Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Crystallographic characterization of the (*R*)-selective amine transaminase from *Aspergillus fumigatus*

Maren Thomsen,‡ Lilly Skalden,‡ Gottfried J. Palm, Matthias Höhne, Uwe T. Bornscheuer and Winfried Hinrichs*

Institute of Biochemistry, University of Greifswald, Felix-Hausdorff-Strasse 4, 17489 Greifswald, Germany

‡ These authors contributed equally to this work.

Correspondence e-mail: winfried.hinrichs@uni-greifswald.de

Received 19 December 2013
Accepted 15 January 2014

PDB reference: (*R*)-selective amine transaminase, 4chi

The importance of amine transaminases for producing optically pure chiral precursors for pharmaceuticals and chemicals has substantially increased in recent years. The X-ray crystal structure of the (*R*)-selective amine transaminase from the fungus *Aspergillus fumigatus* was solved by S-SAD phasing to 1.84 Å resolution. The refined structure at 1.27 Å resolution provides detailed knowledge about the molecular basis of substrate recognition and conversion to facilitate protein-engineering approaches. The protein forms a homodimer and belongs to fold class IV of the pyridoxal-5′-phosphate-dependent enzymes. Both subunits contribute residues to form two active sites. The structure of the holoenzyme shows the catalytically important cofactor pyridoxal-5′-phosphate bound as an internal aldimine with the catalytically responsible amino-acid residue Lys179, as well as in its free form. A long N-terminal helix is an important feature for the stability of this fungal (*R*)-selective amine transaminase, but is missing in branched-chain amino-acid aminotransferases and D-amino-acid aminotransferases.

1. Introduction

During the last decade, interest in transaminases has increased strongly (Koszelewski *et al.*, 2010; Kroutil *et al.*, 2013; Malik *et al.*, 2012; Mathew & Yun, 2012; Rudat *et al.*, 2012; Tufvesson *et al.*, 2011). Many new transaminases have been discovered and applied in organic syntheses to obtain optically pure amines and non-natural amino acids for chemical and pharmaceutical applications (Höhne & Bornscheuer, 2012). This includes oxazolone derivatives used for the treatment of diabetes (Sutin *et al.*, 2007), rivastigmine serving in the treatment of Alzheimer's disease (Fuchs *et al.*, 2010; Rösler *et al.*, 1999), a protected kedarcidine aglycon useful as an antitumour antibiotic (Ogawa *et al.*, 2009), mexiletine for the treatment of cardiac arrhythmia (Koszelewski, Clay *et al.*, 2009; Koszelewski, Pressnitz *et al.*, 2009) and imagabalin, which has been suggested for the treatment of generalized anxiety disorder (Midelfort *et al.*, 2013).

Transaminases belong to the pyridoxal-5′-phosphate (PLP)-dependent enzymes. Besides transamination, the cofactor PLP facilitates a broad variety of other enzymatic reactions such as racemization, decarboxylation and elimination, where it serves as an electron sink to stabilize carbanion intermediates (Christen & Mehta, 2001). The reaction catalyzed by transaminases is the reversible conversion of α-keto acids, ketones and aldehydes to the corresponding amino acids or amines (Hayashi, 1995). The catalysis itself is divided into two half-reactions. During the first half-reaction the amino group of a

suitable amino donor is transferred to PLP to yield pyridoxamine-5′-phosphate (PMP) with the simultaneous release of the co-product, the deaminated donor. In the second half-reaction the amino acceptor is converted to the corresponding amine and PLP is thus regenerated (Eliot & Kirsch, 2004; Jansonius, 1998). Transaminases can be used in the kinetic resolution of racemic amines and amino acids with a maximum yield of one enantiomer or in asymmetric synthesis starting from prochiral ketones to yield the corresponding optically pure amine at up to 100% yield, if a suitable method to shift the equilibrium to amine formation is employed. In particular, the latter method makes them very useful in the production of building blocks for pharmaceuticals (Martens & Schickedanz, 1986; Blaser, 2002).

Transaminases can be divided into α-transaminases, ω-transaminases and amine transaminases based on their substrate scope. Whereas the substrates of α-transaminases require a carboxylate group in the α-position to the carbonyl function, ω-transaminases also accept substrates with several C atoms (Schrewe *et al.*, 2013) between the carbonyl and the carboxylic acid function and, typically, the ketone or aldehyde function is at the (sub-)terminal C atom of the substrate. Amine transaminases convert ketones to amines and do not require a carboxylate group in the substrate (Höhne & Bornscheuer, 2012).

Seven fold classes of PLP-dependent enzymes are currently known, and transaminases have been identified in classes I and IV (Eliot & Kirsch, 2004; Jansonius, 1998). All of the members of these fold classes share the characteristic that the smallest catalytic unit is a homodimer (Eliot & Kirsch, 2004). The monomer can be divided into a large and a small domain. The two active sites lie at the interface between the domains, and amino-acid residues of each monomer contribute to the catalytic centre. The active sites of fold classes I and IV can be regarded as mirror images. Whereas (*S*)-selective amine transaminases occur in fold class I, (*R*)-selective amine transaminases belong to fold class IV (Jansonius, 1998; Eliot & Kirsch, 2004). This assignment also matches observations during protonation in the catalytic mechanism. In (*R*)-selective amine transaminases the *si*-site (Hanson, 1966) of the generated quinoid intermediate is solvent-facing, whereas in the (*S*)-selective amine transaminases it is the *re*-site.

