Biochemical Properties and Crystal Structure of a β -Phenylalanine Aminotransferase from *Variovorax paradoxus*

Ciprian G. Crismaru,^a Gjalte G. Wybenga,^b Wiktor Szymanski,^c Hein J. Wijma,^a Bian Wu,^a Sebastian Bartsch,^a Stefaan de Wildeman,^d Gerrit J. Poelarends,^e Ben L. Feringa,^c Bauke W. Dijkstra,^b Dick B. Janssen^a

Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands^a; Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands^b; Center for Systems Chemistry, Stratingh Institute for Chemistry, University of Groningen, Groningen, The Netherlands^c; DSM Pharmaceutical Products, Geleen, The Netherlands^d; Department of Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Groningen, The Netherlands^e

By selective enrichment, we isolated a bacterium that can use β -phenylalanine as a sole nitrogen source. It was identified by 16S rRNA gene sequencing as a strain of *Variovorax paradoxus*. Enzyme assays revealed an aminotransferase activity. Partial genome sequencing and screening of a cosmid DNA library resulted in the identification of a 1,302-bp aminotransferase gene, which encodes a 46,416-Da protein. The gene was cloned and overexpressed in *Escherichia coli*. The recombinant enzyme was purified and showed a specific activity of 17.5 U mg⁻¹ for (S)- β -phenylalanine at 30°C and 33 U mg⁻¹ at the optimum temperature of 55°C. The β -specific aminotransferase exhibits a broad substrate range, accepting *ortho*-, *meta*-, and *para*-substituted β -phenylalanine derivatives as amino donors and 2-oxoglutarate and pyruvate as amino acceptors. The enzyme is highly enantioselective toward (S)- β -phenylalanine (enantioselectivity [E], >100) and derivatives thereof with different substituents on the phenyl ring, allowing the kinetic resolution of various racemic β -amino acids to yield (R)- β -amino acids with >95% enantiomeric excess (*ee*). The crystal structures of the holoenzyme and of the enzyme in complex with the inhibitor 2-aminooxyacetate revealed structural similarity to the β -phenylalanine aminotransferase from *Mesorhizobium* sp. strain LUK. The crystal structure was used to rationalize the stereo- and regioselectivity of *V. paradoxus* aminotransferase and to define a sequence motif with which new aromatic β -amino acid-converting aminotransferases may be identified.

Various nonproteinogenic β -amino acids occur naturally as free metabolites and as components of biosynthetic products (1). The simplest β -amino acid, β -alanine, occurs in carnosine, coenzyme A, and pantothenic acid. Other β -amino acids are present in bioactive peptides, such as the protease inhibitor bestatin, which contains a (2S,3R)-3-amino-2-hydroxy-4-phenylbutanoyl group, and microcystin, a cyclic nonribosomal peptide that acts as a phosphatase inhibitor and contains both an aliphatic and an aromatic β -amino acid moiety (2–4). Further examples are cryptophycins, which are antitumor agents containing an α -methyl- β -alanine group (5), and paclitaxel, an antitumor agent from *Taxus brevifolia* that contains a (2R,3S)-*N*-benzoyl-3-phenylisoserine group that is derived from (R)- β -phenylalanine (6). Other β -amino acids occur as building blocks in β -lactam antibiotics (4, 7) and in antifungal compounds, such as jasplakinolide (8). In view of the growing importance of pharmaceutical compounds containing β -amino acid groups, there is demand for new tools for their production in enantiopure form. Chemical and biochemical reactions yielding enantiopure β -amino acids have recently been reviewed (9). However, the possibilities for biocatalytic processes are scarce, and most existing options rely on kinetic resolution of racemates instead of the more attractive asymmetric conversions (10, 11).

Microorganisms that can synthesize or degrade specific organic compounds are a rich source of enzymes for application in biocatalytic processes. However, the microbial metabolism of β -amino acids has been poorly investigated. Some information is available about the formation of aliphatic β -amino acids, such as β -lysine, β -leucine, or β -glutamate, which can be formed from α -amino acids by catabolic bacterial aminomutases and which are subject to deamination by aminotransferases (ATs) or ammonia

lyases (12–14). The formation of β -alanine in *Escherichia coli* proceeds by decarboxylation of aspartate (15). Aromatic β -amino acids that occur in secondary-metabolite biosynthesis can be formed from proteinogenic α -amino acids by 4-methylideneimidazole-5-one (MIO)-dependent aminomutases that have a biosynthetic function. The uncommon MIO cofactor also plays a key role in catabolic ammonia lyases that act upon α -histidine, α -phenylalanine, and α -tyrosine (16).

Attractive enzymes for asymmetric synthesis are ATs. These are pyridoxal 5'-phosphate (PLP)-dependent enzymes that transfer amino groups between different metabolites and are ubiquitously present in prokaryotic and eukaryotic cells (17, 18). There is evidence for a role of aminotransferases (also known as transaminases) in the biodegradation of β -amino acids by microorganisms. For example, a study on a β -phenylalanine-utilizing strain of *Mesorhizobium* led to the discovery of a transaminase that converts β -phenylalanine into 3-oxo-3-phenylpropionic acid (19–21).

Based on the latest update of the B6 database, which compiles information on PLP-dependent enzymes, seven fold types can be distinguished (22). The ATs occur in fold types I and IV. Typical

Received 15 August 2012 Accepted 12 October 2012

Published ahead of print 19 October 2012

Address correspondence to Dick B. Janssen, d.b.janssen@rug.nl.

C.G.C. and G.G.W. contributed equally to this study.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.02525-12>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.
doi:10.1128/AEM.02525-12

examples of fold type I ATs are aspartate AT, aromatic AT, and ω -ATs (23). According to a broader classification introduced in the 1980s, based on the reaction that is catalyzed, ATs are divided into two subgroups: I, α -ATs, which catalyze transamination of amino groups at the α -carbon, and II, ω -ATs, which perform transamination at a β -amino group, a γ -amino group, or another, more distal amino group of the substrate (24). According to this older, nonphylogenetic classification, all ATs that convert β -amino acids are confusingly considered to be ω -ATs. Enzymes from subgroup II include β -alanine AT, 4-aminobutyrate AT, ornithine AT, acetylornithine AT, and 7,8-diaminopelargonic acid AT (25, 26), as well as a β -transaminase from *Mesorhizobium* sp. strain LUK (*MesAT*) (27).

Application of ATs in biocatalysis has mainly been investigated for the production of proteinogenic amino acids, unnatural amino acids, and various other amines and amino alcohols (28–34). The catalytic activities can be quite high (apparent k_{cat} values up to 50 s^{-1}) (35), and apart from PLP, which is sometimes added, there is no requirement for an external cofactor. Aminotransferases that catalyze synthesis or conversion of β -amino acids could be attractive for biocatalysis if they are enantioselective (36) and stable and exhibit a wide substrate scope. Here, we describe the gene cloning, the biochemical properties, and the three-dimensional (3D) structure of *VpAT*, an aromatic β -amino acid aminotransferase discovered in a strain of *Variovorax paradoxus* isolated from soil. Based on the crystal structure, we offer a rationale for the regio- and stereoselectivity of β -transaminases and identify a signature sequence motif that allows the discovery of new aromatic β -amino acid-converting aminotransferases.

MATERIALS AND METHODS

Chemicals. PLP and *rac*-3-amino-3-phenylpropionic acid (β -phenylalanine) were purchased from Acros Organics, and 2-oxoglutarate (α -ketoglutarate) disodium salt and Brij 35 were purchased from Fluka. *Ortho*-phthalaldehyde (OPA), dimethyl sulfoxide (DMSO), *trans*-cinnamic acid, sodium pyruvate, *rac*- β -leucine, and 2-aminoxyacetic acid (AOA) were purchased from Sigma-Aldrich. (*S*)- β -Phenylalanine and (*R*)- β -phenylalanine were purchased from PepTech Corp. Racemic and enantiomerically pure *ortho*-, *meta*-, and *para*-substituted β -phenylalanines and α -phenylalanines were either purchased from PepTech Corp. or synthesized according to published procedures (37). Other chemicals were purchased as follows: (*R*)-3-amino-butyric acid (Chemcube), (*R*)-3-amino-5-methyl-hexanoic acid (Fluorochem), β -asparagine (Bachem), and *rac*-3-amino-3-(4-hydroxyphenyl)-propionic acid (β -tyrosine) (Innochemie GmbH).

