Stereoelectronic effects in the reaction of aromatic substrates catalysed by *Halomonas elongata* transaminase and its mutants

Martina Letizia Contente[‡a,b], Matteo Planchestainer[‡a], Francesco Molinari[b]*, Francesca Paradisi[a,c]*

A transaminase from *Halomonas elongata* and four mutants generated by an *in silico*-based design, were recombinantly produced in *E. coli*, purified and applied to the amination of mono-substituted aromatic carbonyl-derivatives. While benzaldehyde derivatives resulted excellent substrates, only NO_2-acetophenones were transformed into the (*S*)-amine with high enantioselectivity. The different behaviour of wild-type and mutated transaminases was assessed by *in silico* substrate binding mode studies.

Introduction

The remarkable work of Codexis-Merck in evolving an amino transferase to achieve a "greener" synthesis of Sitaglipin,[1] highlighted the promising features of this class of enzymes as biocatalytic tool for the manufacturing of chiral amines,[2] valuable building blocks for a wide range of bioactive compounds.[3,4] The
mechanism underlying ω-transaminase activity has been generally assumed similar to that of aspartate transaminase;[5,6] only recently the reaction mechanism of *Chromobacterium violaceum* ω-transaminase has been investigated by means of density functional theory calculations and the energy profile for the conversion of (*S*)-1-phenylethylamine to acetophenone has been calculated. In this study the involved transition states and intermediates have been characterised.[7] The first step in the conversion of carbonyls into amines involves the nucleophilic attack from the amino donor intermediate pyridoxamine-5'-phosphate (PMP) to the amino acceptor carbonyl group.[7] The stereoelectronic features of the substrate play a crucial role; in the case of ketones, especially aromatic ones, the carbonyl carbon is often not sufficiently electrophilic and the reaction is slow and thermodynamically unfavoured.[8] This aspect heavily influences the wide applicability of these enzymes as an alternative to conventional synthesis.[9] On the other hand, unfavourable steric interactions between the enzyme and the substrate may hamper the reactivity and can only be gauged after investigation of the enzyme-substrate binding mode.[10] Beside stereoelectronic effects involved in the enzymatic interactions, water solubility of hydrophobic substrates may also strongly influence the reactivity.

Therefore, in attempt to characterise the electronic behaviour of the amino transferase from *Halomonas elongata* (HEWT),[11] we investigated the role of different substituents on aromatic ketones, aldehydes, and the corresponding (*S*)-amines. The effect on the catalytic efficiency of the wild-type enzyme was initially studied with different mono-substituted acetophenones, having both electron withdrawing and donating groups (Table 1, column 3), where L-alanine was used as amino donor. Wild-type HEWT generally resulted poorly active, showing significant conversions only with *meta*- and *para*-nitroacetophenones. The reasons for this limited reactivity can be related to unfavourable electronic effects for substrates with EDG groups, whereas only strong EWG groups allowed significant yields by favouring the reactivity of the amine of PMP towards the carbonyl group (Fig. 1). The absence of reactivity towards *ortho*-nitroacetophenone can be ascribed to steric effects, as previously reported for the lack of activity towards these substrates with the yeast *P. glucozyma* ketoreductase KRED1-Pglu.[12]

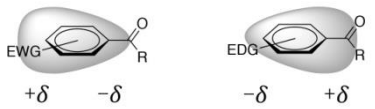


Fig. 1 Representation of the electronic density with EWG and EDG substituted molecules.

Results and discussion

An *in silico* model of the HEWT was generated to aid with the design of key mutations,[13]specifically introduced to probe the steric effect of the substituents. The three-dimensional structure was simulated on the resolved structure of the homologs amino transaminase from *Cromobacterium violaceum.*[14] Due to the medium-high homology between the two proteins (56%), the built model was deemed reliable providing a quality standard score (Z score < 2.5) and it allowed for an accurate evaluation of the catalytic pocket.[15] The enzyme topology for this class of enzymes has been well elucidated and it can be easily described by two pockets

surrounding the cofactor pyridoxal phosphate (PLP).[16,17] The large one accepts bulky groups such as aromatic rings, while the second, smaller one, accommodates short hydrophobic groups.

