	PH 1308	ST 1217
Asp	6±1	11±0	0	9±1	15±1	0	0	100	100	100	0	41±7	78±2
Lys	0	0	0	2±0	0	0	0	0	0	0	0	0	0
Arg	19±3	0	0	0	0	0	0	0	0	0	0	0	5±1
His	8±0	100	9±3	3±0	0	0	4±0	0	0	8±1	0	35±0	84±7
Asn	0	7±0	0	0	0	7±0	0	6±1	0	9±1	0	0	48±3
Gln	12±1	4±0	0	2±0	11±1	0	2±0	0	11±3	5±0	5±0	0	37±1
Ser	0	0	0	0	0	0	0	0	0	0	0	0	11±1
Thr	10±0	0	0	3±0	4±0	0	0	0	0	0	0	8±3	4±0
Cys	0	0	0	2±0	0	0	9±0	2±0	0	6±0	0	0	10±0
Met	10±1	22±9	65±0	69±9	41±3	28±1	77±2	0	0	2±0	0	0	56±8
Ala	0	7±0	13±1	2±0	0	0	2±1	3±0	44±4	32±0	100	0	100
Trp	94±0	100	22±1	10±0	0	5±0	7±3	0	0	85±9	0	0	49±6
Phe	100	82±4	8±1	12±1	0	4±0	10±2	0	0	21±0	0	100	94±5
Val	0	0	82±1	94±6	4±0	94±2	100	0	0	0	0	0	0
Leu	0	7±1	100	100	100	96±6	87±3	0	0	0	0	0	0
Ile	0	0	53±9	71±6	44±1	100	94±0	0	0	0	0	0	0

[a] The values indicate the average of three independent experiments and means ± standard deviation.

histidinol phosphate aminotransferase, both of the class II AATs examined in this study, the MJ0955 AAT and the TTHA0428 AAT (group 7), used phenylalanine and tryptophan as amino group donors.

Five AATs belonging to the class IV subgroup, encoded by the ORFs AF0933, MJ1008, MTH1430, TTHA0124, and TTE0933, were the members of BCATs, which preferred the branched-chain amino acids (group 8, Table 2).

Table 3 The aminotransferase activities of class I AATs for synthesis of unnatural amino acids

Number of keto acid[a]	Specific activity (U/mg)[b]							
	HybAT							TTE2440
	APE0169	PAE2315	PH0207	PTO0509	ST1411	TM1131	TVN0402	
1	11±1	17±5	4±1	4±1	24±4	20±4	7±3	0
2	0	0	0	1±0	4±1	1±0	1±0	0
3	69±6	14±6	4±2	6±1	44±9	179±9	1±0	0
4	46±1	5±1	4±1	0	2±0	88±13	0	0
5	0	0	0	0	0	0	0	0
6	31±1	46±5	67±12	139±44	49±12	135±4	24±6	0
7	35±2	43±3	9±2	12±4	45±16	57±17	17±2	1±0
8	11±1	10±3	1±0	11±2	36±3	18±2	5±1	0
9	1±0	0	0	0	6±1	1±0	0	0
10	0	0	0	0	3±1	0	0	0
11	84±4	106±10	133±13	322±47	106±16	277±18	31±4	2±1
12	83±6	148±13	131±14	231±30	131±12	172±16	76±2	2±1
13	81±3	101±4	38±7	118±10	138±6	190±15	58±10	0
14	26±4	30±2	6±4	13±2	26±5	53±7	3±1	12±1
15	48±6	102±26	22±6	59±11	96±5	26±7	46±3	2±1
16	20±2	6±0	2±1	1±0	7±1	12±5	0	0
17	95±16	116±12	19±3	212±11	151±29	137±9	60±10	0
18	80±7	23±8	5±2	1±1	28±4	148±15	0	11±3

[a] The number of keto acid corresponds to Fig. 1.
[b] One unit of the enzyme was defined as the amount of enzyme catalyzing the production of 1 μmol amino acid per minute. The values indicate the average of three independent experiments and means ± standard deviation.

The substrate specificities of these enzymes resembled mesophilic BCATs (Atiles et al. 2000; Venos et al. 2004). The class IV AATs are well known to contain two distinct members namely, BCAT and D-AAT. All the class IV AATs examined in this study did not use any D-isomers of 16 L-amino acids listed in Table 2.