Enrichment of a β -phenylalanine-degrading microorganism. Samples of grassland soil (1 to 2 g) were used as a source of microorganisms. Minimal medium, pH 7.0, contained (per liter) 5.3 g $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, 1.4 g KH_2PO_4 , 0.2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1.0 g $(\text{NH}_4)_2\text{SO}_4$, 1 ml vitamin solution (38), and 5 ml trace element solution (39). In nitrogen-free minimal medium, $(\text{NH}_4)_2\text{SO}_4$ was replaced by Na_2SO_4 and $\text{Ca}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$ was omitted from the trace element solution. Stock solutions of a carbon source (cinnamic acid or glucose) and a nitrogen source (*rac*- β -phenylalanine) were prepared in 50 mM sodium phosphate buffer (pH 7.5). Inoculated flasks containing 50 ml minimal medium supplemented with 5 mM cinnamic acid and 1 mM *rac*- β -phenylalanine were incubated in the dark at 20°C without shaking. After 2 or 3 transfers, pure cultures were isolated on minimal medium agar plates supplemented with β -phenylalanine as a sole nitrogen source and cinnamic acid as a carbon source. A fast-growing strain, named CBF3, was chosen for further study.

Bacterial strains and plasmids. *E. coli* strains VCS257, DH5 α , MC1061, and C41(DE3) were used as hosts for the construction of a cosmid library and a sublibrary, for proliferation of cloned genes, and for

overexpression of protein, respectively. Plasmids pLAFR3, pZER0-2 (Invitrogen), and pET28b+ (Novagen) were used for DNA libraries, subcloning, and overexpression, respectively.

Preparation of cell extracts. To obtain a high yield of cells expressing aminotransferase, strain CBF3 was grown on glucose (10 mM) with *rac*- β -Phe (2 mM) as a nitrogen source. From a 1-liter culture, about 1 g of wet cells was obtained. The cells were washed with 50 mM Tris- SO_4 buffer (pH 8.0) and then suspended in 3 ml of this buffer containing 0.01% (vol/vol) β -mercaptoethanol. Sonication was performed with a Sonic Wibra cell, followed by centrifugation at 15,000 rpm and 4°C for 1 h. The supernatants were used as cell-free extracts (CFEs).

Enzyme assays and amino acid analysis. To test the enzyme activity, an AT assay cocktail was prepared, consisting of 10 mM *rac*- β -phenylalanine, 5 mM 2-oxoglutarate (or pyruvate), and 50 μM PLP in 50 mM MOPS (morpholinepropanesulfonic acid), pH 7.6. Reactions were started by addition of purified enzyme or CFE and incubated at 30°C. The conversion was monitored by taking samples at different times. To a 50- μl sample, 50 μl 2 M HCl was added to quench the reaction, and the mixture was kept on ice for 5 min. Then, 45 μl 2 M NaOH was added to neutralize the pH and 50 μl of water was added for dilution. Immediately prior to injection, 1 μl of sample was mixed with 2 μl OPA solution and 5 μl of 0.4 M NaBO_3 , pH 10.4, in a high-performance liquid chromatography (HPLC) autosampler (40). The OPA solution was prepared by first dissolving 15 mg OPA in 50 μl absolute ethanol, which then was added to a mixture of 4.42 ml of 0.4 M NaBO_3 , pH 10.4, 15 μl of 30% (wt/vol) Brij 35, and 11 μl of β -mercaptoethanol.

Quantification of glutamate, alanine, and β -phenylalanine was performed with a C_{18} OPA Adsorbosphere column connected to a Jasco HPLC system after prederivatization with OPA. Elution was done with 20 mM sodium acetate, pH 5.5, containing 5% (vol/vol) tetrahydrofuran (THF) as eluent A and acetonitrile as eluent B, with a flow rate of 1 ml/min. Eluent A and eluent B were used with a gradient program as follows: 0 to 5 min, 100:0; 5 to 12 min, from 100:0 to 80:20; 12 to 16 min, 80:20; 16 to 24 min, from 80:20 to 40:60; 24 to 28 min, 40:60; from 28 to 30 min, 40:60 to 100:0; from 30 to 35 min, reequilibration at 100:0. Detection was done with a fluorescence detector, using excitation at 350 nm and measuring emission at 450 nm. Retention times for derivatized L- α -glutamate, L- α -alanine, and β -phenylalanine were 2.3 min, 7.7 min, and 23.2 min, respectively. One unit is defined as the amount of enzyme that catalyzes the formation of 1 μmol of L- α -glutamate min^{-1} at concentrations of 10 mM *rac*- β -phenylalanine and 5 mM 2-oxoglutarate. Protein concentrations were determined with Coomassie brilliant blue.

To determine the pH optimum of *VpAT*, Britton-Robinson buffer was used at a pH range from pH 2 to pH 12. The buffer consists of a mixture of 0.04 M H_3BO_3 , 0.04 M H_3PO_4 , and 0.04 M CH_3COOH titrated to the desired pH with 0.2 M NaOH. A sufficient amount of enzyme was added, and its activity was assayed using *rac*- β -phenylalanine (10 mM) as the amino donor and 2-oxoglutarate (5 mM) as the amino acceptor. The initial reaction rates were plotted against pH.

The optimum temperature was determined by measuring the specific activity of *VpAT* in MOPS buffer (50 mM, pH 7.6) at temperatures between 20°C and 65°C. Enzyme was added, and activity was assayed with *rac*- β -phenylalanine (10 mM) as the amino donor and 2-oxoglutarate (5 mM) as the amino acceptor.

To follow the kinetic resolution of *rac*- β -phenylalanines with *VpAT*, separation of enantiomers was performed using a Crownpak CR(+) HPLC column connected to a UV detector (210 nm), as described previously (37). Because of the low solubility of the racemates, they were tested in 50 mM MOPS (pH 7.6) at a concentration of 3 mM, using 5 mM 2-oxoglutarate and 50 μM PLP at 30°C.

Cloning and sequence analysis. All chemicals used in DNA manipulation procedures were purchased from Roche Diagnostics (Mannheim, Germany) and Qiagen NV (Venlo, The Netherlands) and used as recommended by the manufacturer.

The 16S rRNA gene of strain CBF3 was sequenced after PCR amplifi-

cation. For amplification, two universal primers of the 16S rRNA gene were used, namely, 27F as the forward primer (5'-AGAGTTTGATCMTGGCTCAG-3') and 1492R as the reverse primer (5'-GGYTACCTTGTTACGACTT-3'), with genomic DNA as the template (41). The PCR product was sequenced by GATC Biotech, Konstanz, Germany.

Genomic DNA was isolated from bacterial cells as described previously (42) and subjected to paired-end sequencing by Baseclear BV (Leiden, The Netherlands), using an Illumina GAIIx platform to obtain about 50-bp reads as raw data. The DNA reads were assembled into contigs using CLC Genomics Workbench software (CLC Bio).

For library construction, chromosomal DNA was partially digested with Sau3A, resulting in DNA fragments of 15 kb to 40 kb, which were cloned in BamHI-digested and dephosphorylated cosmid vector pLAFR3 (43, 44). Ligated DNA was packaged *in vitro* and transfected to *E. coli* VCS257 according to the recommendations supplied with the kit (Stratagene). Recombinant clones were stored as glycerol stocks at -20°C . Transformants were grown in 96-well microtiter plates (MTPs) containing 1 ml of LB medium and tetracycline (25 $\mu\text{g}/\text{ml}$) at 30°C and 900 rpm. After 24 h, *rac*- β -phenylalanine was added to a final concentration of 7.5 to 10 mM to each well of the MTP. Screening for AT activity was performed by testing for the formation of acetophenone, which is formed by spontaneous decarboxylation of the expected transaminase product 3-oxo-3-phenylpropionic acid. The assay is based on the reaction of acetophenone with 2,4-dinitrophenylhydrazine (DNPH), forming a hydrazone that appears as an orange-red precipitate (45). For screening, MTPs were covered with a paper filter impregnated with a DNPH solution and incubated for 2 days at 30°C and 900 rpm.

For subcloning, the vector pZER0-2 and the pLAFR3-positive clone were digested with EcoRI, and fragments were ligated. DNA was transformed to *E. coli* DH5 α ElectroMax cells (Invitrogen), and screening was performed as described for the pLAFR3 cosmid library. A 5-kb insert containing the CBF3 *VpAT*-encoding gene was isolated and sequenced by primer walking (GATC Biotech). Sequence comparisons were performed with Clustal Ω (46) and Geneious Pro software version 5.5 (47).

For amplification of the entire *VpAT* gene, two primers were designed, a forward primer (5'-GCGCGCATATGACCCATGCCGCCATAG-3') (the NdeI site is underlined; the start codon is in boldface) and a reverse primer (5'-CGCGCGCTCGAGTTAGTTCGCGCGGGGACAGC-3') (the XhoI site is underlined; the stop codon is in boldface). The 1.3-kb PCR product was cloned using the NdeI and XhoI sites of the pET28b+ plasmid. The *MesAT* I56V/A312S/M414F triple mutant was expressed and purified as previously reported for the *MesAT* wild type (WT) (48).