Ortho- meta- and *para*-nitroacetophenones were selected to map *in-silico* the active site of HEWT (Fig. 2). The steric bulk of the nitro group, when positioned in *ortho* and *meta,* appeared to interfere with tryptophan 56 and tyrosine 149 in the large pocket, as well as with phenylalanine 84 which is located between the two pockets; (Fig 2 A-B). On the other hand, the *para* position is less hindered and the model predicts major interference only with W56 positioned deep into the pocket (Fig.2 C).

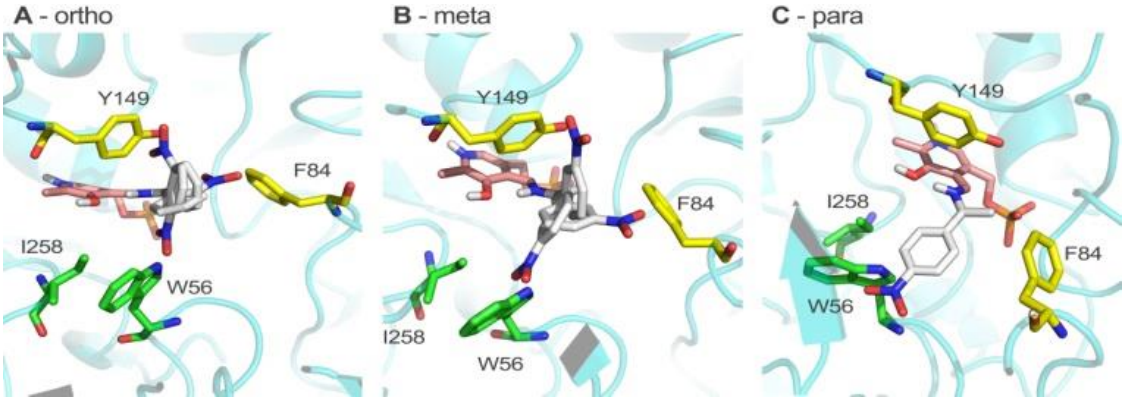
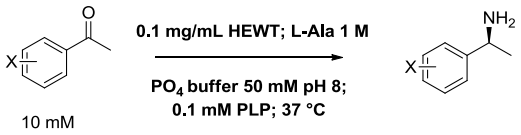


Fig. 2 Model of the catalytic site of HEWT with *ortho- meta-* and *para*-nitroacetophenone.



Entry	Substrate	WT	ee %	W56G	ee %	Y149F	ee %	F84A	ee %	I258A	ee %
1a	Acetophenone	5.0 ± 0.4	>99	6.0 ± 0.1	>99	< 5		5.0 ± 0.3	>99	5.0 ± 0.4	>99
1b	*p*-NO_2	19.6 ± 0.2	>99	15.7 ± 0.5	96	14.2 ± 1.2	97	12.4 ± 0.5	95	19.1 ± 0.9	>99
1c	*m*-NO_2	10.2 ± 0.8	>99	10.7 ± 0.7	95	7.3 ± 0.5	97	6.2 ± 0.6	96	8.6 ± 0.2	95
1d	*o*-NO_2	< 5		15.9 ± 1.6	98	< 5		< 5		< 5	
1e	*p*-F	< 5		< 5		< 5		< 5		< 5	
1f	*m*-F	< 5		< 5		< 5		< 5		< 5	
1g	*o*-F	< 5		<5		< 5		< 5		< 5	
1h	*p*-CF_3	< 5		< 5		< 5		< 5		< 5	
1i	*m*-CF_3	< 5		< 5		< 5		< 5		< 5	
1j	*o*-CF_3	< 5		< 5		< 5		< 5		< 5	
1k	*p*-OCH_3	< 5		< 5		< 5		< 5		< 5	
1l	*m*-OCH_3	< 5		< 5		< 5		< 5		< 5	
1m	*o*-OCH_3	< 5		< 5		< 5		< 5		< 5	
1n	*p*-CH_3	< 5		< 5		< 5		< 5		< 5	
1o	*m*-CH_3	< 5		< 5		< 5		< 5		< 5	
1p	*o*-CH_3	< 5		< 5		< 5		< 5		< 5	

Table 1. Molar conversions of ammination of mono-substituted acetophenones **1a-p** (10 mM) with transaminases (0.1 mg/mL) in phosfate buffer medium 50 mM pH 8, DMSO 10% as cosolvent, 1 M L-Alanine and 0.1 mM PLP. Biotransformations were carried out for 24 h at 37°C under gently agitation.