Six AATs belonging to the class V subgroup showed varied substrate specificities (Table 2). Three out of six class V AATs, encoded by the ORFs MJ0959, TM1400, and TTE2614, mostly preferred aspartate. The amino acid sequence of the MJ0959 AAT was similar to that of AspAT from *Methanobacterium thermoformicicum* (synonym: *Methanothermobacter thermautotrophicus*; Tanaka et al. 1994; Fig. 2). The amino acid utilization patterns of the MJ0959 AAT and AspAT from *M. thermoformicicum* were very similar to each other and also resembled those of the enzymes in group 1 of the class I subgroup. Although the TM1400 and TTE2614 AATs exhibited preferred substrate specificities for aspartate, their amino acid utilization patterns were different from each other and from those of the group 1 enzymes. The TTE2614 AAT preferred not only aspartate but also alanine, tryptophan, and phenylalanine, whereas the preferred substrates for the TM1400 AAT were aspartate and alanine. The AAT encoded by the ORF AF1417 specifically used alanine similarly to the group 2 enzymes of the class I AATs, whereas the AAT encoded by the ORF PH1308 used phenylalanine, histidine, and aspartate. The AAT encoded by the ORF ST1217 had a broad substrate specificity. The enzyme preferred several hydrophobic amino acids such as alanine, methionine, phenylalanine, and tryptophan similarly to the rat mitochondrial alanine/glyoxylate aminotransferase (Oda et al. 1989); however, unlikely to the mitochondrial alanine/glyoxylate aminotransferase, it also preferred the acidic amino acid aspartate.

Evaluation of AAT activities for the synthesis of ring-substituted phenylglycine derivatives

AAT activities for the synthesis of nonproteinogenic amino acids were evaluated by the assay method described previously (Sawai et al. 2007). In the assay, the inhibitions of the ω-AT and the aldehyde dehydrogenase by the ring-substituted phenylglyoxylic acid derivatives listed in Fig. 1 were not observed, and none of these ring-substituted phenylglyoxylic acid derivatives were substrates of the ω-AT and the aldehyde dehydrogenase. Thus, the AAT activities for the synthesis of nonproteinogenic amino acids from the ring-substituted phenylglyoxylic acid derivatives listed in Fig. 1 could be evaluated by this assay method.

Twenty out of the 32 AATs exhibited no or less activities against any of the ring-substituted phenylglyoxylic acid derivatives. None of the class I AATs belonging to groups 1 and 2 synthesized the ring-substituted phenylglycine derivatives. Although the class I AATs belonging to groups 3 and 4 showed preferred specificities for aromatic amino acids, these enzymes did not synthesize ring-substituted phenylglycine derivatives, except the PH1371 AAT. Similarly, none of the class II AATs and none of the class V AATs reacted with any of the ring-substituted phenylglyoxylic acid derivatives. The PH1371 AAT synthesized 1-naphthylglycine from 1-naphthylglyoxylic acid (no. 18 in Fig. 1) but with a low specific activity (1 U/mg).

On the other hand, 12 out of the 32 AATs exhibited the activities with some of the ring-substituted phenylglyoxylic acid derivatives (Tables 3 and 4). Seven out of the 12 AATs were the class I HybATs (group 5; Table 3). All of the tested HybATs were able to synthesize ring-substituted phenylglycine derivatives efficiently. TTE2440 (group 6) was also active in the synthesis of several phenylglycine derivatives with low specific activities (Table 3). Unexpectedly, four of the class IV BCATs (group 8) namely, the AF0933, MJ1008, TTE0933, and TTHA0124 AATs, exhibited the ability to synthesize ring-substituted phenylglycine derivatives, although BCAT MTH1430 did not synthesize any ring-substituted phenylglycine derivatives (Table 4).