The *VpAT* R41A mutant was prepared by site-directed mutagenesis (QuikChange; Stratagene). The R41A forward primer (5'-GGAGCCAACAGCGCTCCGTGCTGTTC-3') and R41A reverse primer (5'-GAACAGCACGGAGGCGCTGTTGGCTCC-3') (mutated codons are in boldface) were used according to the manufacturer's recommendations. All constructs were confirmed by sequencing (GATC Biotech AG, Konstanz, Germany). The *MesAT* I56V/A312S/M414F triple mutant was prepared by site-directed mutagenesis (QuikChange; Stratagene) in three rounds.

Overexpression and purification of *VpAT* in *E. coli*. The pET28b+ construct containing the gene for *VpAT* was used to produce the enzyme with an *N*-terminal His₆ tag (MGSSHHHHHH) followed by a 10-amino-acid linker (SSGLVPRGSH) in *E. coli* C41 (DE3). Cells were grown at 37°C in LB medium with 50 $\mu\text{g}/\text{ml}$ kanamycin. Expression of *VpAT* was induced by adding 0.8 mM isopropyl- β -D-thiogalactopyranoside (IPTG) to the growing cells when the optical density at 600 nm (OD_{600}) reached 0.6. Cultivation was continued for 16 h at 28°C and 170 rpm. Cells were obtained by centrifugation and disrupted by sonication at 4°C , followed by centrifugation for 45 min at 15,000 rpm to obtain CFE. The enzyme was purified in two steps using immobilized metal affinity chromatography (IMAC) (HisTrap HP column; 5 ml; GE Healthcare) and ion-exchange chromatography (IEXC) (Q-Sepharose HP column; 5 ml; GE Healthcare). In the case of IMAC, *VpAT* was eluted at a flow rate of 1 ml/min with 15 column volumes of a linear gradient of 0 to 0.5 M imidazole in a buffer

containing 20 mM Tris-HCl, pH 8.0, 0.5 M NaCl, and 0.01% (vol/vol) β -mercaptoethanol, whereas for IEXC, elution was performed with 15 column volumes of a gradient of 0 to 1 M NaCl in a buffer containing 20 mM Tris-HCl, pH 8.0, and 0.01% (vol/vol) β -mercaptoethanol. For use in crystallization experiments, the fractions containing active enzyme were pooled, concentrated (Ultracel 30K MWCO; Amicon), and applied to a Superdex 200 10/300 GL size exclusion chromatography column (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 7.5, containing 200 mM NaCl. After elution, the fractions corresponding to the protein peak were pooled, concentrated (Amicon), and dialyzed overnight against a buffer containing 20 mM Tris-HCl, pH 7.5, and 10 μM PLP and concentrated to 20 mg/ml.

Protein crystallization. Crystallization experiments were set up at 20°C , and a single crystal was obtained under 0.02 M sodium/potassium phosphate, 0.1 M 1,3-bis[tris(hydroxyl-methyl)-methylamino]-propane, pH 6.5, and 20% (wt/vol) polyethylene glycol (PEG) 3350K. Crystals of *VpAT* grew within a week, after which they were transferred to a cryoprotection solution consisting of the mother liquor with 20% (vol/vol) glycerol. This was achieved in 4 steps of 5 min each, starting with a solution containing 2% (vol/vol) glycerol, followed by solutions with 5, 10, and finally 20% glycerol. Crystals from the last solution were cooled in liquid nitrogen. For the AOA binding study, the same steps were followed, but with the cryoprotection solutions supplemented with 2, 5, 10, and 20 mM AOA.

Diffraction data collection and processing. Diffraction data were collected at beamline ID14-4 of the European Synchrotron Radiation Facility (ESRF) (Grenoble, France). Indexing and integration of reflections was done using XDS (49), and scaling and merging of the data were achieved using SCALA (50) from the CCP4 software suite (51). For molecular replacement, the Phaser program (52) was used with *MesAT* (Protein Data Bank [PDB] code 2YKU [48]) as the input model. The resulting model structure was subjected to successive rounds of automatic model building using ARP/wARP (53) at 1.7-Å resolution, followed by manual model building and manipulation in Coot (54). Refmac5 was used for refinement of the atomic coordinates and atomic B factors (55). After refinement, the model quality was validated with MolProbity (56). Heterocompound coordinate files were obtained from the HIC-Up server (57), while the PRODRG2 server (58) was used to generate the stereochemical restraints. Structural homologues of *VpAT* were obtained with the Dali server (59). PISA from the CCP4 software suite was used for protein interface analysis (60), while PyMOL (<http://www.pymol.org/>) was used to make images of the protein structure. Data collection and refinement statistics are given in Table S2 in the supplemental material.

Substrate docking. Docking of the PLP-(S)- β -phenylalanine intermediate was carried out with Rosetta software, which allows both side chain and backbone flexibility during docking (61, 62). Rosetta's redesign specificity application (63) was used (without permitting mutations); the application allowed optimization of the surrounding protein structure to bind the (S)- β -phenylalanine intermediate while preserving the known binding orientation of the PLP (*MesAT*; PDB code 2YKY). To model the flexibility of the protein, Monte Carlo optimizations of side chain rotamers (collections of thermodynamically accessible conformations) were carried out three times. Each of these optimization rounds was followed by an energy minimization, which also allowed backbone atoms to move. Residues up to 8 Å from the intermediate were allowed to change conformation. Rotamers for the PLP-(S)- β -phenylalanine intermediate were prepared with Yasara (<http://www.yasara.org/>) (64). AM1-BM3 charges of the PLP intermediate were assigned by OEchem (65). A total of 1,040 docking runs were carried out, among which the lowest-energy solution was selected.

Accession numbers. The 16S rRNA gene sequence of *V. paradoxus* CBF3 has been deposited at GenBank under accession number JN990697. The sequences of the DNA contig containing the *VpAT*-encoding gene and of the *VpAT* enzyme were deposited at EMBL under accession numbers HE608883 and CCE46017, respectively. Atomic coordinates and

FIG 1 Partial sequence alignment of VpAT with other aminotransferases made with Clustal Ω (46). The structures of MesAT and SeGSAM are known, and α -helices have been assigned based on these structures. For PoGSAM, they have been assigned on the basis of sequence conservation. Since the structure of AdbpAT has not been elucidated, and since the sequence identity to VpAT is low, secondary-structural elements are not indicated. The proposed aromatic β -transaminase signature sequence is boxed, and its residues are highlighted in dark gray. Light gray shading indicates the boundaries of secondary structure elements. The black squares indicate relevant residues for substrate and cofactor binding. Proteins: VpAT, β -transaminase of *V. paradoxus* (this study) (CCE46017.1; PDB code 4AO9); PoGSAM, glutamate-1-semialdehyde 2,1-aminomutase of *Polaromonas* sp. strain JS666 (ABE43415.1); MesAT, β -transaminase of *Mesorhizobium* sp. strain LUK (ABL74379.1; PDB code 2YKY); SeGSAM, glutamate-1-semialdehyde 2,1-aminomutase from *S. elongatus* (ABB56677.1; PDB code 2HOZ); AdbpAT, β -alanine:pyruvate transaminase of *A. denitrificans* (AAP92672.1). The numbers 6, 11, 17, 23, 29, 35, 40, and 46 refer to proteins that are listed in Table S3 in the supplemental material.

structure factors have been deposited in the Protein Data Bank (<http://www.pdb.org>) under accession codes 4AO9 for the VpAT holoenzyme and 4AOA for VpAT complexed with 2-aminoxyacetic acid.

RESULTS

Isolation of a β -phenylalanine-degrading bacterium. The isolation of a bacterial strain possessing β -phenylalanine transaminase activity was carried out by an enrichment procedure using *rac*- β -phenylalanine as the sole nitrogen source. After growth was observed, a pure culture was obtained by repeated transfer to fresh medium and streaking onto minimal medium plates supplemented with β -phenylalanine and *trans*-cinnamic acid. The most rapidly growing strain (named CBF3) was selected for further investigation. The 16S rRNA gene (1,517 bp) of strain CBF3, identified by PCR and paired-end genome sequencing, has 99% sequence identity to the 16S rRNA genes of *V. paradoxus* S110 (CP001635.1) and *V. paradoxus* EPS (CP002417.1) (66). This affiliates strain CBF3 with the species *V. paradoxus*.

Activity assays with HPLC analysis showed that strain CBF3 possesses a β -phenylalanine AT activity that converts β -phenylalanine and 2-oxoglutarate (or pyruvate) to 3-oxo-3-phenylpropionic acid and L- α -glutamate (or L- α -alanine). Attempts to purify VpAT using wild-type CBF3 as the source of enzyme failed due to low protein recovery after partial purification.

