ARTICLES
PUBLISHED ONLINE: 29 APRIL 2012 | DOI: 10.1038/NCHEM.1338
nature chemistry

Engineering methylaspartate ammonia lyase for the asymmetric synthesis of unnatural amino acids

Hans Raj[1], Wiktor Szymański[2,3], Jandré de Villiers[1], Henriëtte J. Rozeboom[4], Vinod Puthan Veetil[1], Carlos R. Reis[1], Marianne de Villiers[1], Frank J. Dekker[5], Stefaan de Wildeman[6], Wim J. Quax[1], Andy-Mark W. H. Thunnissen[4], Ben L. Feringa[3], Dick B. Janssen[2] and Gerrit J. Poelarends[1]*

The redesign of enzymes to produce catalysts for a predefined transformation remains a major challenge in protein engineering. Here, we describe the structure-based engineering of methylaspartate ammonia lyase (which in nature catalyses the conversion of 3-methylaspartate to ammonia and 2-methylfumarate) to accept a variety of substituted amines and fumarates and catalyse the asymmetric synthesis of aspartic acid derivatives. We obtained two single-active-site mutants, one exhibiting a wide nucleophile scope including structurally diverse linear and cyclic alkylamines and one with broad electrophile scope including fumarate derivatives with alkyl, aryl, alkoxy, aryloxy, alkylthio and arylthio substituents at the C2 position. Both mutants have an enlarged active site that accommodates the new substrates while retaining the high stereo- and regioselectivity of the wild-type enzyme. As an example, we demonstrate a highly enantio- and diastereoselective synthesis of *threo*-3-benzyloxyaspartate (an important inhibitor of neuronal excitatory glutamate transporters in the brain).

Optically pure α-amino acids are highly valuable as tools for biological research and as chiral building blocks for pharmaceuticals and (agro)chemicals[1]. Industrial and academic interest in these compounds, combined with the potential advantages of replacing conventional chemical processes with biocatalysis, has fueled the development of enzymatic synthesis routes for enantiomerically pure α-amino acids. Various enzymatic synthesis routes have been developed, for example using hydantoinases, dehydrogenases, acylases, aminotransferases, ammonia lyases and amidases[1–5], the most attractive of which are those based on asymmetric synthesis in which a prostereogenic substrate is converted to an optically pure α-amino acid with 100% theoretical yield. For example, the asymmetric addition of ammonia to the double bonds of unsaturated acids catalysed by ammonia lyases is a very attractive strategy for the synthesis of chiral α-amino acids[5]. This strategy makes use of readily available starting substrates without a requirement for cofactor recycling, implementation of dynamic kinetic resolution strategies or additional catalysts. However, it is limited by the narrow substrate range of ammonia lyases.

3-Methylaspartate ammonia lyase (MAL) catalyses the reversible addition of ammonia (**2a**) to mesaconate (**1b**) to yield *threo*-(2*S*,3*S*)-3-methylaspartate (*threo*-**3a**) and *erythro*-(2*S*,3*R*)-3-methylaspartate (*erythro*-**3a**) as products (Table 1)[6–10]. MAL is used by the bacterium *Clostridium tetanomorphum* as part of a degradative pathway that converts (*S*)-glutamic acid via *threo*-**3a** to yield acetyl-coenzyme A[6,11]. The crystal structure of MAL and that of the isozyme from *Citrobacter amalonaticus* have been solved by X-ray crystallography[12,13]. Based on kinetic isotope measurements[14,15], structural studies[12,13,16] and mutagenesis experiments[8], a mechanism has emerged for the MAL-catalysed reaction. In this proposed mechanism, an (*S*)- or (*R*)-specific catalytic base abstracts the C3 proton of the respective stereoisomer of **3a** to generate an enolate intermediate that is stabilized by coordination to the essential active site Mg^{2+} ion. Collapse of this intermediate results in the elimination of ammonia and yields mesaconate.