An *in silico* mutagenesis approach was used to evaluate optimal substitutions.[18] W56G on the large pocket, and F84A near the small one, yielded a lower docking energy favouring the substrates accommodation with a greater effect on the *ortho-* substitution (ESI, Table S.2); Y149A also was indicative of a better docking, however this amino acid is involved in PLP stacking, therefore a more

conservative substitution (Y149F) was selected. Finally, I258 appeared to be critical for enlarging the pocket; it is spatially near W56 and prevents its side chain to rotate. A I258A mutation will allow a higher rotational flexibility to the tryptophan ring leaving more space for the substrate to dock (Fig. 3). Notably, the selected amino acids are highly conserved among this class of amino transaminase (96% among 487 protein sequences aligned and Y149 is 100% retained).[19]

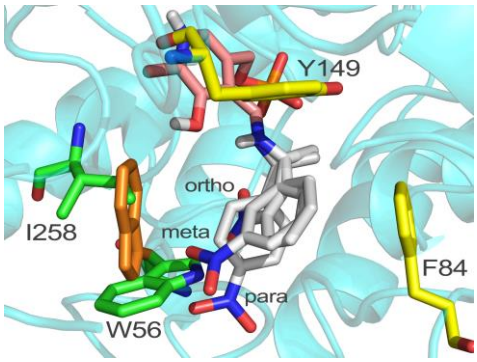
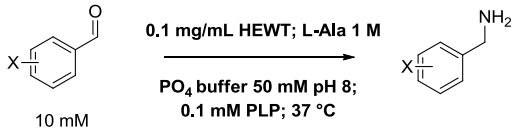


Fig.3. Simulation of tryptophan rotation after the substitution of isoleucine 258 with alanine.

The four generated mutants of HEWT were thus investigated with *para*- *meta*- and *ortho*- mono-substituted acetophenones as amino acceptor (Tab. 1, columns 4-7). The biotransformations were performed over 24 hours and under the reaction conditions selected; the only substrates transformed were the NO_2-acetophenones (entry **1b-d**).

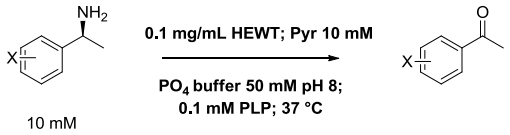


The enantioselectivity was also monitored validating that both the WT and the mutants selectively produced the (*S*)-amines, confirming that the mutations did not affect the stereo-preference of the catalyst. The W56G variant is the only one able to transform the *ortho*-nitroderivative. In fact, the replacement of tryptophan with glycine reduced hindrance and facilitated substrate binding (Fig 2 A). Due to the low reactivity of acetophenones, the same study was performed on more reactive mono-substituted benzaldehydes (Tab 2, entry **2a-p**), using L-alanine as amino donor. As predicted by the molecular modelling, the substitution of bulky amino acids such as W56 and F84 with glycine and alanine respectively, and the replacement of Y149 with phenylalanine (lacking the OH group fundamental for steric and electronic interaction with the substrate), gave an enhanced molar conversion for all the *ortho*-substituted aldehydes into the corresponding primary amines with respect to the wild type transaminase. However, in the case of *meta*-derivatives the activity of the different transaminases was predominantly dependent on substituent bulk and/or hydrophobicity whereas the electronic properties had negligible effects. For *para*-substituted aldehydes, I258A was the best variant both with EWG and EDG, as steric effects at this position prevailed on the electronic ones; mutating I258 allowed for greater flexibility of the neighbouring W56 side chain substantially increasing the size of the active site. The direct mutation of W56 with glycine yielded very similar results with respect to the wild type, highlighting the role of this amino acid in stabilising the *para*-substituents. A number of chiral amines with different substitution patterns on the aromatic ring have also been included in this study (Tab 3, **entry 3a-h**) using as amino acceptor sodium pyruvate. Amines are generally better accepted substrates than the corresponding ketones, and steric rather than electronic effects can easily be identified.