Isolation of the VpAT gene. A DNA library, consisting of about 4,000 clones, was constructed in cosmid pLAFR3. The insert

size of the DNA fragments was between 15 and 25 kb. The genome size of *V. paradoxus* is about 6.7 Mb (67); thus, the number of clones obtained was sufficient for 9-fold coverage of the whole genome. Screening individual clones for AT activity yielded two positive clones, one of which was investigated further. The insert was subcloned into vector pZErO-2, and rescreening for acetophenone formation yielded six positive hits. Restriction analysis identified a shared 5-kb EcoRI fragment that likely contains the AT-encoding gene.

DNA sequence analysis showed the presence of a 1,302-bp gene encoding an aminotransferase, which was subsequently transferred to the pET28b+ expression vector. The encoded 434-amino-acid protein has a theoretical pI of 6.06 and a calculated molecular mass of 46.42 kDa (http://web.expasy.org/compute_pi/). The sequence of a 20.6-kb contig found by paired-end genome sequencing indicated that around the VpAT gene there were no regulatory regions or open reading frames related to other enzymes of amino acid metabolism.

Protein sequence analysis. Sequence analysis showed that VpAT has 55% sequence identity to glutamate-1-semialdehyde-2,1-aminomutase from *Polaromonas* sp. strain JS666 (PoGSAM), which has been reported to have activity with aromatic β -amino acids (ABE43415.1) (Fig. 1), 51% sequence identity to the aromatic β -phenylalanine aminotransferase (MesAT) (ABL74379.1) from *Mesorhizobium* sp. LUK, and 35% sequence identity to a

glutamate-1-semialdehyde-2,1-aminomutase from *Synechococcus elongatus* (SeGSAM) (ABB56677.1). These data suggest that VpAT, just like MesAT and PoGSAM, is a fold type I aminotransferase, and furthermore, that VpAT belongs to subgroup II transaminases, which is based on an enzyme's substrate specificity (27, 48, 68). VpAT has 19% sequence identity to a transaminase from *Alcaligenes denitrificans* that has been reported to have activity toward aliphatic β -amino acids (AdbpAT) (AAP92672.1) (21). A sequence comparison between VpAT, MesAT, PoGSAM, and AdbpAT shows that the amino acid residues that are involved in cofactor binding are conserved, in addition to several residues that, based on the structures of VpAT and MesAT, are involved in substrate and cofactor binding (Fig. 1).

Purification of VpAT expressed in *E. coli*. The recombinant protein was overproduced with an N-terminal His₆ tag in *E. coli* strain C41(DE3). VpAT was mainly present as a soluble protein. The enzyme was purified by three chromatography steps. Size exclusion chromatography indicated a molecular mass of approximately 100 kDa, suggesting that VpAT exists as a dimer in solution. The purified protein showed a single band of about 48 kDa in an SDS-PAGE gel (see Fig. S1 in the supplemental material). The overall yield from 1 liter of culture was 40 to 50 mg (see Table S1 in the supplemental material). The specific activity of the purified enzyme was 17.5 U mg⁻¹ at 30°C, corresponding to a k_{cat} of 11.8 s⁻¹ per monomer. The VpAT R41A mutant was overexpressed and purified under the same conditions as the wild-type enzyme, resulting in similar amounts of purified protein.

Catalytic properties. The pH activity profile of VpAT showed that the enzyme has high activity over a broad pH range (4 to 11.2) at 30°C. The optimum temperature of the enzyme was tested by measuring the specific activity at temperatures between 20°C and 65°C. VpAT exhibits a maximum specific activity of 33 U mg⁻¹ at 55°C, which is about 2-fold higher than the specific activity at 30°C (see Fig. S2 in the supplemental material). These data show that VpAT is more active toward (S)- β -phenylalanine than other β -transaminases reported previously (21, 27, 68). The activity of VpAT with pyruvate is 85% of that with α -ketoglutarate as the amino acceptor.

The relationship between the reaction rate and the substrate concentration displayed Michaelis-Menten kinetics with substrate inhibition (69). The apparent K_m and k_{cat} values (per monomer) for (S)- β -phenylalanine in the presence of 10 mM α -ketoglutarate were 1.5 mM and 11.8 s⁻¹, respectively, with an apparent substrate inhibition constant (K_i) of 40.3 mM. When using 10 mM (S)- β -phenylalanine, the apparent K_m and k_{cat} for α -ketoglutarate were 0.3 mM and 10.6 s⁻¹, respectively, with an apparent K_i of 82.4 mM.

Substrate scope and enantioselectivity (E) of VpAT. To investigate the substrate scope of the enzyme, a number of β -amino acids and α -amino acids were tested (Table 1). The results show that VpAT acts exclusively on β -amino acids and that the enzyme prefers aromatic β -amino acids (Table 1, entries 1 to 3 and 8 to 20) over aliphatic β -amino acids (Table 1, entries 21 to 27). The most favorable substituent positions on the phenyl ring of β -phenylalanine for activity were *para*- and *meta*-substituted analogues, while *ortho*-substituted analogues were poorly converted. The β -phenylalanine derivatives bearing linear alkyl substituents at the *para* position, i.e., ethyl or propyl (Table 1, entries 13 to 14), were converted more efficiently than other substituted substrates (Table 1, entries 12 and 16 to 20). However, introduction of a

branched alkyl substituent, i.e., *iso*-propyl, at the *para* position results in a decrease of activity (Table 1, entry 15). The enzyme showed no activity for α -phenylalanine or its *meta*-chloro-, *meta*-bromo-, and *para*-hydroxy-substituted (α -tyrosine) derivatives, which contrasts with MesAT, for which activity was reported with α -phenylalanine and other α -amino acids (27). However, we could not confirm the activity of MesAT toward α -phenylalanine.

VpAT further preferentially converts the (S)-enantiomers of aromatic β -amino acids (Table 1, entries 1 to 3, 9, 10, 12, and 16) and the (R)-enantiomers of aliphatic β -amino acids (which have a similar configuration of functional groups around the chiral carbon atom but a change in Cahn-Ingold-Prelog priority). However, the activity of VpAT toward several aliphatic β -amino acids was quite low (Table 1, entries 22 to 23). We also investigated the regio- and enantioselectivity of VpAT with several racemic substrates (Table 1, entries 1, 9, 10, 12, and 16). The enzyme appeared highly enantioselective toward the (S)-enantiomers of β -phenylalanine and its *meta*- and *para*-ring-substituted derivatives, thereby producing highly pure (R)-enantiomer preparations with high enantiomeric excess (*ee*; >95%).

3D structure of VpAT. The crystal structure of the VpAT holoenzyme was solved by molecular replacement, using the structure of MesAT (PDB code 2YKU) (48) as a search model, and subsequently refined at 1.5-Å resolution. VpAT crystallizes in space group $P2_12_12_1$ with two molecules (chains A and B) per asymmetric unit, which are related by noncrystallographic 2-fold symmetry. Chains A and B are very similar to each other, with root mean square deviation (RMSD) values of C- α atom positions of approximately 0.2 Å. For both monomers, electron density is absent for the first 21 N-terminal residues. This includes the N-terminal His₆ tag, 10 amino acids included as a linker in the expression construct, and the first amino acid (methionine) of the VpAT polypeptide. The interface between the monomers buries an area of about 4,400 Å², which equals a quarter of the total surface area of the monomer pair (Fig. 2). Since the active sites of chains A and B are located at this interface and shaped by residues from both monomers, this suggests that the two monomers in the asymmetric unit form the functional VpAT dimer observed in solution (see above).

The structure of the VpAT monomer. The VpAT monomer has a curved shape and consists of a PLP-binding domain and a domain formed by the N and C termini of the polypeptide chain (NC domain; residues 1 to 86 and 320 to 434) (Fig. 2). The two domains line a cleft into which the PLP cofactor protrudes. The monomer consists of 11 α -helices of at least 2 turns and 13 β -strands that form a central 7-stranded mixed β -sheet in the PLP-binding domain and two 3-stranded antiparallel β -sheets in the NC domain. The overall structure of the enzyme is similar to that of aspartate aminotransferase, the archetype of a fold type I aminotransferase (*z*-score, 22; RMSD, 4.3 Å for 308 C- α atoms; 17% sequence identity; PDB code 1BKG) (70, 71), but is most similar to that of MesAT (*z*-score, 60; RMSD, 1.1 Å for 429 C- α atoms; 50% sequence identity; PDB code 2YKU) (48).