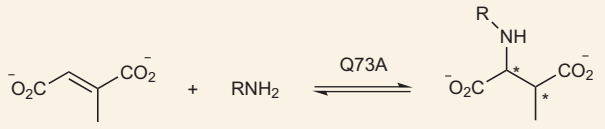
Unfortunately, the substrate scope of MAL is very narrow. Only with a few small substituted amines and fumarates is a reasonable activity observed, yielding a limited number of substituted aspartic acids[17–19]. Accordingly, it would be very attractive to extend the accessible range of aspartic acid derivatives by redesigning MAL to convert various unnatural substrates, thereby enlarging its biocatalytic applicability. In this Article, we describe the use of an efficient engineering strategy based on saturation mutagenesis at carefully chosen sites lining the substrate binding pocket[20–24], with a view to expanding both the nucleophile and electrophile spectrum of MAL. We have obtained two single-active-site mutants, one that accepts a large set of non-native amines and one that accepts various non-native fumarate derivatives. These engineered MALs have potential as catalysts for the asymmetric synthesis of a large variety of substituted aspartic acids, which are valuable as tools for biological research and as building blocks for chemical and pharmaceutical synthesis[25–31]. This potential is illustrated by the mutant MAL-catalysed synthesis of *threo*-3-benzyloxyaspartate, a widely used non-transportable blocker for all subtypes of neuronal excitatory amino-acid transporters[25,26].

Results

Expansion of the nucleophile spectrum of MAL. The crystal structure of MAL complexed with the natural substrate *threo*-**3a**[13]

[1]Department of Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands, [2]Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, [3]Center for Systems Chemistry, Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, [4]Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands, [5]Department of Pharmaceutical Gene Modulation, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands, [6]DSM Pharmaceutical Products, PO Box 18, 6160 MD, Geleen, The Netherlands. *e-mail: g.j.poelarends@rug.nl

Table 1 | Q73A-catalysed amine (2a–2u) addition to mesaconate (1b).



1b 2 3

Entry	Nucleophile	R-group	Product	Conv. (%)*	d.e. (%)	e.e. (%)
1	**2a**	H	**3a**	79†	5 (*threo*)‡	ND‖
2	**2b**	Methyl	**3b**	74	>95 (*threo*)‡	ND
3	**2c**	Ethyl	**3c**	51	>95 (*threo*)‡	>99¶
4	**2d**	Propyl	**3d**	53	>95 (*threo*)‡	>99¶
5	**2e**	Butyl	**3e**	60	>95 (*threo*)‡	>99¶
6	**2f**	Pentyl	**3f**	29	>95 (*threo*)‡	ND
7	**2g**	Hexyl	**3g**	17	>95 (*threo*)‡	ND
8	**2h**	Isopropyl	**3h**	8	>95 (*threo*)‡	>99¶
9	**2i**	Cyclobutyl	**3i**	65	>95 (*threo*)‡	>99¶
10	**2j**	Cyclopentyl	**3j**	36	>95 (*threo*)‡	ND
11	**2k**	Cyclohexyl	**3k**	21	>95 (*threo*)‡	ND
12	**2l**	(Cyclopropyl)methyl	**3l**	61	>95 (*threo*)‡	>99¶
13	**2m**	Benzyl	**3m**	6	>95 (*threo*)‡	ND
14	**2n**	Cyclopropyl	**3n**	75	>95 (*threo*)§	ND
15	**2o**	Ethoxy	**3o**	99	>95 (*threo*)§	ND
16	**2p**	2-Hydroxyethyl	**3p**	37†	>95 (*threo*)§	ND
17	**2q**	3-Hydroxypropyl	**3q**	56	80 (*threo*)§	ND
18	**2r**	2-Methoxyethyl	**3r**	60	>95 (*threo*)§	ND
19	**2s**	*N*-Methyl-2-aminoethyl	**3s**	80	>95 (*threo*)§	ND
20	**2t**	2-Aminoethyl	**3t**	90	>95 (*threo*)§	ND
21	**2u**	3-Aminopropyl	**3u**	78	>95 (*threo*)§	ND

*Reactions were allowed to proceed for 7 days to detect the formation of low amounts of *erythro* product isomers, if present. †Reaction was allowed to proceed for 14 days. ‡The d.e. (defined as excess of *threo* isomer over *erythro* isomer) of the amino-acid product was determined by comparison of its ^{1}H NMR signals in the crude reaction mixture to those of synthesized authentic standards with known *threo* or *erythro* configuration. §The purified amino-acid product could be tentatively assigned the *threo* configuration on the basis of analogy. ‖ND, not determined. ¶The e.e. of the isolated product was determined by chiral HPLC using a synthesized authentic standard with *threo*-(DL) configuration.