The substrate preferences for phenylglyoxylic acid derivatives were different between the HybAT and BCAT

Table 4 The aminotransferase activities of class IV AATs for synthesis of unnatural amino acids

Number of keto acid[a]	Specific activity (U/mg)[b]			
	AF0933	MJ1008	TTE0933	TTHA0124
1	26±2	52±12	55 ±12	6±1
2	2±1	1±0	1±0	0
3	0	1±0	0	0
4	2±1	3±0	5±1	0
5	0	0	0	0
6	7±2	2±0	0	6±4
7	27±1	66±12	72±8	43±11
8	0	0	0	0
9	0	0	0	0
10	0	0	0	0
11	33±7	58±7	159±6	80±8
12	40±6	63±3	137±2	60±9
13	27±3	47±4	101±11	63±8
14	0	0	0	4±0
15	27±2	46±7	76±19	30±5
16	25±2	35±2	42±5	10±3
17	0	0	0	3±1
18	12±1	10±1	3±1	1±0

[a] The number of keto acid corresponds to Fig. 1.
[b] One unit of the enzyme was defined as the amount of enzyme catalyzing the production of 1 μmol amino acid per minute. The values indicate the average of three independent experiments and means ± standard deviation.

enzymes; for instance, most of the HybATs preferred 2-methoxyphenylglyoxylic acid (no. 6 in Fig. 1), 2,3-difluoro-4-ethoxy-phenylglyoxylic acid (no. 14), and 4-bromophenylglyoxylic acid (no. 17), while most of the BCATs preferred 2.3,4,5,6-pentafluorophenylglyoxylic acid (no. 16). Furthermore, the substrate specificities varied across the HybATs. For instance, when the substrate specificities of the TM1131 and PAE2315 HybATs were compared, 2,4- (no. 3) and 2,5-dimethylphenylglyoxylic acid (no. 4) were preferred substrates for the TM1131 HybAT but not for the PAE2315 HybAT. The opposite relationships between these two enzymes were observed with 3,4,5-trifluorophenylglyoxylic acid (no. 15).

Discussion

In the present comprehensive study, we determined the substrate specificities of the 32 thermophilic AATs each of which contained a histidine tag at its C-terminal end. The presence of the histidine tag may not have affected the substrate specificity of each enzyme as the C-terminal part of the polypeptide extends outwardly in all AATs whose three-dimensional structures have been published in Protein Data Bank (PDB, http://www.pdbj.org/). Actually, the comparison of the published data on the substrate specificities of the PH1371 AroAT (Matsui et al. 2000), the APE2248 AlaAT (Koma et al. 2006), and the TTHA0046 AspAT (Okamoto et al. 1996) with the date obtained in this study demonstrated that the C-terminal histidine tags did not affect the substrate specificities of these enzymes.

Thanks to the plasticity of AATs in terms of their substrate specificities, we were interested in synthesizing phenylglycine and its derivatives that have wide applications as chiral building blocks and intermediates for the synthesis of peptidemimetic pharmaceuticals (Davies et al. 2000; Ingallinella et al. 2002) by using thermophilic AATs. We found that some of the HybAT and BCAT enzymes show significant activities to convert ring-substituted phenylglyoxylic acids to ring-substituted phenylglycines. Particularly, the HybAT enzymes exhibited substrate specificities broader than the BCAT enzymes (Tables 3 and 4). There was a good correlation between the ability of AATs to utilize beta-branched amino acids, Val and Ile, and that to synthesize ring-substituted phenylglycines. Both the BCATs and HybATs accepted the beta-branched amino acids, while all other AATs did not.

Since the class IV AATs have been reported to be specific either to D-amino acids or to branched-chain amino acids, it could not be expected that BCATs act on the synthesis of the ring-substituted phenylglycine derivatives. Thermophilic BCATs had several characteristics in common with thermophilic HybAT; both are capable of acting on branched-chain amino acids and efficiently synthesizing many kinds of ring-substituted phenylglycine derivatives. However, there are differences between BCATs and HybATs. All thermophilic BCATs did not synthesize (no. 8, 9, and 10 in Fig. 1) or poorly synthesized (no. 14) phenylglycine derivatives containing a methoxy group at position 4, although some HybATs synthesized 4-methoxyphenylglycine derivatives from these substrates (Tables 3 and 4). In addition, all thermophilic BCATs synthesized *tert*-leucine (specific activity=0.9−2.3 U/mg) similarly to mesophilic BCAT (Taylor et al. 1998), although all HybATs were not able to synthesize such an amino acid. This suggests that the structures of the substrate-binding sites were quite different between BCATs and HybATs. Actually, both AATs are classified into different folding types of PLP-dependent enzymes having different PLP and substrate recognition mechanisms (Hwang et al. 2005).