The PLP cofactor. Each VpAT dimer contains one PLP cofactor per monomer, which is positioned close to the surface (Fig. 3). The PLP cofactor is anchored to the protein via a covalent imine bond, with the ϵ -amino group of residue K267, which is conserved in the sequence alignment of homologous PLP-dependent enzymes. In addition, the three phosphate oxygen atoms are at hydrogen-bonding distance from the amide protons of residues

TABLE 1 Substrate range of VpAT^a

Type	Entry	Amino donor	Relative activity (%)	Conversion (%)	ee (%), E
I (β -phenylalanine derivatives)	1	<i>rac</i> - β -Phenylalanine	100	49	>99 (R), >200
	2	(S)- β -Phenylalanine	100	99 ^c	
	3	(R)- β -Phenylalanine	NA ^b	— ^d	
	4	<i>rac</i> -2-Methyl- β -phenylalanine	5	19	ND ^e
	5	<i>rac</i> -2-Fluoro- β -phenylalanine	15	26	ND
	6	<i>rac</i> -2-Chloro- β -phenylalanine	4	32	ND
	7	<i>rac</i> -2-Bromo- β -phenylalanine	4	25	ND
	8	<i>rac</i> -3-Methyl- β -phenylalanine	54	64	ND ^e
	9	<i>rac</i> -3-Fluoro- β -phenylalanine	48	48	>98 (R), >200
	10	<i>rac</i> -3-Chloro- β -phenylalanine	56	47	>95 (R), >100
	11	<i>rac</i> -3-Bromo- β -phenylalanine	53	99	ND
	12	<i>rac</i> -4-Methyl- β -phenylalanine	100	53	>95 (R), >200
	13	<i>rac</i> -4-Ethyl- β -phenylalanine	134	47	ND
	14	<i>rac</i> -4-Propyl- β -phenylalanine	126	48	ND
	15	<i>rac</i> -4- <i>iso</i> -Propyl- β -phenylalanine	61	25	ND
	16	<i>rac</i> -4-Trifluoromethyl- β -phenylalanine	73	49	>95 (R), >100
	17	<i>rac</i> -4-Fluoro- β -phenylalanine	87	59	ND
	18	<i>rac</i> -4-Bromo- β -phenylalanine	62	61	ND
	19	<i>rac</i> -4-Nitro- β -phenylalanine	32	6	ND
	20	<i>rac</i> -4-Hydroxy- β -phenylalanine	61	39	ND
II (other β -amino acids)	21	<i>rac</i> -3-Amino-butyric acid	1	6	ND
	22	(R)-3-Amino-butyric acid	3.1	11	
	23	(R)-3-Amino-5-methyl-hexanoic acid	48	75	
	24	<i>rac</i> - β -Asparagine	1	4	ND
	25	<i>rac</i> - β -Leucine	38	51	ND
	26	β -Alanine	NA	—	
	27	β -Glutamic acid	NA	—	
III (α -amino acids)	29	(S)- α -Phenylalanine	NA	—	
	30	<i>rac</i> -3-Chloro- α -phenylalanine	NA	—	
	31	<i>rac</i> -3-Bromo- α -phenylalanine	NA	—	
	32	(S)- α -Tyrosine	NA	—	

^a Reaction mixtures (0.4-ml volume) contained an amino donor (3 mM) and α -ketoglutarate (5 mM) as the acceptor. Initial rates were measured, and conversion was determined after 16 h at 30°C. The activity for β -phenylalanine, corresponding to 17.5 U mg⁻¹, was taken as 100%.

^b NA, no activity (less than 0.001 U mg⁻¹).

^c Conversion after 0.5 h.

^d —, no conversion observed after 16 h.

^e ND, not determined.

G132, T133, and T300* (monomer B). Two of the phosphate oxygen atoms also make a hydrogen bond to a water molecule, which in turn is hydrogen bonded to peptide backbone atoms. The phosphate ester oxygen atom is not involved in hydrogen bond forma-

tion. The PLP cofactor is positioned between V242 (at the *Si* face of the C-4' of the aldimine) and Y159 (at the *Re* face of the C-4' of the aldimine), while the pyridine nitrogen makes a hydrogen bond with the side chain of D240. The interactions of the PLP cofactor

FIG 2 Structures of the monomer and dimer of VpAT. (A) Monomer of VpAT with the PLP cofactor (magenta) positioned in the cleft between the PLP-binding domain (left half of the monomer, containing the 7-stranded mixed β -sheet) and the NC domain (right half of the monomer, containing the two 3-stranded antiparallel β -sheets and the N and C termini of the polypeptide chain, indicated by N and C, respectively). (B) Two monomers of VpAT assemble into a dimer, here viewed along the 2-fold noncrystallographic axis relating the two monomers. The two PLP cofactors are shown in magenta and are linked to the polypeptide backbone via a Schiff base with K267. The α -helices are blue, the β -strands are yellow, and the loops are red. N and C indicate the N and C termini, respectively.

FIG 3 Wall-eyed stereo presentation of VpAT bound with an inhibitor and a docked intermediate. (A) The PLP cofactor (yellow) is shown bound to AOA (yellow). Monomer A of VpAT is shown in light blue, whereas monomer B is shown in orange. (B) AOA bound in VpAT. The holoenzyme without inhibitor is shown in green. The X-ray structure with AOA bound is shown with different colors for monomer A (light blue), monomer B (orange), and the inhibitor (yellow). (C) X-ray structure of holo-VpAT with docked (*S*)- β -phenylalanine. The X-ray structure of the holoenzyme is shown in green, while the modeled structure with docked (*S*)- β -phenylalanine is in light blue for monomer A and orange for monomer B. The docked intermediate is shown in yellow. The side chains of residues interacting with the docked intermediate and the backbone atoms of A212 are also shown. R41 (P pocket) is the residue that binds the α -carboxylate of β -phenylalanine. In the O pocket, R398 is the arginine switch residue (compare panels B and C).

with residues K267, V242, Y159, and D240 keep the PLP cofactor fixed in the active site (Fig. 3A).

The binding of 2-aminooxyacetic acid. AOA, a mimic of β -alanine, is a known inhibitor of aminotransferases (72) and also inhibits *VpAT*. To investigate how AOA binds in *VpAT*, enzyme crystals were soaked with AOA. The structure of *VpAT* with AOA (Fig. 3A and B) shows that the amino group of the inhibitor covalently binds to the C-4A atom of the PLP cofactor, as also described for aspartate aminotransferase (73) and recently for *MesAT* (48). The ether-oxygen atom (O χ 1) of AOA is positioned close (3.0 Å) to the ϵ -amino group of residue K267. The carboxylate group binds via a salt bridge to the N ϵ - and N η 2-atoms of R41. The binding of AOA in the active site of *VpAT* is similar to the way in which AOA binds in the active site of *MesAT* (PDB code 2YKV). AOA inhibits aminotransferase activity because the amine-ether oxygen bond of AOA cannot be weakened by K267, and hence, hydrolysis and formation of the pyridoxamine 5'-phosphate (PMP) intermediate does not occur (Fig. 3B).

Docking of PLP-(S)- β -phenylalanine in the active site of *VpAT*. To investigate how an aromatic β -amino acid, such as (S)- β -phenylalanine, binds in the active site of *VpAT*, initially, a structure was determined after a *VpAT* crystal had been soaked with (S)- β -phenylalanine. However, this structure only revealed the presence of a PMP intermediate in the active site of the enzyme. Since trapping of (S)- β -phenylalanine in the active site of *VpAT* apparently had not occurred, the external (S)- β -phenylalanine aldimine intermediate was modeled into the active site of *VpAT* using Rosetta, which allows docking with flexible backbone and side chains. In the docked structure, the aromatic ring of the substrate is bound between V43, Y76, Y159, and A212 (Fig. 3C), similar to the way (S)- β -phenylalanine is bound to *MesAT* (48). We previously termed this region the O pocket, because it is positioned on the 3-hydroxyl side of the PLP cofactor (48) (see below). Similar to the conformational changes observed in *MesAT* upon formation of the external aldimine intermediate with (S)- β -phenylalanine (48), in *VpAT* the arginine switch residue R398 in the O pocket is also displaced during docking, freeing space for the phenyl ring of the substrate. Instead, the R398 side chain N η 2 atom forms a hydrogen bond to the carbonyl oxygen of A212. Also, V43 and Y159 have reoriented to better accommodate the docked external aldimine intermediate. In the absence of (S)- β -phenylalanine, the docking procedure resulted in minor shifts (less than 0.5 Å) of the backbone atoms around the active site. While the *ortho* hydrogen atoms of the phenyl ring of the bound intermediate are buried, the *para* and *meta* hydrogen atoms of the phenyl ring are solvent exposed in the docked structure. The carboxylate of the docked (S)- β -phenylalanine is buried in a region we term the P pocket, which is at the phosphate group side of the PLP cofactor. Here, residue R41 is directly involved in binding of (S)- β -phenylalanine by making a salt bridge to the carboxylate group (Fig. 3C). The importance of R41 for the binding of (S)- β -phenylalanine is underscored by the observation that the R41A mutant was inactive when tested with (S)- β -phenylalanine as amino donor and α -ketoglutarate as amino acceptor.