suggests that the side chains of Q73 and Q172 are involved in the formation of the binding pocket for the amino group of the substrate (Fig. 1a). To engineer this putative amine binding pocket, three focused libraries were generated by saturation mutagenesis at positions 73, 172 and 73/172, using NNS degeneracy[32]. As the model reaction for activity screening, the addition of methylamine (**2b**) to mesaconate (**1b**) was chosen (Table 1). This small alkylamine is a poor substrate for wild-type MAL. We reasoned that if mutants displaying substantial activity were to be found in any of the libraries, then they could also be tested for their ability to process larger and more challenging amines (for example, amines **2c–2u**).

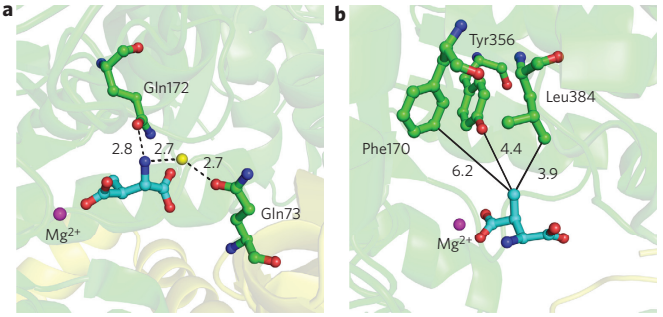


Figure 1 | Crystal structure of wild-type MAL in complex with the natural substrate *threo*-(2*S*,3*S*)-3-methylaspartate. a, Close-up of the active site showing the hydrogen-bond interactions between the substrate's amino group and the side chains of Gln73 (via a water molecule) and Gln172. The carbon atoms of the active-site residues are shown in green, and those of the substrate in cyan. Hydrogen bonds are represented as dashed lines. The magnesium ion and water molecule are shown as magenta and yellow spheres, respectively. **b**, Close-up of the active site showing the observed distances (in ångstroms, atoms connected by solid lines) between the substrate's methyl group and the side chains of the three residues (Phe170, Tyr356 and Leu384) that are involved in the formation of the alkyl binding pocket. Colour scheme as in **a**. The figures were prepared with PyMOL (http://www.pymol.org).

The three libraries were used to transform *Escherichia coli* cells. The double-site library was screened by evaluating ~1,200 transformants, and ~400 transformants were evaluated for each single-site library. Although the library at position 172 failed to contain any active mutants, screening of the libraries corresponding to positions 73 and 73/172 resulted in the identification of six mutants with pronounced activity towards methylamine addition. DNA sequencing revealed that all positive hits carry mutations at position 73, namely Q73A, Q73G, Q73N, Q73P, Q73S and Q73T. These mutants and the wild-type enzyme were overexpressed, purified to homogeneity, and assayed for their ability to catalyse ammonia or methylamine addition to mesaconate. When compared to wild-type MAL, all mutants display a significant decrease in ammonia addition activity and a large increase in methylamine addition activity (Supplementary Fig. S1). The Q73A mutant showed the highest catalytic rate for methylamine addition to mesaconate and was therefore selected for further study.

A comparison of the k_{cat}/K_m values shows that the Q73A mutant is at least 140-fold more efficient in the addition of methylamine than wild-type MAL (Supplementary Table S1). Wild-type MAL, however, is ~500-fold more efficient in the ammonia addition reaction. Hence, the Q73A mutation moved the specificity of MAL away from ammonia and towards methylamine (>70,000-fold shift in nucleophile specificity as defined by k_{cat}/K_m values). ^{1}H NMR spectroscopic analysis showed that the Q73A-catalysed ammonia and methylamine additions to mesaconate lead to the same amino-acid products as the corresponding wild-type MAL-catalysed reactions