Entry	Substrate	WT	W56G	Y149F	F84A	I258A
2a	Benzaldehyde	22.2 ± 1.9	31.6 ± 0.8	32.0 ± 0.9	27.3 ± 0.9	22.5 ± 0.7
2b	*p*-NO_2	72.4 ± 3.2	70.2 ± 1.0	59.3 ± 0.9	56.5 ± 1.1	80.5 ± 1.0
2c	*m*-NO_2	65.4 ± 1.3	68.4 ± 1.1	57.0 ± 0.8	51.0 ± 1.9	50.6 ± 1.2
2d	*o*-NO_2	69.2 ± 1.2	89.3 ± 1.1	88.8 ± 1.2	80.6 ± 0.4	64.8 ± 1.5
2e	*p*-F	26.2 ± 1.5	35.4 ± 1.1	23.7 ± 0.7	24.7 ± 0.3	52.0 ± 1.3
2f	*m*-F	64.6 ± 3.9	79.2 ± 1.3	80.0 ± 0.7	76.2 ± 1.0	65.3 ± 1.0
2g	*o*-F	67.9 ± 1.5	91.2 ± 1.4	96.1 ± 1.8	94.1 ± 1.0	90.6 ± 0.7
2h	*p*-CF_3	70.8 ± 2.7	67.0 ± 0.8	74.5 ± 0.6	72.9 ± 0.6	97.5 ± 1.4

2i	*m*-CF_3	73.3 ± 2.2	75.5 ± 0.5	70.8 ± 1.2	68.1 ± 0.2	66.7 ± 1.3
2j	*o*-CF_3	90.7 ± 1.3	99.7 ± 0.6	98.2 ± 0.4	94.1 ± 0.2	92.1 ± 0.2
2k	*p*-OCH_3	17.7 ± 1.7	21.1 ± 1.2	14.8 ± 1.5	18.4 ± 0.7	38.1 ± 0.2
2l	*m*-OCH_3	22.6 ± 1.3	26.9 ± 0.5	23.5 ± 2.5	20.3 ± 0.6	24.1 ± 0.6
2m	*o*-OCH_3	19.5 ± 0.3	48.1 ± 0.3	46.9 ± 0.4	41.0 ± 0.2	20.6 ± 0.4
2n	*p*-CH_3	14.1 ± 0.6	18.6 ± 0.7	17.1 ± 0.3	20.3 ± 0.8	28.5 ± 0.5
2o	*m*-CH_3	25.5 ± 0.8	35.2 ± 1.0	27.0 ± 0.2	20.4 ± 0.6	26.9 ± 1.7
2p	*o*-CH_3	39.6 ± 1.4	47.1 ± 0.7	49.9 ± 0.4	44.4 ± 0.7	36.5 ± 1.6

Table 2. Molar conversions of ammination of mono-substituted benzaldehydes **2a-p** (10 mM) with transaminases (0.1 mg/mL) in phosfate buffer medium 50 mM pH 8, DMSO 10% as cosolvent, 1 M alanine and 0.1 mM PLP. Biotransformations were carried out for 24 h at 37°C under gently agitation



Entry	Substrate	WT	W56G	Y149F	F84A	I258A
3a	(*S*)-PEA	99.9 ± 0.2	99.7 ± 0.3	99.5 ± 0.5	99.6 ± 0.2	83.5 ± 1.0
3b	(*S*)-*p*-F	99.9 ± 0.1	95.4 ± 1.3	95.4 ± 0.1	47.7 ± 1.3	95.9 ± 0.4
3c	(*S*)-*m*-F	99.9 ± 0.2	95.5 ± 0.2	91.7 ± 1.4	34.8 ± 1.9	93.2 ± 2.5
3d	(*S*)-*o*-F	82.9 ± 2.0	84.6 ± 1.4	68.3 ± 1.2	12.8 ± 0.1	20.2 ± 0.1
3e	(*S*)-*p*-OCH_3	99.3 ± 0.6	98.0 ± 1.0	97.9 ± 0.6	70.7 ± 0.6	98.3 ± 0.1
3f	(*S*)-*m*-OCH_3	97.9 ± 1.0	94.8 ± 0.9	97.9 ± 0.3	75.3 ± 0.2	98.2 ± 0.6
3g	(*S*)-*o*-OCH_3	71.7 ± 1.9	78.0 ± 1.1	46.2 ± 2.4	10.3 ± 1.0	72.1 ± 1.2
3h	(*S*)-*p*-CH_3	99.7 ± 0.3	98.6 ± 1.5	97.3 ± 0.8	76.7 ± 0.6	98.4 ± 0.1