To enable the production of enantiopure compounds, amine transaminases with both enantiopreferences are required. In 2010, Höhne and coworkers discovered 17 (*R*)-selective amine transaminases using an *in silico* search (Höhne *et al.*, 2010). To find these putative (*R*)-selective amine transaminases sequences, the *in silico* search was based on the determination of specific sequence motifs which characterize either D-amino-acid aminotransferases (D-ATAs) or branched-chain amino-acid aminotransferases (BCATs) to filter out motifs for (*R*)-selective amine transaminases. Based on these criteria, the sequences of BCATs and D-ATAs could be excluded and the remaining sequences (approximately 0.4% of all investigated sequences) were experimentally confirmed to be (*R*)-selective amine transaminases (Höhne *et al.*, 2010). Structures of α-transaminases (Schwarzenbacher *et al.*, 2004; Han *et al.*, 2006) and also of a few (*S*)-selective amine transaminases have been published and investigated (Humble *et al.*, 2012; Sayer *et al.*, 2013; Steffen-Munsberg *et al.*, 2013), but a structural analysis of an (*R*)-selective amine transaminase has not been published to date. Presently, a homology model of an amine transaminase from *Arthrobacter* sp. is the only existing toehold (Savile *et al.*, 2010).

In this paper, we present the crystal structure analysis of the (*R*)-selective amine transaminase from the fungus *Aspergillus fumigatus*.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the outermost resolution shell.

Data set	Native	Anomalous
Beamline	BL14.1, BESSY II	BL14.1, BESSY II
Detector	Pilatus 6M	Pilatus 6M
Wavelength (Å)	0.91841	1.77122
Temperature (K)	100	100
Space group	$C222_1$	$C222_1$
Unit-cell parameters (Å)	a = 102.2, b = 120.9, c = 135.4	a = 102.2, b = 120.9, c = 135.4
Resolution range (Å)	50.0–1.27 (1.35–1.27)	50.0–1.84 (1.95–1.84)
No. of unique reflections	426722 (68273)	135117 (17260)
Multiplicity	3.38 (3.3)	5.6 (2.6)
$R_{\rm merge}$† (%)	6.3 (60.3)	3.9 (9.5)
Mean $I/\sigma(I)$	13.2 (2.0)	30.29 (8.34)
$CC_{1/2}$‡ (%)	99.9 (73.0)	99.9 (98.8)
Completeness (%)	99.1 (97.9)	95.9 (75.8)
Overall B factor from Wilson plot (Å^2)	17.4	18.8
Total rotation/increment (°)	180/0.1	360/0.1

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple measurements. ‡ $CC_{1/2}$ is the percentage correlation between intensities from random half data sets (Karplus & Diederichs, 2012).

2. Materials and methods

2.1. Expression and purification

The expression, purification and crystallization of the (*R*)-selective amine transaminase from *A. fumigatus* were performed as reported previously (Thomsen *et al.*, 2013).

2.2. Crystallization and diffraction data collection

For cryoprotection, a solution consisting of 35% glycerol, 20 m*M* tricine pH 7.5, 10 µ*M* PLP was used. X-ray diffraction data were collected at 100 K on beamline 14.1 at the BESSY II synchrotron, Berlin, Germany. Two data sets were collected from one crystal. The first data set at a wavelength of 0.918 Å was obtained using the highest intensity of the storage ring and the second was collected at a wavelength of 1.771 Å to obtain the highest anomalous signal of the S atoms present in the protein. The resolution of the anomalous data set was limited by the detector size. All diffraction images were processed with *XDS* (Kabsch, 2010) using the graphical user interface *XDSapp* (Krug *et al.*, 2012). Data-collection and processing statistics are summarized in Table 1.

2.3. Structure determination of the holoenzyme by S-SAD phasing

The crystal structure of the (*R*)-selective amine transaminase from *A. fumigatus* was determined by single-wavelength anomalous dispersion sulfur (S-SAD) phasing using the 'native crystals SAS' protocol of the automated crystal structure determination platform *Auto-Rickshaw* (Panjikar *et al.*, 2005). SAD was preferred over MAD because higher CC (all/weak) parameters were obtained in *SHELXD*. The automated SAS protocol incorporates *SHELXC* (Sheldrick *et al.*, 2001) for data preparation as well as *SHELXD* (Schneider & Sheldrick, 2002) to find heavy-atom positions. With two cysteine and ten methionine residues per monomer, we searched for 24 S-atom positions per asymmetric unit. The resolution limit for substructure determination and initial phasing was set to 2.5 Å. The best solution obtained resulted in CC (all/weak) of 27.11/16.93 and a PATFOM of 3.68. The program *ABS* (Hao, 2004) determined the correct hand of the substructure, which was subsequently used by *SHELXE* (Sheldrick, 2002) for initial phasing. SAD phasing statistics are listed in Table 2. Density modification was performed with *DM* (Cowtan, 1994). Automatic tracing using *ARP/wARP* (Perrakis *et al.*, 1999) yielded 97% of the polypeptide model at 1.84 Å resolution. Manual completion of the model was carried out with *Coot* (Emsley & Cowtan, 2004). Final refinement with anisotropic *B* factors was carried out with data extending to 1.27 Å resolution using *REFMAC*5 (Murshudov *et al.*, 2011). The quality of the refined protein model was validated using *MolProbity* (Chen *et al.*, 2010). Refinement statistics are listed in Table 3. All molecular graphics were prepared using *PyMOL* (Delano, 2002).

2.4. Docking studies

The docking studies were performed with *YASARA* (Krieger *et al.*, 2002) with default parameters using the polypeptide chains of the homodimer of our crystallographic model. The enantiomeric PLP adducts of (*R*)- and (*S*)-α-methylbenzylamine [(*R*)-α- and (*S*)-α-MBA] were generated in *YASARA* and energy minimization was performed to the lowest energy conformation. The completely flexible ligands were then alternatively docked into the active site. The chosen simulation cell was defined to be 18 × 17 × 18 Å around the catalytic residue Lys179. All residues of the active site and the active-site loop were included. H atoms were added in riding positions. The correct solution of the docking analysis was distinguished by the orientation of the cofactor PLP. The comparison of the docked enantiomeric PLP adducts with the PLP in the solved crystal structure led to the final assignment of the correct enantiomer.