Recently, the substrate recognition mechanism was proposed in the PH0207 HybAT (Chon et al. 2005) with concomitant determination of the 3D structure (PDB: 1X0M). In addition, the 3D-structure of the TM1131 HybAT (PDB: 1VP4) was also determined. Thus, the putative residues for substrate recognition of the PH0207 HybAT (Val46, Tyr94, Gln130, Tyr154, and Leu299) and of the TM1131 HybAT (Ile 21, Tyr69, Gln107, Tyr131, and Leu278) were suggested to be composed of the similar amino acids and placed in a similar spatial configuration with each other but not in the same manner as the PH1371 AroAT (Tyr59, Phe121, Val122, Met260, and Thr264; PDB: 1DJU; Matsui et al. 2000). In particular, the replacement of Phe121 (PH1371) to Tyr154 (PH0207) or Tyr131 (TM1131) suggests that the "edge-to-face interaction" (Burley and Petsko 1986) observed in the PH1317 AroAT (Matsui et al. 2000) between the aromatic ring of Phe121 and the phenyl group of an aromatic amino acid substrate disappears in the active pocket of HybATs. The amino acids corresponding to Phe121 in the PH1371 AroAT are conserved in the TTE0301 (Phe125) and MJ0001 (Phe117) AATs that acted on phenylalanine but not on phenylglycine derivatives, while Tyr125 is substituted for Phe in the TTE2440 AroAT capable of synthesizing several ring-substituted phenylglycine derivatives. Therefore, the substitution of Tyr for Phe may contribute to broaden the substrate specificities of AATs in the synthesis of ring-substituted phenylglycine derivatives. In the future, comparative structural studies of thermophilic HybATs, BCATs, and AroATs will allow knowledge-based designs of novel AATs with relaxed substrate specificities convenient for the synthesis of nonproteinogenic amino acids.

Acknowledgments We thank professor Isao Shimizu, Consolidated Research Institute for Advanced Science and Medical Care, Waseda University, for providing us the keto acids and Ms. Ayako Hashimoto for

technical assistance. This research work was done in part at the "Center for Practical Chemical Wisdom" supported by Global COE program of Ministry of Education, Culture, Sports, Science and Technology.

References

Ager DJ, Fotheringham IG (2001) Methods for the synthesis of unnatural amino acids. Curr Opin Drug Discov Devel 4:800–807

Ager DJ, Fotheringham IG, Li T, Pantaleone DP, Senkpeil RF (2000) The large scale synthesis of "unnatural" amino acids. Enantiomer 5:235–243

Andreotti G, Cubellis MV, Nitti G, Sannia G, Mai X, Marino G, Adams MWW (1994) Characterization of aromatic aminotransferases from the hyperthermophilic archaeon *Thermococcus litoralis*. Eur J Biochem 220:543–549

Andreotti G, Cubellis MV, Nitti G, Sannia G, Mai X, Adams MWW, Marino G (1995) An extremely thermostable aromatic aminotransferase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Biochim Biophys Acta 1247:90–96

Atiles MW, Dudley EG, Steele JL (2000) Gene cloning, sequencing, and inactivation of the branched-chain aminotransferase of *Lactococcus lactis* LM0230. Appl Environ Microbiol 66:2325–2329

Birolo L, Arnone MI, Cubellis MV, Andreotti G, Nitti G, Marino G, Sannia G (1991) The active site of *Sulfolobus solfataricus* aspartate aminotransferase. Biochem Biophys Act 1080:198–204

Burley SK, Petsko GA (1986) Dimerization energetics of benzene and aromatic amino acid side chain. J Am Chem Soc 108:7995–8001

Chon H, Matsumura H, Shimizu S, Maeda N, Koga Y, Takano K, Kanaya S (2005) Crystal structure of a human kynurenine aminotransferase II homologue from *Pyrococcus horikoshii* OT3 at 2.20 Å resolution. Proteins 61:685–688

Creary X (1987) Reaction of organometallic reagents with ethyl trifluoroacetate and diethyl oxalate. Formation of trifluoromethyl ketones and α-keto esters via stable tetrahedral adducts. J Org Chem 52:5026–5030

Cubellis MV, Rozzo C, Nitti G, Arnone MI, Marino G, Sannia G (1989) Cloning and sequencing of the gene coding for aspartate aminotransferase from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*. Eur J Biochem 186:375–381