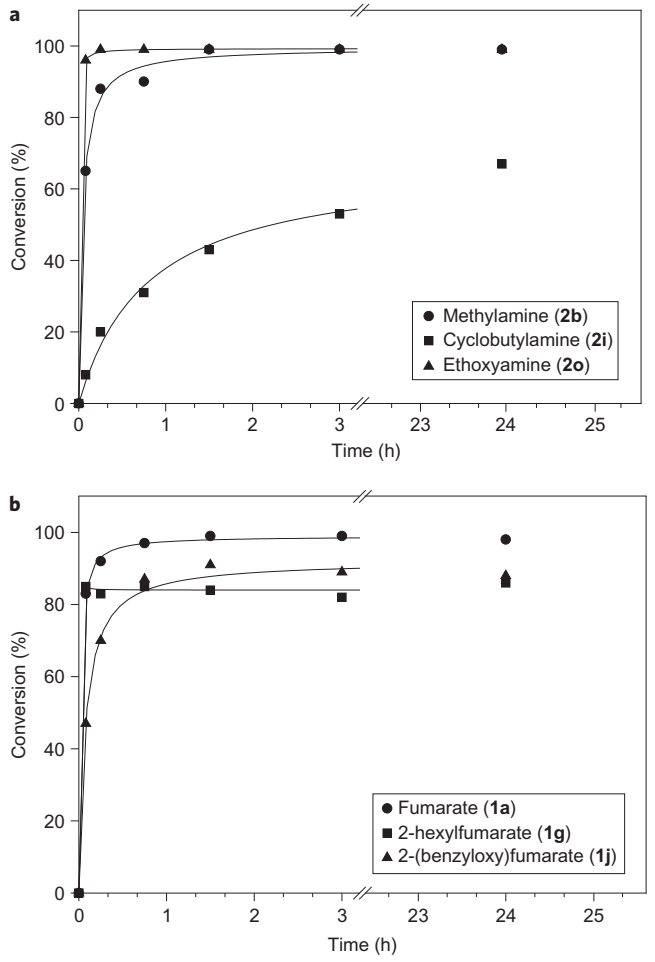


Figure 2 | Enzyme-catalysed transformations under optimized reaction conditions. a, Progress curves for the Q73A-catalysed (0.05 mol%) addition of methylamine (**2b**, 3 M), cyclobutylamine (**2i**, 2 M) or ethoxyamine (**2o**, 3 M) to mesaconate (26 mM), as monitored by ^{1}H NMR spectroscopy. **b**, Progress curves for the L384A-catalysed (0.05 mol%) addition of ammonia (5 M) to fumarate (**1a**), 2-hexylfumarate (**1g**) or 2-(benzyloxy)fumarate (**1j**) (each at 30 mM), as monitored by ^{1}H NMR spectroscopy.

(Supplementary Fig. S2). This demonstrates that the Q73A mutation does not affect the regio- or diastereoselectivity of the enzyme. Whereas the enzyme-catalysed ammonia addition reaction yields an ~1:1 mixture of *threo* and *erythro* isomers of 3-methylaspartate (Table 1, **3a**)[8], the enzyme-catalysed methylamine addition reaction is highly diastereoselective and gives exclusively the *threo* isomer of *N*,3-dimethylaspartate (**3b**) (Table 1, Supplementary Fig. S2)[19].

To investigate its nucleophile scope, we used ^{1}H NMR spectroscopy to determine the ability of the Q73A mutant to add a variety of structurally different amines (**2c–2u**, Table 1) to mesaconate (tenfold molar excess of amine over mesaconate, 0.01 mol% of biocatalyst, room temperature). These amines are not accepted as substrates by wild-type MAL, with the exception of **2c**, **2n**, **2p** and **2t**, which show minimal conversion (<3%) following prolonged (two weeks) incubation. Given that MAL catalyses the fast *anti*-addition and much slower *syn*-addition of ammonia to mesaconate[8], leading to *threo* and *erythro* isomers of the corresponding product, respectively, we allowed the unnatural amine addition reactions to run for seven days to detect the formation of low amounts of *erythro* product isomers, if present. Remarkably, the Q73A mutant processed all amines tested (Table 1), including those having bulky substituents such as hexyl, cyclohexyl and benzyl, which demonstrates its surprisingly broad nucleophile scope. Control experiments showed that the amines do not react with mesaconate in the absence of enzyme.