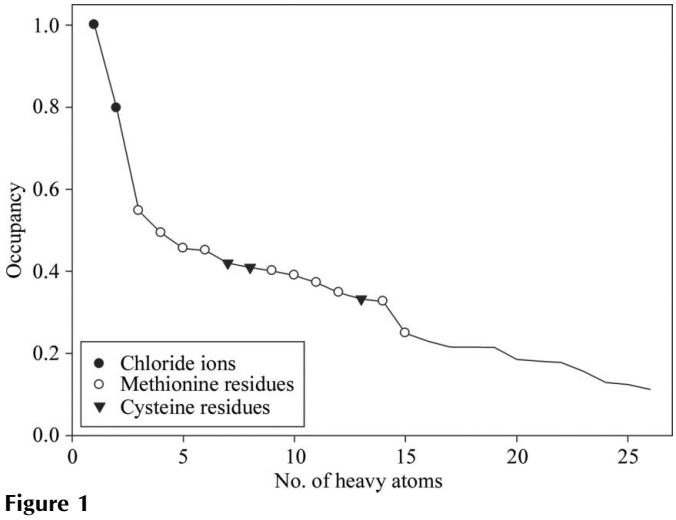

Figure 1
Occupancies of the heavy-atom sites found by *SHELXD* (Schneider & Sheldrick, 2002). Meaningful heavy atoms are labelled by locating their positions in the refined model.

Table 2
Summary of SAD phasing.

SHELXD	
CC (all)	27.11
CC (weak)	16.93
PATFOM	3.68
SHELXE	
CC between E_{obs} and E_{calc}	20.47
CC for partial structure against native data	47.86
FOM	0.725
MapCC	0.901
No. of residues built by *ARP/wARP*	626

Table 3
Refinement statistics.

Resolution (Å)	50.0–1.27
Working/test reflections	208810/10991
R/R_{free}† (%)	10.3/12.7
No. of protein residues	639
No. of water/glycerol molecules	994/2
No. of ions ($Cl^-/K^+/Na^+$)	4/4/2
R.m.s.d. from ideality	
Bond lengths (Å)	0.014
Bond angles (°)	1.842
Average *B* factors (Å^2)	
Protein (5910 atoms)	14.7
Water (1031 atoms)	32.1
Others (96 atoms)	15.8
Ramachandran statistics‡ (%)	
Most favored regions	97.64
Outliers	0
PDB code	4chi

† $R = \sum_{hkl} \big||F_{obs}| - |F_{calc}|\big| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} is analogous to the *R* factor for 5% of the diffraction data excluded from refinement. ‡ Categories were defined by *MolProbity*.

3. Results

3.1. Structure analysis

The phasing contributions of the chloride ions and the S atoms of Cys and Met are shown in Fig. 1. Interestingly, the highest occupancy is observed for two chloride ions, but not for the other possible elements (S, P or K), even taking into account that these atoms show clear signals in the final anomalous electron-density map. The initial phasing based on the anomalous diffraction at 1.84 Å resolution was sufficient for automatic tracing. Refinement using the high-resolution

data converged to an R and R_{free} of 10.3 and 12.7%, respectively.

The final model contains 639 amino-acid residues of two polypeptide chains (*A* and *B*), two PLP molecules, four potassium ions, four chloride ions, two glycerol molecules and 994 water molecules. Ions are assigned according to electron density and meaningful chemical terms and refinement conditions. Both polypeptide chains are well defined by the electron-density maps (Supplementary Fig. S2[1]) and the final model is consistent with the anomalous map (Supplementary Fig. S3). Some residues with poor electron density at the N- and C-termini (monomer *A*, Met1 and Ser322–His332; monomer *B*, Met1–Ala2 and Ser322–His332) were excluded from the structural model. The cofactor PLP was modelled into the active site using $F_{obs} - F_{calc}$ difference maps and refined with a summed occupancy of 0.8. The occupancies of the two PLP states were assigned so that the *B* factors were consistent with those of neighbouring residues. The remaining occupancy of 0.2 was filled with a single phosphate in the same position as the phosphate group of the cofactor.

Additional positive difference electron density was observed in the substrate-binding site within covalent bond distance of the cofactor PLP in each monomer (Supplementary Fig. S4). Owing to the low occupancy of the ligand, we could not conclusively model this density. All compounds used in purification and crystallization and common metabolites were ruled out; also, GC-MS-analysis of acid-denatured and heat-denatured enzyme did not uncover the identity of this ligand. Besides the tricine molecule in the buffer, no amine or carbonyl compounds were added after cell disruption. Nevertheless, D-amino acids were also tested as possible ligands. In this case, the α-carboxyl group could not be modelled into the small binding pocket.

Alternative conformations were modelled for 138 amino-acid side chains out of 639 residues (~20%). Some peptide backbone O atoms could also be modelled in alternative conformations. Differences in the main-chain conformation could be detected for residues Thr204–Gly206. The final refinement and validation statistics are shown in Table 3.

3.2. Overall fold

The (*R*)-selective amine transaminase crystallized in space group $C222_1$ with two monomers in the asymmetric unit forming a homodimer (Fig. 2). Each polypeptide chain is constituted of 332 amino-acid residues with a molecular weight of 37.1 kDa. The tertiary structure of one subunit consists of the typical fold of enzymes belonging to the fold class IV of PLP-dependent enzymes, as first described for D-ATA from *Bacillus* sp. (Sugio *et al.*, 1995). The subunit divides into a small domain (N-terminus to Pro144) with an α/β-structure, an inter-domain loop (Tyr145–Met149) and a large domain (Ala150 to the C-terminus) with a pseudo-barrel structure (Fig. 2). The enzyme belongs to fold class IV of PLP-dependent enzymes, and the overall structure is very similar to those of BCATs and D-ATAs, with the best fit to the BCAT from *Thermus thermophilus* [PDB entry 1wrv; root-mean-square difference on C^{α} atoms (r.m.s.d.) of 1.8 Å, fitting 297 residues; RIKEN Structural Genomics/Proteomics Initiative, unpublished work] and the D-ATA from *Bacillus* sp. YM-1