Table 3. Molar conversions of oxidation of mono-substituted amines **3a-h** (10 mM) with transaminases (0.1 mg/mL) in phosfate buffer medium 50 mM pH 8, DMSO 10% as cosolvent, 10 mM pyruvate and 0.1 mM PLP. Biotransformations were carried out for 24 h at 37°C under gently agitation

Interestingly, among the mutants, only W56G had better activity with the *ortho*-substituents (*o*-F- and *o*-OCH_3-amine), while the other mutations showed a major destabilizing effect. The molar conversion of F84A indeed dropped to just over 10% conversion with both substrates in the same reaction time. This could be attributed to an impairment effect of the first step in the catalytic cycle where the substrate displaces the catalytic lysine from the PLP. F84 appears essential in the case of substituted aromatic amines in stabilising the binding of the substrate as well as assisting the catalytic mechanism. However, this effect is not observed with the natural substrate (*S*)-PEA probably due to a very efficient reaction that masks the negative impact of the mutation. These data were also confirmed by very different K_m for *o*-amines with W56G and F84A compared with wild type (0.61 *o*-F_{W56G}; 1.80 *o*-F_{F84A}; 1.38 *o*-F_{WT} and 0.45 *o*-$OCH_{3\ W56G}$; 3.97 *o*-$OCH_{3\ F84A}$; 1.98 *o*-$OCH_{3\ WT}$ expressed in mM). W56G and I258A, predicted as the best mutated transaminases for *para*-derivatives, in oxidising conditions showed very similar results with respect to wild type enzyme. The mutations can effectively alter the efficiency of the enzymes in just one direction, therefore also altering the equilibrium of the reaction.

Experimental

Expression and purification of the HEWT in E. coli

Protein expression and purification was performed following previously reported protocols.[11]The proteins yields after the IMAC purification were estimate to be: WT 100 mg, W56G 93 mg, F84A 47 mg, Y149G 32 mg, and I258A 88 mg for 1 Liter of expression media. The assay derived from Schatzle, *et al*.[20] was used as standard enzymatic assay following the procedure applied in Cerioli, *et al*.[11] to assess the enzymes activity in standard conditions and evaluate then their stability (see ESI).

HEWT mutant generation

The SupC gene of HEWT harboured in a pRSETb plasmid was mutated using the QuikChangeXL single point mutation kit provided by Agilent Technologies®. The oligonucleotide primers designed exploiting the QuickChange Primer Design tool (Agilent Technologies®)

were synthesised from Eurofins Genomics®. The four HEWT mutants were achieved using the following primers: W56G 5'-cggcatggccgggcttggctgcgtgaatctc-3'; F84A 5'-ctgccgtactacaacaccgcctttcaagaccacgcatcc-3'; Y149F 5'-gtcgcgagaacgcctttcacggctcca-3'; I258A 5'-cgccgacgaggt ggccctgcggcttcggg-3' (mutated codon is underlined).

Ketones and aldehydes reactions

The batch reactions to evaluate the conversion after 24 h, with wild-type and mutated HEWT, were performed in 1.5 mL micro centrifuge tubes; 200 µL reaction mixture in 50 mM phosphate buffer pH 8.0, containing 10 mM amino acceptor substrate dissolved in 10% DMSO, 1.0 M L-Alanine (amino donor substrate), 0.1 mM PLP, and 0.1 mg/mL enzyme, was left under gentile shaking at 37 °C. 10 µL aliquots were quenched with trifluoroacetic acid (TFA) 0.2% at 24 hours and then analysed by HPLC equipped with a Supelcosil LC-18-T column (250 mm x 4.6 mm, 5 µm particle size; Supelco, Sigma-Aldrich, Germany). The compounds were detected using an UV detector at 210 nm, or 250 nm after an isocratic run with 25% acetonitrile/75% water with TFA 0.1% v/v at 25 °C with a flow rate of 1 mL/min. The retention times in minutes are listed in table S.1. Calibration curve for substrate and product were carried out with commercial standard and the conversion was then estimated, when possible, both from the substrate signal decreasing and product signal forming.