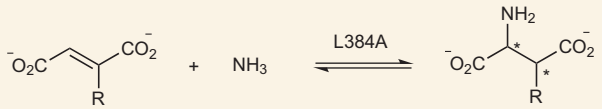
For amines **2c–2m** (Table 1, entries 3–13), the single products of the Q73A-catalysed additions to mesaconate were identified as the *threo* isomers of the corresponding *N*-substituted 3-methylaspartic acids (**3c–3m**) by comparison of their ^{1}H NMR signals in the crude reaction mixtures to those of chemically synthesized authentic standards with known relative configuration (Supplementary Fig. S3). No other regio- or diastereoisomers were observed. For a few selected amines (**2c–2e**, **2h**, **2i** and **2l**), preparative-scale reactions were performed (1.2 mmol **1b**). Purification of the amino-acid products and analysis by high-performance liquid chromatography (HPLC) on a chiral stationary phase (Supplementary Fig. S4) revealed that the Q73A mutant is also highly enantioselective, producing **3c–3e**, **3h**, **3i** and **3l** with >99% enantiomeric excess (e.e.) (Table 1). For the other amines (Table 1, entries 14–21), the Q73A-catalysed addition reactions were performed on a 3.1 mmol scale (**1b**). The products were isolated and identified as the corresponding *N*-substituted 3-methylaspartic acids (**3n–3u**). Although the relative configuration of products **3n–3u** has not been determined by comparison to authentic standards, we assume the relative configuration to be *threo* for all products **3b–3u**.

To further demonstrate the preparative usefulness of the Q73A mutant, we optimized the reaction conditions for a few structurally distinct amines (**2b**, **2i** and **2o**). With 0.05 mol% biocatalyst and a 115-fold molar excess of **2b** or **2o** (3 M each) over mesaconate (26 mM), the reactions were complete within 20–60 min at 25 °C, achieving final conversions of almost 100% (Fig. 2a). For amine **2i**, a lower concentration (2 M) had to be used to avoid protein precipitation. Nevertheless, the reaction reached equilibrium (~65% conversion) in just a few hours. These results clearly demonstrate the potential of the Q73A mutant for application in the selective synthesis of various *N*-substituted 3-methylaspartic acids.

Expansion of the electrophile spectrum of MAL. The structure of MAL in complex with *threo*-**3a**[13] suggests that the side chains of F170, Y356 and L384 are involved in the formation of the binding pocket for the substrate's methyl group (Fig. 1b). To engineer this putative alkyl binding pocket, three focused libraries were generated by saturation mutagenesis at positions 170, 356 and 384 using NNS degeneracy. As the model reaction for activity screening, the addition of ammonia to 2-hexylfumarate (Table 2, **1g**) was chosen. Compound **1g** is not converted at all by wild-type MAL, even after an incubation period of several weeks. The hope was that active mutants with the ability to aminate electrophiles with large substituents at the C2 position (for example, **1e–1g**, **1i**, **1j**, **1l–1n**) could be identified in this way.

The three libraries were introduced into *E. coli* cells, after which the library corresponding to position 384 was screened by evaluating ~400 transformants; ~200 transformants were evaluated for each of the other two libraries. Although the libraries corresponding to positions 170 and 356 failed to give any active mutants, screening of the library corresponding to position 384 resulted in several positive hits. Analysis by DNA sequencing revealed that all positive hits carry either an alanine or glycine mutation at position 384. These two mutants and the wild-type enzyme were overexpressed, purified to homogeneity, and tested for their ability to catalyse the addition of ammonia to either mesaconate or 2-hexylfumarate. Gratifyingly, the L384G and L384A mutants exhibit pronounced amination activity towards 2-hexylfumarate (Supplementary Fig. S5). Both mutants display a significant decrease in amination activity towards mesaconate when compared to wild-type MAL. In other words, the mutations moved the electrophile specificity of MAL away from the natural substrate mesaconate and towards the unnatural substrate 2-hexylfumarate. Because the L384A mutant showed

Table 2 | MAL- and L384A-catalysed ammonia addition to fumarates 1a–1n.