Identification of aromatic β -amino acid aminotransferases. The *VpAT* structure with bound AOA, a β -alanine mimic, and the docking results with β -phenylalanine show that R41 is important for binding the α -carboxylate of β -amino acids, as is the corresponding residue R54 in *MesAT* (48). The side chain of R41 interacts with E75, which may assist the function of R41 by orienting its

side chain and shaping the carboxylate-binding P pocket. We therefore hypothesized that the presence of an R41-E75 pair in an aminotransferase protein sequence indicates enzymatic activity with β -amino acids. A comparison of the crystal structures of *VpAT* and *MesAT* showed that the R41-E75 salt bridge is conserved in *MesAT* (R54-E88). In addition, a sequence alignment between *VpAT* and a glutamate-1-semialdehyde-2,1-aminomutase from *Polaromonas* sp. JS666 (*PoGSAM*), for which no structure is available but which has activity toward aromatic β -amino acids (68), also showed conservation of R41 and E75 (as R51 and E85, respectively) (Fig. 1). In contrast, the crystal structure of *S. elongatus* glutamate-1-semialdehyde-2,1-aminomutase (*SeGSAM*) (PDB code 2HOZ) (74), which is structurally very similar to that of *VpAT* (z-score, 47; RMSD, 1.8 Å for 404 C- α atoms; 35% sequence identity), shows that no Arg and Glu residues equivalent to R41 and E75 of *VpAT* are present, in agreement with the enzyme's lack of activity on β -amino acids.

We subsequently tested whether R41 and E75 could be used to define a sequence motif with which aromatic β -amino acid aminotransferases can be distinguished from aminotransferases with a similar fold but with a different substrate specificity. For this, BLAST searches were performed using the *VpAT*, *MesAT*, and *PoGSAM* (ABE43415.1) protein sequences. For each query sequence, 2,000 homologous sequences were retrieved with $\geq 18\%$ sequence identity. After omitting duplicates, the resulting 2,222 sequences were included in a multiple-sequence alignment and inspected for residues that are important for substrate binding.

We used the P-pocket R41-E75 pair as the starting point to identify a motif with which aminotransferases with activity toward aromatic β -amino acids can be distinguished from aminotransferases with a similar fold but a different activity. The R41-E75 starting pair was subsequently broadened by including the O-pocket residues V43 and Y76, since the equivalent residues in *MesAT* (I56 and Y89) were shown to be involved in substrate binding. At position 43, Val, Ile, and Ala were allowed, but hydrophobic residues with more bulky side chains, such as Leu, Phe, Tyr, and Trp were excluded, since they were expected to prevent aromatic substrate binding for steric reasons. At position 75, apart from Glu, we also allowed Asp, Gln, and Asn, since they may still be able to interact with the side chain of R41. At position 76, Phe and Trp were also allowed, since they are frequently found in the same position as Tyr in other fold type I aminotransferases. We next extended the motif with conserved residues at positions that appear structurally important by adding a conserved Pro at position 50, which is in the center of a hydrophobic cluster in the *VpAT* and *MesAT* structures, and an Asp at position 65 and a Gly at position 66, which form a β -turn, in which the Asp is hydrogen bonded to A23 and S24, residues that are located in the α -helix that precedes loop 1.

The motif that was finally defined, R-X-[AVI]-X(6)-P-X(14)-D-G-X(8)-[EDNQ]-[YFW] (Fig. 1), retrieved 46 unique sequences (see Table S3 in the supplemental material) that we hypothesize may represent fold type I aminotransferases with activity toward aromatic β -amino acids.

DISCUSSION

Catalytic properties and enantioselectivity of *Variovorax* AT. Isolation of microorganisms that use β -phenylalanine as a sole nitrogen source yielded a Gram-negative bacterium of the species

V. paradoxus that produces a β-amino acid selective aminotransferase. The gene encoding this enzyme was cloned, sequenced, and overexpressed to a high level in *E. coli*. Sequence analysis showed that the enzyme is a fold type I AT, just like several other subgroup II ω-ATs. *VpAT* shows the highest sequence identity with biochemically characterized β-phenylalanine ATs (27, 68), including the *Mesorhizobium* aminotransferase (*MesAT*), described by Kim et al. (27), the structure of which was recently solved (48). *MesAT* and *VpAT* have similar activities with α-ketoglutarate and pyruvate as amino acceptors (27) but differ from α-transaminases, which generally show a clear preference for either pyruvate, α-ketoglutarate, or oxaloacetate (75–77). Striking differences between *MesAT* and *VpAT* are the facts that *VpAT* has a higher catalytic activity, is active over a very broad range of pH values, and accepts only β-amino acids. On the other hand, *VpAT* shows the same enantioselectivity for (S)-β-phenylalanine and a very high enantioselectivity ($E > 100$), just like *MesAT* (19) and *PoGSAM* (68). This allows a biotechnologically relevant kinetic resolution of racemic mixtures, from which the remaining (R)-β-phenylalanines can be obtained with very high *ee* (Table 1, entries 1, 9, 10, 12, and 16).

Although the differences in activity are difficult to explain, a comparison of the residues surrounding the active site of *VpAT* with those of *MesAT* shows that V43, S298* (monomer B), and F400, which line the hydrophobic O pocket of the active site of *VpAT*, differ from the residues found at these positions in *MesAT* (I56, A312*, and M414). A triple mutant of *MesAT* (I56V/A312S/M414F) was generated and showed a 2-fold increase (3.3 U mg^{-1}) in activity with (S)-β-phenylalanine compared to the *MesAT* WT (1.6 U mg^{-1} [27]), indicating an important role of these residues in activity with (S)-β-phenylalanine.

Structure of *VpAT* and architecture of the active site. The structure of *VpAT* revealed the overall topology of the enzyme and the position of the cofactor and provided insight into the architecture of the active site of *VpAT*. Comparison with the structure of *MesAT* shows that the two enzymes are structurally similar (RMSD, 1.0 Å for 423 C-α atoms; 51% sequence identity) and confirms that they both belong to the fold type I aminotransferase family (70). As in *MesAT*, both monomers in *VpAT* contribute to the binding of the PLP cofactor, suggesting that dimerization is essential for catalytic activity (Fig. 3).

Previously, we proposed to use the notation P pocket (pointing in the same direction as the PLP phosphate) and O pocket (at the side of the hydroxyl substituent of the PLP) to define substrate binding sites of amino acid-converting fold type I ATs (48). This notation is more generally applicable than using L(arge) and S(mall) pocket (78), since the size of a pocket at a topologically conserved position may vary, whereas the O pocket and the P pocket are topologically fixed (Fig. 4).

Computer-aided docking of (S)-β-phenylalanine suggested binding of β-phenylalanine with the carboxylate group in the P pocket, which is opposite to the binding mode of the α-phenylalanine analogue 3-phenylpropionate in the active site of aromatic amino acid aminotransferase from *Paracoccus denitrificans* (PDB code 1AY8) (79) but similar to the binding mode of β-phenylalanine in *MesAT* (48). The docking results further indicated that several residues in the O pocket, including R398, must rearrange to enable binding of the phenyl ring of β-phenylalanine (Fig. 3C). This switch was also observed for *MesAT* complexed with (S)-β-phenylalanine (PDB code 2YKY) and 2-oxoglutarate (PDB code

FIG 4 Schematic representation of (S)-β-phenylalanine (structure in bold-face) bound to PLP (external aldimine) docked in the active site of *VpAT*. A number of relevant amino acid residues for binding and catalysis are shown.

2YKX) and appears necessary to allow binding of both the phenyl group of (S)-β-phenylalanine AT and the α-carboxylate of 2-oxoglutarate (48). In class I ATs accepting aromatic α-amino acids, a so-called arginine switch has also been identified (e.g., R292 in AroAT), but here, the arginine switch and cosubstrate non-α-carboxylate-binding site are located in the P pocket, in agreement with the reverse mode of substrate binding of β-Phe in *VpAT* and *MesAT* compared to α-phenylalanine in AroAT (80). The docking further indicated that the *para* and *meta* positions of the phenyl ring are solvent exposed while the *ortho* position is buried in the O pocket of the protein (Fig. 3C). This is in agreement with the higher catalytic activities observed for substrates substituted at the *para* and *meta* positions relative to *ortho*-substituted substrates (Table 1, entries 4 to 7), for which steric hindrance by S298* (monomer B) and Y76 may prevent productive binding (Fig. 3C). The importance of steric crowding at the *ortho* position is confirmed by the relatively high activity with the *ortho*-fluorine-substituted substrate compared to the compounds with larger substituents, i.e., chlorine and bromine (Table 1, entries 5 to 7). Thus, it appears that the substrate specificity of *VpAT* with regard to substituted β-phenylalanine derivatives is mainly determined by the residues that line the hydrophobic O pocket.