Amines reactions

The batch reactions to evaluate the conversion after 24h, with wild-type and mutated HEWT, were performed in 1.5 mL micro centrifuge tubes; 200 µL reaction mixture in 50 mM phosphate buffer pH 8.0, containing 10 mM amino donor substrate dissolved in 10% DMSO, 10 mM Pyruvate (amino acceptor substrate), 0.1 mM PLP, and 0.1 mg/mL enzyme, was left under gentile shaking at 37 °C. The conversion was evaluated by HPLC with the same procedures described above. Retention times are listed in Table S.1 in ESI. The kinetic profiles for the amine substrates were measured using a Bio-Tek EPOC2 microplate reader. The reaction mixture was set up in 300 µL of activity buffer (50 mM phosphate buffer pH 8.0) containing a variable concentration of substrate amino donor (0.01 to 2 mM), 10 mM of pyruvate, 0.1 mM PLP and an appropriate amount of enzyme. The reaction rate was detected at 245 nm over 20 minutes at 25°C. The molar extinction coefficient was calculated for each product in order to calculate the reaction rate in mM/s (data not shown). The data were elaborated by fitting them with the Lineweaver–Burk double reciprocal plot which yielded the Michaelis-Menten parameters (Km and Kcat).

Chirality evaluation

To evaluate the enantioselectivity of the enzymes, the chiral amine products were assessed by chiral HPLC analysis, following derivatization to the corresponding amides by adapting the method reported by Andrade et al. with Codex® ATA Screening Kit. Addition of 100 µL 5 M NaOH to the sample was followed by extraction in 900 µL methyl-tert-butyl ether (MTBE) and then by addition of 5 µL triethylamine and 5 µL acetic anhydride. A normal-phase HPLC method was performed on an Agilent Technologies 1200 equipped with a Chiralpak® IC (250mm x 4.6 mm, 5 µm particle size). Detection at 254 nm after an isocratic run 95:05 Heptane:Ethanol at 25 °C with a flow rate of 1 mL/min. Employing this method, product and starting materials were separated with the following retention times (in minutes): acetophenone (7.1), derivatized (R)-1-phenylethylamine (16.7) and (S)-1-phenylethylamine (19.1), 4-nitroacetophenone (12.4), derivatized (R)-4-nitrophenylethylamine (22.4) and (S)-4-nitrophenylethylamine (24.2), 3-nitroacetophenone (13.6), derivatized (R)-3-nitrophenylethylamine (23.7) and (S)-3-nitrophenylethylamine (26.7), 2-nitroacetophenone (14.2), derivatized (R)-3-nitrophenylethylamine (24.8) and (S)-3-nitrophenylethylamine (27.1) .

Protein modelling and structural analysis

The protein model was build using the SWISS-MODEL Homology Modeling[15] tool, superimposing the HEWT protein sequence on the resolved structure of the homologues amino transaminase from *Cromobacterium violaceum* (PDB: 4a72).[14] Substrate docking was performed exploiting the open source tool Autodock Vina.[21] The results elaboration and the structural evaluation was achieved using the molecular visualization system PyMOL (open source license). Sequence analysis and alignment was performed with CunSurf, software for multiple blast provided by BioSoft.[22]

Conclusions

Transaminase from *Halomonas elongata* (HEWT) showed a wide variety of catalytic abilities towards different aromatic substrates (ketones, aldehydes, amines). Acetophenone derivatives were poor substrates for the wild-type enzyme, since only *meta*- and *para*-NO_2 acetophenones were reduced with molar conversions ranging from 10 to 20%. Four mutants were designed following *in silico* substrate-enzyme binding mode studies. Mutant W56G proved able to catalyse also the amination of *ortho*-NO_2 acetophenone, as a result of reduced interferences and facilitated substrate binding. Variants W56G and F84A, which have an enlarged binding pocket, showed better performances with *ortho*-substituted benzaldehyde derivatives, whereas I258A was the best mutant for *para*-derivatives. Since the investigated mutations involved highly conserved amino acids, these findings might reflect the behaviour of homologs ω-transaminases. Therefore, as general rule, the mutation of large pocket tryptophan with a smaller residue yields to an improvement in the reactivity of *ortho*-substituted aromatic compounds. The change of the large pocket isoleucine provides instead an increase in activity of *para*-substituted molecules. On the other hand, F84 between the two

pockets was very sensible to substitution and when it was changed for smaller alanine, it gave a worst conversion of *ortho* and *para* aromatic amines. Consequently, this position has a fundamental role in the oxidation step of amino transaminase and it should be preserved in order to maintain the optimal enzyme efficiency.