1 2a 4

Entry	Electrophile	R-group	Product	Conv. (%)* wild-type	Conv. (%)* L384A	Conv. (%)*,† L384A	d.e. (%)§	e.e. (%)
1	**1a**	H	**4a**	100	99	99 (80)‡	-	>99 (*S*)§§
2	**1b**	Methyl	**4b**	80	75	73 (60)‡	56 (*threo*)‖	ND‖‖
3	**1c**	Ethyl	**4c**	74	76	73 (45)‡	>95 (*threo*)‖	ND
4	**1d**	Propyl	**4d**	57	67	66 (46)‡	>95 (*threo*)‖	ND
5	**1e**	Butyl	**4e**	0	59	57 (36)‡	2¶	ND
6	**1f**	Pentyl	**4f**	0	52	52 (23)‡	>95¶	ND
7	**1g**	Hexyl	**4g**	0	53	53 (48)‡	>95¶	ND
8	**1h**	Ethoxy	**4h**	46	43	45 (24)‡	70 (*threo*)**	ND
9	**1i**	Phenoxy	**4i**	0	44	46 (20)‡	>95¶	ND
10	**1j**	Benzyloxy	**4j**	0	60	65 (50)‡	>95 (*threo*)††	>99¶¶
11	**1k**	Ethylthio	**4k**	36	50	81 (22)‡	36 (*threo*)‡‡	ND
12	**1l**	Phenylthio	**4l**	0	34	45 (13)‡	30 (*threo*)‡‡	ND
13	**1m**	Benzylthio	**4m**	0	42	55 (30)‡	20 (*threo*)‡‡	ND
14	**1n**	Benzyl	**4n**	0	90	90 (55)‡	>95 (*threo*)**	ND

*Reactions were allowed to proceed for 7 days. †Preparative scale. ‡Yield (%) of isolated product after ion exchange chromatography. §The d.e. is defined as excess of *threo*-isomer over *erytho*-isomer. ‖The relative configuration of the isolated product(s) was assigned by ^{1}H NMR using the product(s) of the corresponding wild-type MAL-catalysed reaction as reference sample and by referring to the literature[17]. ¶The relative configurations have not been determined. **The relative configuration of the isolated product(s) was assigned by referring to the literature[29,31]. ††The relative configuration of isolated product was assigned by ^{1}H NMR spectroscopy using an authentic standard with known *threo* configuration and by referring to the literature[30]. ‡‡The isolated amino-acid products were first esterified and acetylated, after which the relative configuration of the products could be tentatively assigned by ^{1}H NMR spectroscopy on the basis of comparison to chemically synthesized reference molecules (Supplementary Methods). §§The e.e. value of isolated product was determined by chiral HPLC using authentic standards with known *R* or *S* configuration. ‖‖ND, not determined. ¶¶The e.e. value of isolated product was determined by chiral HPLC using authentic *threo*-(DL)-**4j** as standard.

higher activity towards a range of fumarate derivatives, it was selected for further study.

Kinetic parameters were determined for the L384A mutant using substrates **1a–1g**, which have varying 2-alkyl chain length (Table 2), and compared to those measured for the wild-type enzyme (Supplementary Table S2). As previously reported[17,18], wild-type MAL only displays activity towards fumarates with a short alkyl substituent (**1a–1d**). Mutant L384A, however, shows activity for all tested substrates. It displays both higher affinity (lower K_m values) and higher activity (larger k_{cat} values) for substrates with a long alkyl chain (**1d–1g**) than for those with a short alkyl chain (**1a–1c**). Although wild-type MAL is at least 100-fold more efficient in the amination of **1a–1c**, the L384A mutant is ~13-fold more efficient in the amination of **1d**. In the case of the large substrates **1e–1g**, which are not converted at all by wild-type MAL, it was not possible to determine the rate enhancement, but it must be much higher than that observed for **1d**.

To further explore the electrophile scope of the L384A mutant, its ability to catalyse the amination of substrates with aryl, alkoxy, aryloxy, alkylthio and arylthio substituents at the C2 position (**1h–1n**) was tested using ^{1}H NMR spectroscopy (Table 2). The wild-type enzyme only displays amination activity for compounds **1h** and **1k**, which, like substrates **1a–1d**, have small R-groups at the C2 position (Table 2, entries 1–4, 8, 11). Strikingly, the L384A mutant exhibits activity towards all substrates tested (Table 2), demonstrating its broad electrophile scope. Control experiments showed that the fumarate derivatives do not undergo amination in the absence of enzyme.