A sequence motif for aromatic β-amino acid aminotransferases. A combination of sequence analysis, structural data, and information on substrate range may be used to define aminotransferase sequence motifs that can be applied to identify enzymes with certain substrate specificities (81). As observed in the *VpAT* structure bound with AOA (Fig. 3A and B), as well as in the modeled *VpAT*–(S)-β-phenylalanine structure (Fig. 3C), residues R41 in loop 1 and E75 in loop 2 of the NC domain contribute to shaping the carboxylate-binding site for β-amino acids in the P pocket. For this reason, and also taking into account that E75 is conserved in *MesAT* and *PoGSAM* but not in *SeGSAM* (Fig. 1), we used the conserved functional P-pocket R41/E75 pair as the starting point to identify a motif that may distinguish aminotransferases with activity toward aromatic β-amino acids from aminotransferases with a similar fold but different activity. Extension of the R41/E75 motif with residues selected on the basis of sequence conservation or position in the structure yielded a motif, R-X-[AVI]-X(6)-P-X(14)-D-G-X(8)-[EDNQ]-[YFW], that comprises

36 residues in the N-terminal segment of the NC domain. When used to screen a set of over 2,200 (putative) class I aminotransferases, it allowed the selection of 46 sequences (see Table S3 in the supplemental material), among which were the sequences of the conserved β -selective enzymes *VpAT*, *MesAT*, and *PoGSAM*. Although several retrieved sequences have been annotated as glutamate-1-semialdehyde 2,1-aminomutase (see Table S3 in the supplemental material), we propose that the sequences selected by the motif are ATs with activity toward aromatic β -amino acids. Although our motif may be applied to identify aminotransferases in a set of fold type I aminotransferases, it probably cannot predict β -aminotransferase activity in less related sequences or enzymes lacking activity with aromatic substrates. For example, *AdbpAT* (21), which is inactive with aromatic β -amino acids and has an enantioselectivity for aliphatic β -amino acids opposite that of *VpAT* (Table 1) and *MesAT* (27), is not covered by the motif. This may suggest that *AdbpAT* belongs to a different phylogenetic cluster and may have a different mode of substrate binding.

Outlook. Although more experiments or structure and modeling studies are needed, we are confident that the newly discovered β -phenylalanine aminotransferase from *V. paradoxus* strain CBF3, along with its promising biocatalytic properties and its 3D structure, is an attractive template for future β -aminotransferase-engineering efforts toward the synthesis of enantiomerically pure β -amino acids.

ACKNOWLEDGMENTS

We thank O. May and B. Kaptein from DSM for helpful discussions. We acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of synchrotron radiation facilities and thank the beamline staff of ID14-4 for their assistance.

This research was financially supported by the Netherlands Ministry of Economic Affairs and the B-Basic partner organizations (<http://www.b-basic.nl>) through B-Basic, a public-private NWO-ACTS program. The work of S. Bartsch was supported by the EU FP7 under Grant Agreement number 266025 (BIONEXGEN). Hein J. Wijma was supported by NWO (Netherlands Organization for Scientific Research) through an ECHO grant.

REFERENCES

- Seebach D, Beck A, Bierbaum D. 2004. The world of beta- and gamma-peptides comprised of homologated proteinogenic amino acids and other components. *Chem. Biodivers.* 1:1111–1239.
- Burley SK, David PR, Lipscomb WN. 1991. Leucine aminopeptidase: bestatin inhibition and a model for enzyme-catalyzed peptide hydrolysis. *Proc. Natl. Acad. Sci. U. S. A.* 88:6916–6920.
- Geueke B, Kohler H. 2007. Bacterial beta-peptidyl aminopeptidases: on the hydrolytic degradation of beta-peptides. *Appl. Microbiol. Biotechnol.* 74:1197–1204.
- Juaristi E, Soloshonok V. 2005. Enantioselective synthesis of β -amino acids. Wiley-VCH, New York, NY.
- Shih C, Gossett LS, Gruber JM, Grossman CS, Andis SL, Schultz RM, Worzalla JF, Corbett TH, Metz JT. 1999. Synthesis and biological evaluation of novel cryptophycin analogs with modification in the beta-alanine region. *Bioorg. Med. Chem. Lett.* 9:69–74.
- Mekhail TM, Markman M. 2002. Paclitaxel in cancer therapy. *Expert Opin. Pharmacother.* 3:755–766.
- Ojima I, Lin S, Wang T. 1999. Recent advances in the medicinal chemistry of taxoids with novel beta-amino acid side chains. *Curr. Med. Chem.* 6:927–954.
- Crews P, Manes LV, Boehler M. 1986. Jasplakinolide, a cyclodepsipeptide from the marine sponge, *Jaspis* sp. *Tetrahedron Lett.* 27:2797–2800.
- Weiner B, Szymanski W, Janssen DB, Minnaard AJ, Feringa BL. 2010. Recent advances in the catalytic asymmetric synthesis of beta-amino acids. *Chem. Soc. Rev.* 39:1656–1691.
- Liljeblad A, Kanerva LT. 2006. Biocatalysis as a profound tool in the preparation of highly enantiopure beta-amino acids. *Tetrahedron* 62: 5831–5854.
- Rehdorf J, Mihovilovic MD, Bornscheuer UT. 2010. Exploiting the regioselectivity of Baeyer-Villiger monooxygenases for the formation of beta-amino acids and beta-amino alcohols. *Angew. Chem. Int. Ed. Engl.* 49:4506–4508.
- Chirpich TP, Zappia V, Costilow RN, Barker HA. 1970. Lysine 2,3-aminomutase. Purification and properties of a pyridoxal phosphate and S-adenosylmethionine-activated enzyme. *J. Biol. Chem.* 245:1778–1789.
- Ruzicka FJ, Frey PA. 2007. Glutamate 2,3-aminomutase: a new member of the radical SAM superfamily of enzymes. *Biochim. Biophys. Acta* 1774: 286–296.
- Wu B, Szymanski W, Heberling MM, Feringa BL, Janssen DB. 2011. Aminomutases: mechanistic diversity, biotechnological applications and future perspectives. *Trends Biotechnol.* 29:352–362.
- Cronan JE, Jr. 1980. Beta-alanine synthesis in *Escherichia coli*. *J. Bacteriol.* 141:1291–1297.
- Christianson CV, Montavon TJ, Festin GM, Cooke HA, Shen B, Bruner SD. 2007. The mechanism of MIO-based aminomutases in beta-amino acid biosynthesis. *J. Am. Chem. Soc.* 129:15744–15745.
- Christen P, Metzler DE. 1985. Transaminases. Wiley, New York, NY.
- Stirling DI. 1992. The use of aminotransferases for the production of chiral amino acids and amines. In Collins AN, Sheldrake GN, Crosby J (ed), *Chirality in industry*. Wiley, New York, NY.
- Kim J, Kyung D, Yun H, Cho BK, Kim BG. 2006. Screening and purification of a novel transaminase catalyzing the transamination of aryl beta-amino acid from *Mesorhizobium* sp LUK. *J. Microbiol. Biotechnol.* 16:1832–1836.
- Shin JS, Kim BG. 2001. Comparison of the omega-transaminases from different microorganisms and application to production of chiral amines. *Biosci. Biotechnol. Biochem.* 65:1782–1788.
- Yun H, Lim S, Cho BK, Kim BG. 2004. Omega-amino acid:pyruvate transaminase from *Alcaligenes denitrificans* Y2k-2: a new catalyst for kinetic resolution of beta-amino acids and amines. *Appl. Environ. Microbiol.* 70:2529–2534.
- Percudani R, Peracchi A. 2009. The B6 database: a tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families. *BMC Bioinformatics* 10:273.
- Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. 2010. The Pfam protein families database. *Nucleic Acids Res.* 38:D211–D222.
- Yonaha K, Suzuki K, Minei H, Toyama S. 1983. Distribution of omega-amino acid-pyruvate transaminase and aminobutyrate-alpha-ketoglutarate transaminase in microorganisms. *Agric. Biol. Chem. Tokyo* 47:2257–2265.
- Christen P, Mehta P. 2001. From cofactor to enzymes. The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Chem. Rec.* 1:436–447.
- Mehta PK, Hale TI, Christen P. 1993. Aminotransferases—demonstration of homology and division into evolutionary subgroups. *Eur. J. Biochem.* 214: 549–561.
- Kim J, Kyung D, Yun H, Cho BK, Seo JH, Cha M, Kim BG. 2007. Cloning and characterization of a novel beta-transaminase from *Mesorhizobium* sp. strain LUK: a new biocatalyst for the synthesis of enantiomerically pure beta-amino acids. *Appl. Environ. Microbiol.* 73:1772–1782.
- Hwang BY, Cho BK, Yun H, Koteshwar K, Kim BG. 2005. Revisit of aminotransferase in the genomic era and its application to biocatalysis. *J. Mol. Catal. B Enzym.* 37:47–55.
- Koszelewski D, Lavandera I, Clay D, Guebitz GM, Rozzell D, Kroutil W. 2008. Formal asymmetric biocatalytic reductive amination. *Angew. Chem. Int. Ed. Engl.* 47:9337–9340.
- Koszelewski D, Tauber K, Faber K, Kroutil W. 2010. Omega-transaminases for the synthesis of non-racemic alpha-chiral primary amines. *Trends Biotechnol.* 28:324–332.
- Rozzell JD. May 1985. Production of amino acids by transamination. US patent 4518692.
- Stewart JD. 2001. Dehydrogenases and transaminases in asymmetric synthesis. *Curr. Opin. Chem. Biol.* 5:120–129.
- 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.