Acknowledgements

The authors wish to thank Cariplo Foundation for funding the project "INnovative Biocatalytic OXidations – INBOX" (Rif.2014-0568) "Ricerca integrata biotecnologie industriali" (Dr. Contente) and Science Foundation Ireland through the Synthesis and Solid State Pharmaceutical Centre grant number 12/RC/2275 (Planchestainer and reagents).

Notes and references

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

2 D. Koszelewski, I. Lavandera, D. Clay, D. Rozzell and W. Kroutil, *Adv. Synth. Catal.*, 2008, **350**, 2761.

3 J. Rudat, B. R. Brucher and C. Syldatk, *AMB Express*, 2012, **2**, 11.

4 D. Koszelewski, K. Tauber, K. Faber, W. Kroutil, *Trends Biotechnol.*, 2010, **28**, 324.

5 K. E. Cassimjee, B. Manta and F. Himo *Org. Biomol. Chem*., 2015, **13**, 8453.

6 G. D. Smith, R. Harrison and R. Eisenthal, *Neurochem. Res.*, 1996, **21**, 1061.

7 J. F. Kirsch, G. Eichele, G. C. Ford, M. G. Vincent, J. N. Jansonius, H. Gehring and P. Christen, *J. Mol. Biol.*, 1984, **174**, 497.

8 J. S. Shin and B. G. Kim, *J. Org. Chem.*, 2002, **67**, 2848.

9 D. Koszelewski, I. Lavandera, D. Clay, G. M. Guebitz, D. Rozzell and W. Kroutil, *Angew. Chemie - Int. Ed.*, 2008, **47**, 9337.

10 E. S. Park, S. R. Park, S. W. Han, J. Y. Dong and J. S. Shin, *Adv. Synth. Catal.*, 2014, **356**, 212.

11 L. Cerioli, M. Planchestainer, J. Cassidy, D. Tessaro and F. Paradisi, *J. Mol. Catal. B Enzym.*, 2015, **120**, 141.

12 M. L. Contente, I. Serra, L. Palazzolo, C. Parravicini, E. Gianazza, I. Eberini, A. Pinto, B. Guidi, F. Molinari, D. Romano, *Org. Biomol. Chem*, 2016, **14**, 3404.

13 S. Sirin, R. Kumar, C. A. Martinez, M. J. Karmilowicz, P. Ghosh, Y. A. Abramov, V. Martin and W. Sherman, *J. Chem. Inf. Model.*, 2014, **54**, 2334.

14 M. S. Humble, K. E. Cassimjee, M. Hãkansson, Y. R. Kimbung, B. Walse, V. Abedi, H. J. Federsel, P. Berglund and D. T. Logan, *FEBS J.*, 2012, **279**, 779.

15 K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195.

16 E. S. Park, S. R. Park, S. W. Han, J. Y. Dong and J. S. Shin, *Adv. Synth. Catal.*, 2014, **356**, 212.

17 B.-K. Cho, H.-Y. Park, J.-H. Seo, J. Kim, T.-J. Kang, B.-S. Lee and B.-G. Kim, *Biotechnol. Bioeng.*, 2008, **99**, 275.

18 A. Nobili, Y. Tao, I. V Pavlidis, T. van den Bergh, H.-J. Joosten, T. Tan and U. T. Bornscheuer, *Chembiochem*, 2015, **16**, 1.

19 F. Steffen-Munsberg, C. Vickers, H. Kohls, H. Land, H. Mallin, A. Nobili, L. Skalden, T. van den Bergh, H.-J. Joosten, P. Berglund, M. Höhne and U. T. Bornscheuer, *Biotechnol. Adv.*, 2015, **33**, 566.

20 S. Schätzle, M. Höhne, E. Redestad, K. Robins and U. T. Bornscheuer, *Anal. Chem.*, 2009, **81**, 8244.

21 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2009, **31**.

22 H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko and N. Ben-Tal, *Submitted*, 2016, 1.