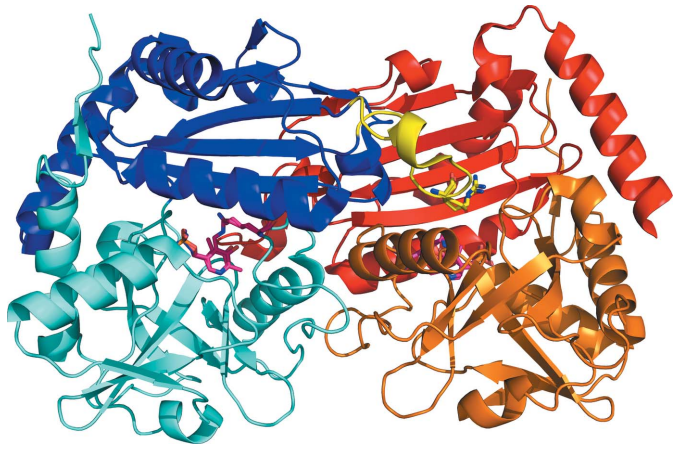

Figure 2
Overall structure of the (*R*)-selective amine transaminase from *A. fumigatus* viewed normal to the molecular dyad. The subunits of the homodimer are shown in blue and red, respectively. The monomer is divided into colour-coded domains: the small domain (blue, red) with the active-site loop (yellow) and the large domain (cyan, orange). The active-site loop derived from the left subunit is shown in yellow (with Arg126 as a stick model). The cofactor PLP is bound to Lys179 at the domain interface (shown as a stick model in pink).

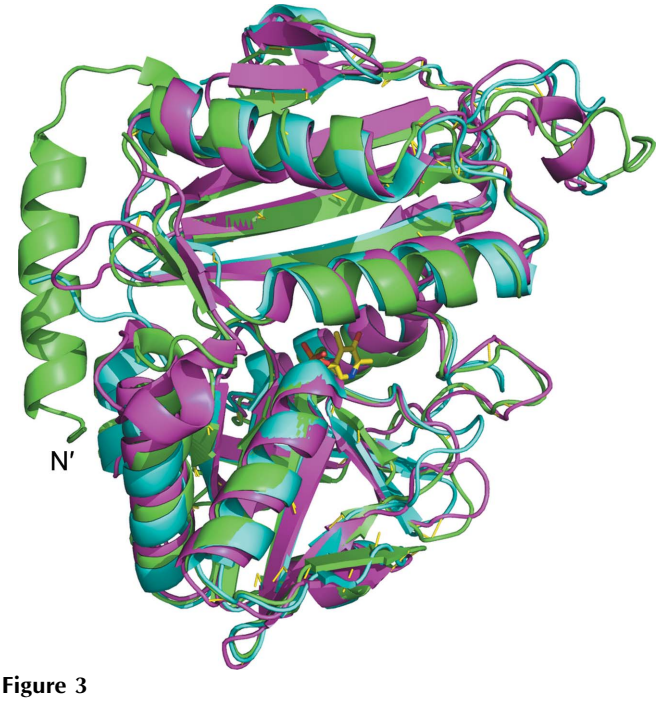

Figure 3
Comparison of the overall monomer fold between the (*R*)-selective amine transaminase from *A. fumigatus* (green), D-amino-acid aminotransferase (PDB entry 3lqs; r.m.s.d. 2.0 Å, cyan) and branched-chain amino-acid aminotransferase (PDB entry 1wrv; r.m.s.d. 1.8 Å, violet) distinctly shows the unique long N-terminal helix found in the (*R*)-selective amine transaminase. The cofactor PLP is shown as a stick model in yellow.

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: DZ5319).

(PDB entry 3lqs; r.m.s.d. of 2.0 Å, fitting 280 residues; Lepore *et al.*, 2010). However, the (*R*)-selective amine transaminase from *A. fumigatus* has an additional long N-terminal α-helix (Met4–Arg20) which has a significant effect on protein stability, as discussed below (Fig. 3). Other differences in the backbone folding compared with BCATs and D-ATAs are only visible for loop regions on the surface. Residues from each domain as well as residues from the other subunit of the dimer participate in forming the active site. The active-site loop (Gly121*–Asn135*; residues labelled with an asterisk belong to the other subunit) limits access to the active site and is contributed by the other subunit.

3.3. Cofactor binding

Well defined electron density is observed for the cofactor PLP, which is located at the bottom of the active site between the small and the large domains (Fig. 2). There are two distinct states observed for the cofactor. One state (occupancies of 0.5 and 0.4 in monomers *A* and *B*, respectively) is covalently bound to the active-site residue Lys179 of the large domain