34. Wu B, Szymanski W, Crismaru CG, Feringa BL, Janssen DB. 2012. C-N lyases catalyzing addition of ammonia, amines and amides to C=C and C=O bonds, p 749–778. In Drauz K, Groger H, May O (ed), *Enzyme catalysis in organic synthesis*. Wiley-VCH, New York, NY.
35. Nowicki C, Hunter GR, Montemartini-Kalisz M, Blankenfeldt W, Hecht H, Kalisz HM. 2001. Recombinant tyrosine aminotransferase from *Trypanosoma cruzi*: structural characterization and site directed mutagenesis of a broad substrate specificity enzyme. *Biochim. Biophys. Acta* 1546: 268–281.
36. Rudat J, Brucher BR, Syltatk C. 2012. Transaminases for the synthesis of enantiopure beta-amino acids. *AMB Express*. 2:11.
37. Szymanski W, Wu B, Weiner B, de Wildeman S, Feringa BL, Janssen DB. 2009. Phenylalanine aminomutase-catalyzed addition of ammonia to substituted cinnamic acids: a route to enantiopure alpha- and beta-amino acids. *J. Org. Chem.* 74:9152–9157.
38. Janssen DB, Scheper A, Witholt B. 1984. Biodegradation of 2-chloroethanol and 1,2-dichloroethane by pure bacterial cultures. *Prog. Ind. Microbiol.* 20:169–178.
39. Gabor E, de Vries E, Janssen D. 2004. Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases. *Environ. Microbiol.* 6:948–958.
40. Hill DW, Walters FH, Wilson TD, Stuart JD. 1979. High performance liquid chromatographic determination of amino acids in the picomole range. *Anal. Chem.* 51:1338–1341.
41. Woese CR. 1987. Bacterial evolution. *Microbiol. Rev.* 51:221–271.
42. Poelarends G, Wilkens M, Larkin M, van Elsas J, Janssen D. 1998. Degradation of 1,3-dichloropropene by *Pseudomonas cichorii* 170. *Appl. Environ. Microbiol.* 64:2931–2936.
43. Sambrook J, Russell DW. 2001. *Molecular cloning: a laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
44. Staskawicz B, Dahlbeck D, Keen N, Napoli C. 1987. Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinia*. *J. Bacteriol.* 169:5789–5794.
45. Brady OL, Elsmie GV. 1926. The use of 2,4-dinitrophenylhydrazine as a reagent for aldehydes and ketones. *Analyst* 51:77–78.
46. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539.
47. Drummond AJ AB, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A. 2010. Geneious v5.5, 5.5.3 ed. Biomatters, Auckland, New Zealand.
48. Wybenga GG, Crismaru CG, Janssen DB, Dijkstra BW. 2012. Structural determinants of the beta-selectivity of a bacterial aminotransferase. *J. Biol. Chem.* 287:28495–28502.
49. Evans P. 2006. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* 62:72–82.
50. Kabsch W. 2010. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* 66:133–144.
51. Bailey S. 1994. The CCP4 Suite—programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50:760–763.
52. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* 40: 658–674.
53. Langer G, Cohen SX, Lamzin VS, Perrakis A. 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* 3:1171–1179.
54. Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60:2126–2132.
55. Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53:240–255.
56. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66:12–21.
57. Kleywegt GJ, Jones TA. 1998. Databases in protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 54:1119–1131.
58. Schuttelkopf AW, van Aalten DM. 2004. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* 60:1355–1363.
59. Holm L, Rosenstrom P. 2010. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38:W545–W549.
60. Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372:774–797.
61. Davis IW, Baker D. 2009. RosettaLigand docking with full ligand and receptor flexibility. *J. Mol. Biol.* 385:381–392.
62. Davis IW, Raha K, Head MS, Baker D. 2009. Blind docking of pharmaceutically relevant compounds using RosettaLigand. *Protein Sci.* 18: 1998–2002.
63. Richter F, Leaver-Fay A, Khare SD, Bjelic S, Baker D. 2011. De novo enzyme design using Rosetta3. *PLoS One* 6:e19230. doi:10.1371/journal.pone.0019230.
64. Krieger E, Koraimann G, Vriend G. 2002. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins* 47:393–402.
65. OpenEye Scientific Software, Inc. 2010. OEChem TK programming library for chemistry and cheminformatics. OpenEye Scientific Software, Inc., Santa Fe, NM.
66. Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7:203–214.
67. Han JI, Choi HK, Lee SW, Orwin PM, Kim J, Laroe SL, Kim TG, O'Neil J, Leadbetter JR, Lee SY, Hur CG, Spain JC, Ovchinnikova G, Goodwin L, Han C. 2011. Complete genome sequence of the metabolically versatile plant growth-promoting endophyte *Variovorax paradoxus* S110. *J. Bacteriol.* 193:1183–1190.
68. Bea HS, Park HJ, Lee SH, Yun H. 2011. Kinetic resolution of aromatic beta-amino acids by omega-transaminase. *Chem. Commun. (Camb)* 47: 5894–5896.
69. Tipton KF. 1996. Patterns of enzyme inhibition, p 115–174. In Engel PC (ed), *Enzymology*. Academic Press, San Diego, CA.
70. Grishin NV, Phillips MA, Goldsmith EJ. 1995. Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci.* 4:1291–1304.
71. Schneider G, Kack H, Lindqvist Y. 2000. The manifold of vitamin B6 dependent enzymes. *Structure* 8:R1–R6.
72. John RA, Charteris A. 1978. The reaction of amino-oxyacetate with pyridoxal phosphate-dependent enzymes. *Biochem. J.* 171:771–779.
73. Markovic-Housley Z, Schirmer T, Hohenester E, Khomutov AR, Khomutov RM, Karpeisky MY, Sandmeier E, Christen P, Jansonius JN. 1996. Crystal structures and solution studies of oxime adducts of mitochondrial aspartate aminotransferase. *Eur. J. Biochem.* 236:1025–1032.
74. Stetefeld J, Jenny M, Burkhard P. 2006. Intersubunit signaling in glutamate-1-semialdehyde-aminomutase. *Proc. Natl. Acad. Sci. U. S. A.* 103: 13688–13693.
75. Lowe PN, Rowe AF. 1985. Aspartate-2-oxoglutarate aminotransferase from *Trichomonas vaginalis*: identity of aspartate-aminotransferase and aromatic amino-acid aminotransferase. *Biochem. J.* 232:689–695.
76. Shin JS, Yun H, Jang JW, Park I, Kim BG. 2003. Purification, characterization, and molecular cloning of a novel amine:pyruvate transaminase from *Vibrio fluvialis* JS17. *Appl. Microbiol. Biotechnol.* 61:463–471.
77. Sung MH, Tanizawa K, Tanaka H, Kuramitsu S, Kagamiyama H, Soda K. 1990. Purification and characterization of thermostable aspartate-aminotransferase from a thermophilic *Bacillus* species. *J. Bacteriol.* 172: 1345–1351.
78. Shin JS, Kim BG. 2002. Exploring the active site of amine:pyruvate aminotransferase on the basis of the substrate structure-reactivity relationship: how the enzyme controls substrate specificity and stereo selectivity. *J. Org. Chem.* 67:2848–2853.
79. Okamoto A, Nakai Y, Hayashi H, Hirotsu K, Kagamiyama H. 1998. Crystal structures of *Paracoccus denitrificans* aromatic amino acid aminotransferase: a substrate recognition site constructed by rearrangement of hydrogen bond network. *J. Mol. Biol.* 280:443–461.
80. Eliot AC, Kirsch JF. 2004. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* 73: 383–415.
81. Hohne M, Schatzle S, Jochens H, Robins K, Bornscheuer UT. 2010. Rational assignment of key motifs for function guides in silico enzyme identification. *Nat. Chem. Biol.* 6:807–813.