Davies AJ, Ashwood MS, Cottrell IF (2000) A facile synthesis of substituted phenylglycines. Synth Commun 30:1095–1102

Fernandez FJ, Vega MC, Lehmann F, Sandmeier E, Gehring H, Christen P, Wilmanns M (2004) Structural studies of the catalytic reaction pathway of a hyperthermophilic histidinol-phosphate aminotransferase. J Biol Chem 279:21478–21488

Hwang BY, Cho BK, Yun H, Koteshwar K, Kim B (2005) Revisit of aminotransferase in the genomic era and its application to biocatalysis. J Mol Catal B Enzyme 37:45–55

Ingallinella P, Fattori D, Altamura S, Steinkühler C, Koch U, Cicero D, Bazzo R, Cortese R, Bianchi E, Pessi A (2002) Prime site binding inhibitors of a serine protease: NS3/4A of hepatitis C virus. Biochemistry 41:5483–5492

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23:403–405

Koma D, Sawai T, Harayama S, Kino K (2006) Overexpression of the genes from thermophiles in *Escherichia coli* by high-temperature cultivation. Appl Microbiol Biotechnol 73:172–180

Marino G, Nitti G, Arnone MI, Sannia G, Gambacorta A, De Rosa M (1988) Purification and characterization of aspartate aminotransferase from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*. J Biol Chem 263:12305–12309

Matsui I, Matsui E, Sakai Y, Kikuchi H, Kawarabayashi Y, Ura H, Kawaguchi S, Kuramitsu S, Harata K (2000) The molecular structure of hyperthermostable aromatic aminotransferase with novel substrate specificity from *Pyrococcus horikoshii*. J Biol Chem 275:4871–4879

Mehta PK, Hale TI, Christen P (1993) Aminotransferases: demonstration of homology and division into evolutionary subgroups. Eur J Biochem 214:549–561

Oda T, Miyajima H, Suzuki Y, Ito T, Yokota S, Hoshino M, Ichiyama A (1989) Purification and characterization of the active serine: pyruvate aminotransferase of rat liver mitochondria expressed in *Escherichia coli*. J Biochem 106:460–467

Okamoto A, Kato R, Masui R, Yamagishi A, Oshima T, Kuramitsu S (1996) An aspartate aminotransferase from an extremely thermophilic bacterium, *Thermus thermophilus* HB8. J Biochem 119:135–144

Pojitkov AE, Efremenko EN, Varfolomeev SD (2000) Unnatural amino acids in enzymes and proteins. J Mol Catal B Enzyme 10:47–55

Sawai T, Koma D, Hara R, Kino K, Harayama S (2007) A high-throughput and generic assay method for the determination of substrate specificities of thermophilic a-aminotransferases. J Microbiol Methods 71:32–38

Scheffer JR, Wang K (2001) Enantioselective photochemical synthesis of a β-lactam via the solid state ionic chiral auxiliary method. Synthesis 8:1253–1257

Schneider G, Käck H, Lindqvist Y (2000) The manifold of vitamin B_6 dependent enzymes. Structure 8:R1–R6

Shin JS, Kim BG (1999) Asymmetric synthesis of chiral amines with w-transaminase. Biotechnol Bioeng 65:206–211

Sonnleitner B, Fiechter A (1983) Advantages of using thermophiles in biotechnological processes: expectations and reality. Trends Biotechnol 1:74–80

Tanaka K, Yamamoto S, Moriya T, Taniguchi M, Hayashi H, Kagamiyama H, Oi S (1994) Aspartate aminotransferase from a thermophilic formate-utilizing methanogen, *Methanobacterium thermoformicicum* strain SF-4: relation to serine and phosphoserine aminotransferases, but not to the aspartate aminotransferase family. J Biochem 115:309–317

Taylor PP, Pantaleone DP, Senkpeil RF, Fotheringham IG (1998) Novel biosynthetic approaches to the production of unnatural amino acids using transaminases. Trends Biotechnol 16:412–418

Venos ES, Knodel MH, Radford CL, Berger BJ (2004) Branched-chain amino acid aminotransferase and methionine formation in *Mycobacterium tuberculosis*. BMC Microbiol 4:39

Word DE, Kengen SWM, van Der Oost J, de Vos WM (2000) Purification and characterization of the alanine aminotransferase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its role in alanine production. J Bacteriol 182:2559–2566