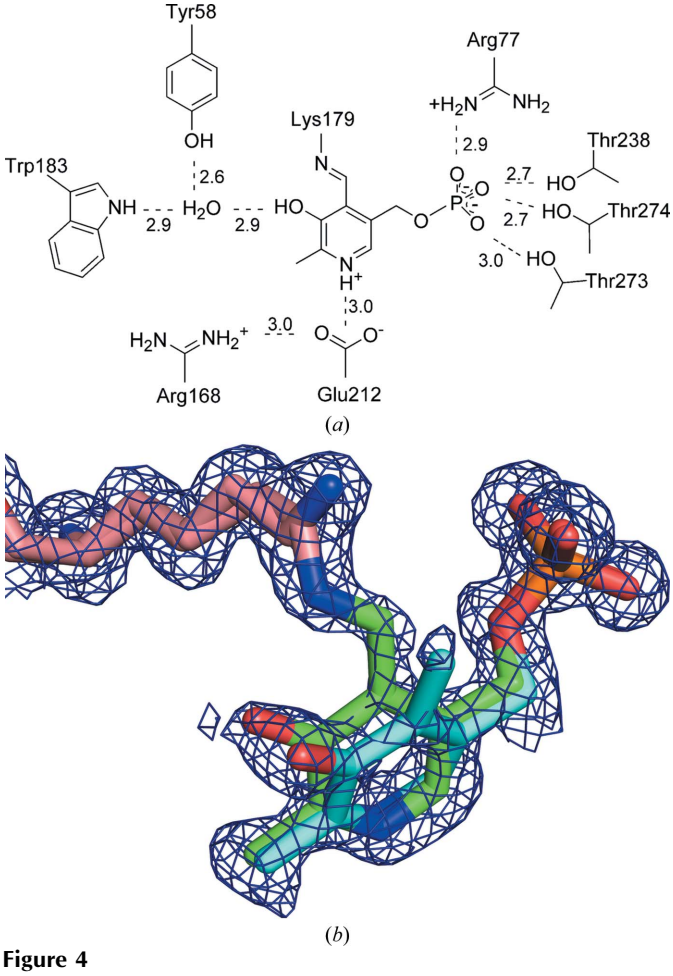

Figure 4
(*a*) Schematic drawing of the coordination of the cofactor PLP (distances are in Å) and (*b*) $2F_{obs} - F_{calc}$ map contoured at 1σ. Note the covalently bound (green C atoms) and unbound (cyan C atoms) states with Lys179. The aldehyde O atom O4A of the free PLP, or alternatively the N atom N4A of PMP, could not be identified in the electron-density map and is therefore omitted in all figures.

(Fig. 4*a*), whereas the other state (occupancies of 0.3 and 0.4 in monomers *A* and *B*, respectively) represents an adduct with an unidentified ligand (see §3.1) and a free lysine (Fig. 4*b*). The covalently bound PLP shows the typical distorted aldimine bond of PLP-dependent enzymes. The bond angles deviate from the ideal 120° and the internal aldimine bond is out of the plane of the pyridoxyl ring by about 11.5 and 14° in monomers *A* and *B*, respectively. This typical geometry has been found in various crystal structures of other PLP-dependent enzymes and it is supposed that the release of strain on breaking the internal aldimine bond enhances the catalytic ability (Dubnovitsky *et al.*, 2005; Hayashi *et al.*, 1998). In the free state Lys179 has a distinct alternative conformation and is involved in a hydrogen-bonding network with Arg77 and the phosphate group of PLP *via* a water molecule. The pyridoxyl ring shows two distinct orientations, and the phosphate group is tightly bound and is involved in several hydrogen-bond interactions (with His74, Arg77, Thr273, Thr274, Ile237 and Thr238) as an anchor for the cofactor. Residues Ile237 and Thr238 are located at the N-terminus of helix α7 (according to Sugio *et al.*, 1995) such that the dipole moment of the helix additionally facilitates the coordination of the phosphate group. The pyridoxyl ring is sandwiched between Leu234 and the peptide bond of Gly215 to Phe216. The ion pair formed by the highly conserved Glu212 and the N atom of the pyridoxyl ring (N1) provides an electron sink during the reaction mechanism. This glutamate is in turn coordinated by the conserved Arg168. Anchoring of PLP by the phosphate and N1 coordination

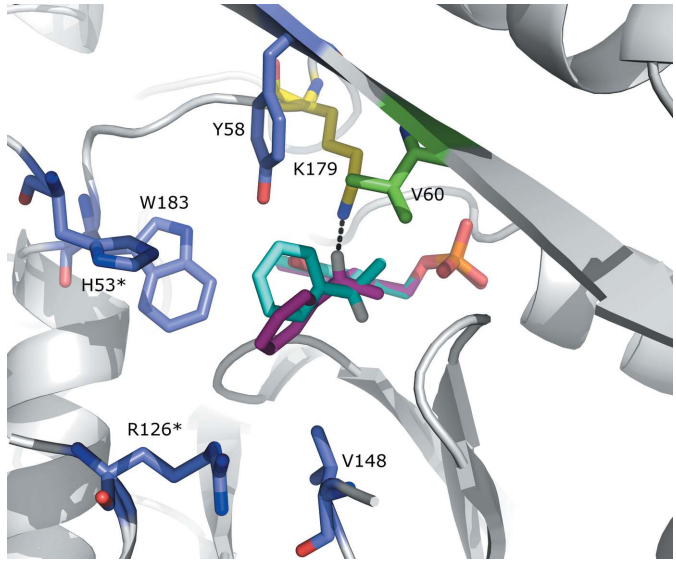

Figure 5
Comparison of the docking studies of (*R*)-α-MBA-PLP (violet) and (*S*)-α-MBA-PLP (cyan) into the active site of the (*R*)-selective amine transaminase from *A. fumigatus*. On the basis of the docking experiments the active site can be divided into a small binding pocket hosting the methyl group and a large binding pocket which is responsible for coordinating the aromatic ring. Only for the (*R*)-enantiomer is the catalytic lysine residue (yellow) at a reasonable distance (2.8 Å, dashed line) from the H atom (white) of α-MBA-PLP to initiate deamination. Residues defining the small binding pocket are coloured green, while residues of the large binding pocket are shown in blue.

limits movement of the pyridoxyl ring to a rotation around the C5—C5A bond by about 19° (Fig. 4*b*).

3.4. Deletion of the N-terminal helix

Previously, we performed crystallization studies on the (*R*)-amine transaminase from *Neosartorya fischeri* (96% identity to the amine transaminase from *A. fumigatus*). Presently, the diffraction images of the obtained crystals are not indexable. Based on a homology model built from D-amino-acid amino-transferase (PDB entry 3daa; Peisach *et al.*, 1998) with a flexible N-terminus, an N-terminal deletion of 22 amino-acid residues was introduced to improve the crystallization quality. Whereas the wild-type amine transaminase could be over-expressed as soluble protein, this N-terminal deletion resulted in insoluble protein. The insolubility of this amine transaminase variant could not be prevented by changing the expression temperature, varying the inducer concentration, altering the induction time or the co-expression of chaperones.

In the solved X-ray structure of (*R*)-amine transaminase from *A. fumigatus*, the N-terminus forms a long helix corresponding to the sequence that was deleted in the (*R*)-amine transaminase mutant from *N. fischeri*. It is obvious that the N-terminal helix is important for the soluble expression of fungal amine transaminases, but remarkable hydrophobic patches on the surface of the modelled truncated enzyme are not observed.

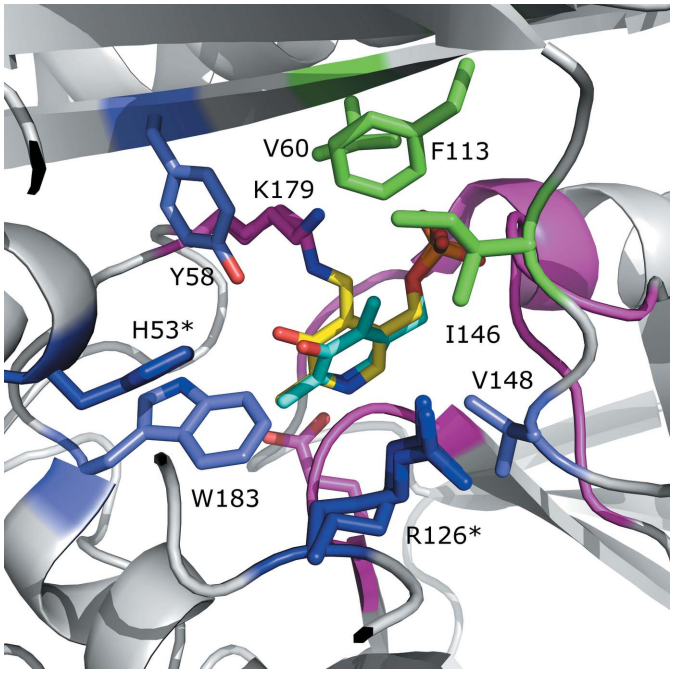

Figure 6
Active-site architecture. The residues forming the small binding pocket are shown in green and the amino-acid residues responsible for forming the large binding pocket are shown in blue (residues which originate from the other subunit are shown in dark blue and are marked with asterisks; PLP-binding residues are coloured violet).

3.5. Structural design of the active site

Via docking studies performed with the program *YASARA*, it could be demonstrated that the active site of the (*R*)-selective amine transaminase from *A. fumigatus* is divided into a small and a large binding pocket (Fig. 5). Docking was performed with the substrate adducts of (*R*)-α-methylbenzylamine [(*R*)-α-MBA] and (*S*)-α-MBA to PLP, which starts the first deamination cycle. The pyridoxyl rings and the phosphate group of the modelled substrate adducts superposed very well with the free PLP state (r.m.s.d. of 0.14 Å) of the X-ray structure, indicating good quality of the docking results. In every docking run the methyl group was bound in the small binding pocket formed by the residues Val60, Phe113 and Ile146 (Fig. 6). The aromatic ring was coordinated in the large binding pocket which is built by the residues His53*, Tyr58, Arg126*, Val148 and Trp183. Although (*S*)-α-MBA-PLP could be docked without clashes, the enantioselectivity can be explained by the orientation of (*R*)-α-MBA-PLP and (*S*)-α-MBA-PLP to the catalytically active lysine residue. This residue initiates the deamination reaction by deprotonation and is only at a reasonable distance (2.8 Å) for abstraction of the proton from (*R*)-α-MBA-PLP which points directly towards the Lys179 N^{ε} atom. In contrast, the proton of (*S*)-α-MBA-PLP points in the opposite direction and cannot be abstracted by Lys179.

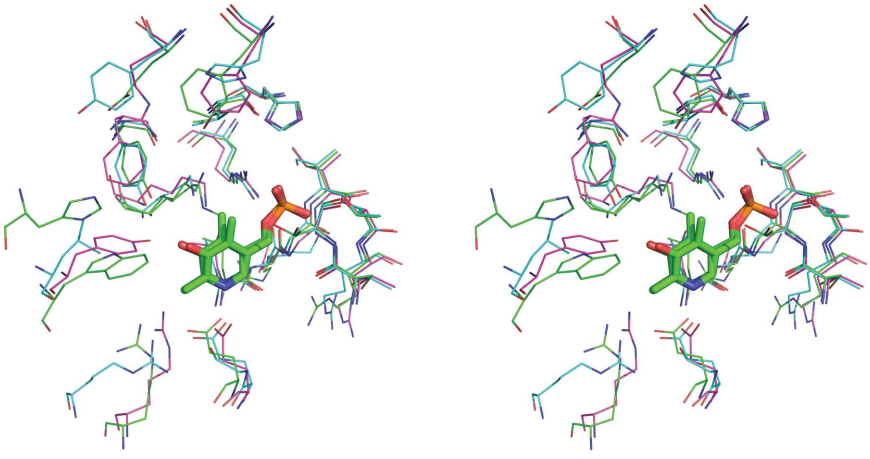

Figure 7
Stereo representation of the active-site comparison between the (*R*)-selective amine transaminase from *A. fumigatus* (green), D-amino acid aminotransferase (PDB entry 3lqs; cyan) and branched-chain amino-acid aminotransferase (PDB entry 1wrv; violet). The inter-domain loop is omitted for clarity.

3.6. Active-site comparison

Whereas the binding of the cofactor and the backbone are conserved, a comparison of the active-site residues responsible for substrate recognition of the (*R*)-selective amine transaminases (*R*-ATAs) with the active sites of BCATs and D-ATAs shows clearly that no amino-acid residues other than the catalytic Lys179 and Glu212 are conserved (Fig. 7). When considering

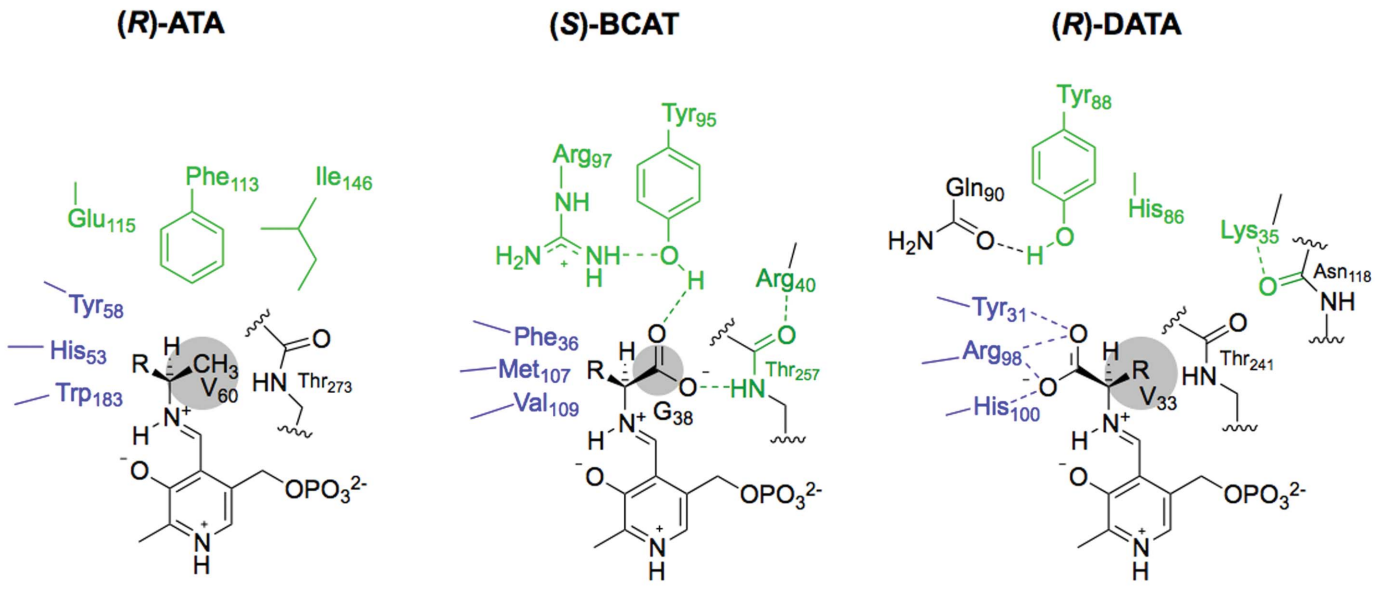

Figure 8
Schematic drawing of the active-site arrangement of enzymes of fold class IV in comparison to the active site of the (*R*)-selective amine transaminase from *A. fumigatus* [(*R*)-ATA], the (*S*)-branched-chain amino-acid aminotransferase [(*S*)-BCAT; PDB entry 1iye] and D-amino-acid aminotransferase [(*R*)-DATA; PDB entry 3daa]. Grey spheres indicate the space-filling requirements of residues Val60, Gly38 and Val33, respectively.

substrates to be converted by (*R*)-selective amine transaminases, the substitution of the carboxyl group by a methyl group inverts the priority according to the Cahn–Ingold–Prelog rule. Hence, the active-site architecture, based on the definition of the small and large pockets, was postulated to be more similar to the BCATs than to the D-ATAs. This can now be verified by the crystal structure with the docking analysis (see §3.5) as well as the sequence motif Y/(F)V*Z* (with *Z* preferably being glutamate) postulated to be an important feature in the structural design of the active site of (*R*)-selective amine transaminases (Höhne *et al.*, 2010). Whereas the large pocket of the D-ATAs is mostly built by small hydrophobic residues, the pocket volume is reduced in BCATs and in the (*R*)-selective amine transaminase by bulky amino-acid residues (Fig. 8). A search of the *DALI* secondary-structure database (Holm & Rosenström, 2010) revealed the inter-domain loop as an active-site limiting feature. The loop in question is two amino-acid residues longer than the equivalent loop found in D-ATAs and therefore contributes to the small binding pocket of the active site. Whereas the small pocket of the D-ATAs harbours predominantly positively charged residues to coordinate the carboxylate group, in BCATs and *R*-ATAs aromatic side chains form a hydrophobic environment for the mostly hydrophobic substituents that are accepted. As mentioned above, the entrance of the active site of the *R*-ATAs is limited by the active-site loop. Interestingly, similar to dual substrate recognition by the (*S*)-amine transaminase (Steffen-Munsberg *et al.*, 2013), the active-site loop of the *R*-ATA also has a highly flexible Arg126 (slightly increased *B* factors and two alternative conformations of residues Arg126, Gly127 and Ser128). By flipping in and out of the active site, Arg126 could facilitate the coordination of the negatively charged carboxylate of the amino acceptor pyruvate as well as the binding of uncharged substrate in the same pocket of the active site. This assumption needs to be investigated further *via* mutagenesis.

4. Conclusion

In the era of rational protein design, crystal structure or NMR analyses at atomic resolution are the most valuable tools for protein-engineering experiments. The crystal structure of the (*R*)-selective amine transaminase elucidated here for the enzyme from *A. fumigatus* provides essential information and insights into understanding how substrate recognition occurs in (*R*)-selective amine transaminases and distinguishes them from other enzymes of fold class IV. For further understanding of substrate binding and enantioselectivity, soaking and co-crystallization experiments are in progress.

Note added in proof. A crystal structure analysis of the (R)-selective ω-transaminase from *Aspergillus terreus* has recently been published (Lyskowski *et al.*, 2014).

MT thanks the 'Landesgraduiertenkolleg Mecklenburg-Vorpommern' for financial support. We thank the European Union (KBBE-2011-5, grant No. 289350) for financial support within the European Union Seventh Framework Programme. Diffraction data were collected on BL14.1 operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron-storage ring (Berlin-Adlershof, Germany). We also thank Professor M. Lalk (University Greifswald, Germany) for the GC-MS analysis.

References

Blaser, H.-U. (2002). *Adv. Synth. Catal.* **344**, 17.

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

Christen, P. & Mehta, P. K. (2001). *Chem. Rec.* **1**, 436–447.

Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

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

Dubnovitsky, A. P., Ravelli, R. B. G., Popov, A. N. & Papageorgiou, A. C. (2005). *Protein Sci.* **14**, 1498–1507.

Eliot, A. C. & Kirsch, J. F. (2004). *Annu. Rev. Biochem.* **73**, 383–415.

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

Fuchs, M., Koszelewski, D., Tauber, K., Kroutil, W. & Faber, K. (2010). *Chem. Commun.* **46**, 5500.

Han, Q., Robinson, H., Gao, Y. G., Vogelaar, N., Wilson, S. R., Rizzi, M. & Li, J. (2006). *J. Biol. Chem.* **281**, 37175–37182.

Hanson, K. R. (1966). *J. Am. Chem. Soc.* **88**, 2731–2742.

Hao, Q. (2004). *J. Appl. Cryst.* **37**, 498–499.

Hayashi, H. (1995). *J. Biochem.* **118**, 463–473.

Hayashi, H., Mizuguchi, H. & Kagamiyama, H. (1998). *Biochemistry*, **37**, 15076–15085.

Höhne, M. & Bornscheuer, U. T. (2012). *Enzymes in Organic Synthesis*, edited by O. May, H. Gröger & W. Drauz, pp. 779–820. Weinheim: Wiley-VCH.

Höhne, M., Schätzle, S., Jochens, H., Robins, K. & Bornscheuer, U. T. (2010). *Nature Chem. Biol.* **6**, 807–813.

Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.

Humble, M. S., Cassimjee, K. E., Håkansson, M., Kimbung, Y. R., Walse, B., Abedi, V., Federsel, H. J., Berglund, P. & Logan, D. T. (2012). *FEBS J.* **279**, 779–792.

Jansonius, J. N. (1998). *Curr. Opin. Struct. Biol.* **8**, 759–769.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.

Koszelewski, D., Clay, D., Rozzell, D. & Kroutil, W. (2009). *Eur. J. Org. Chem.* **2009**, 2289–2292.

Koszelewski, D., Pressnitz, D., Clay, D. & Kroutil, W. (2009). *Org. Lett.* **11**, 4810–4812.

Koszelewski, D., Tauber, K., Faber, K. & Kroutil, W. (2010). *Trends Biotechnol.* **28**, 324–332.

Krieger, E., Koraimann, G. & Vriend, G. (2002). *Proteins*, **47**, 393–402.

Kroutil, W., Fischereder, E. M., Fuchs, C. S., Lechner, H., Mutti, F. G., Pressnitz, D., Rajagopalan, A., Sattler, J. H., Simon, R. C. & Siirola, E. (2013). *Org. Process Res. Dev.* **17**, 751–759.

Krug, M., Weiss, M. S., Heinemann, U. & Mueller, U. (2012). *J. Appl. Cryst.* **45**, 568–572.

Lepore, B. W., Liu, D., Peng, Y., Fu, M., Yasuda, C., Manning, J. M., Silverman, R. B. & Ringe, D. (2010). *Biochemistry*, **49**, 3138–3147.

Lyskowski, A., Gruber, C., Steinkellner, G., Schurmann, M., Schwab, H., Gruber, K. & Steiner, K. (2014). *PLoS One*, **9**.

Malik, M. S., Park, E.-S. & Shin, J.-S. (2012). *Appl. Microbiol. Biotechnol.* **94**, 1163–1171.

Martens, J. G. K. & Schickedanz, M. (1986). *Arch. Pharm.* **319**, 461–465.

Mathew, S. & Yun, H. (2012). *ACS Catal.* **2**, 993–1001.

Midelfort, K. S., Kumar, R., Han, S., Karmilowicz, M. J., McConnell, K., Gehlhaar, D. K., Mistry, A., Chang, J. S., Anderson, M., Villalobos, A., Minshull, J., Govindarajan, S. & Wong, J. W. (2013). *Protein Eng. Des. Sel.* **26**, 25–33.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Ogawa, K., Koyama, Y., Ohashi, I., Sato, I. & Hirama, M. (2009). *Angew. Chem. Int. Ed.* **48**, 1110–1113.

Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.

Peisach, D., Chipman, D. M., Van Ophem, P. W., Manning, J. M. & Ringe, D. (1998). *Biochemistry*, **37**, 4958–4967.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Mol. Biol.* **6**, 5.

Rösler, M., Anand, R., Cicin-Sain, A., Gauthier, S., Agid, Y., Dal-Bianco, P., Stähelin, H. B., Hartman, R. & Gharabawi, M. (1999). *BMJ*, **318**, 633–638.

Rudat, J., Brucher, B. R. & Syldatk, C. (2012). *AMB Express*, **2**, 11.

Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W. & Hughes, G. J. (2010). *Science*, **329**, 305–309.

Sayer, C., Isupov, M. N., Westlake, A. & Littlechild, J. A. (2013). *Acta Cryst.* D**69**, 564–576.

Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.

Schrewe, M., Ladkau, N., Bühler, B. & Schmid, A. (2013). *Adv. Synth. Catal.* **355**, 1693–1697.

Schwarzenbacher, R. *et al.* (2004). *Proteins*, **55**, 759–763.

Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.

Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R. & Uson, I. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 333–351. Dordrecht: Kluwer Academic Publishers.

Steffen-Munsberg, F., Vickers, C., Thontowi, A., Schätzle, S., Tumlirsch, T., Svedendahl Humble, M., Land, H., Berglund, P., Bornscheuer, U. T. & Höhne, M. (2013). *ChemCatChem*, **5**, 150–153.

Sugio, S., Petsko, G. A., Manning, J. M., Soda, K. & Ringe, D. (1995). *Biochemistry*, **34**, 9661–9669.

Sutin, L., Andersson, S., Bergquist, L., Castro, V. M., Danielsson, E., James, S., Henriksson, M., Johansson, L., Kaiser, C., Flyrén, K. & Williams, M. (2007). *Bioorg. Med. Chem. Lett.* **17**, 4837–4840.

Thomsen, M., Skalden, L., Palm, G. J., Höhne, M., Bornscheuer, U. T. & Hinrichs, W. (2013). *Acta Cryst.* F**69**, 1415–1417.

Tufvesson, P., Lima-Ramos, J., Jensen, J. S., Al-Haque, N., Neto, W. & Woodley, J. M. (2011). *Biotechnol. Bioeng.* **108**, 1479–1493.