The L384A-catalysed aminations of **1a–1n** were also performed on a preparative (3.1 mmol) scale (Table 2). Reactions (tenfold molar excess of ammonia over unsaturated acid, 0.01 mol% of biocatalyst, room temperature) were allowed to run for seven days to detect the formation of low amounts of *erythro* product isomers, if present. The products were isolated and identified as aspartate (**4a**) and its 3-substituted derivatives **4b–4n**, respectively. While these L384A-catalysed amination reactions provided high regioselectivities, varying degrees of diastereoselectivity were obtained, with the single or predominant diastereoisomer having the *threo* configuration (Table 2). Analysis of products **4a** and **4j** by chiral HPLC revealed that the L384A-catalysed amination reactions are highly enantioselective, producing these amino acids with >99% e.e. (Table 2).

The preparative usefulness of the L384A mutant was further demonstrated by transformations using 0.05 mol% biocatalyst, 5 M ammonia and 30 mM of the structurally distinct unsaturated acids **1a**, **1g** or **1j** (167-fold molar excess of ammonia over unsaturated acid). Under these optimized conditions, the reactions were complete within 10–60 min at 25 °C, achieving final conversions of ~85–100% (Fig. 2b). These results clearly demonstrate the potential of the L384A mutant for applications in the selective synthesis of valuable α-amino acids.

Structural basis for the expanded substrate spectrum. To obtain insight into how the active sites of the Q73A and L384A mutants are remodelled to accommodate the new substrates, we determined their X-ray crystal structures at resolutions of 2.0 Å and 1.9 Å, respectively (Supplementary Table S3). A comparison of the Q73A and L384A structures with that of wild-type MAL revealed no significant changes in overall structure nor in main-chain and side-chain conformations of the active-site residues (Fig. 3). These results demonstrate that the expanded substrate spectrum of the Q73A and L384A mutants is not due to major structural changes but is related to the replacement of a large active site residue by a smaller one, resulting in an enlarged active site that can accommodate the new substrates (Fig. 3).

To further demonstrate that the active sites of the Q73A and L384A mutants can accommodate larger substrates than the active site of wild-type MAL, we performed comparative docking simulations. Figure 4a shows an overlay of the docking model of wild-type MAL in complex with the natural substrate *threo*-**3a** and the docking model of mutant Q73A complexed with the unnatural substrate *threo*-(2*S*,3*S*)-2-(cyclohexylamino)-3-methylbutanedioic acid (**3k**), respectively. This comparison clearly shows that compound **3k** does not fit into the active site of wild-type MAL because residue Q73, instead of the shorter A73, partially fills the volume that is occupied by the cyclohexylamino moiety of substrate **3k** in

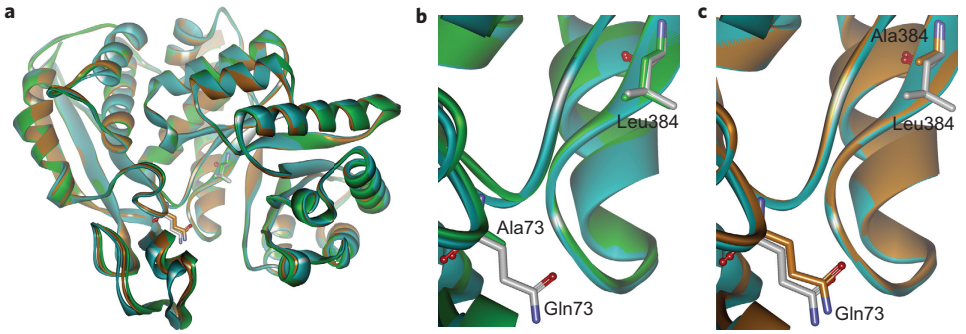


Figure 3 | Structural comparison of wild-type MAL and the Gln73Ala and Leu384Ala mutants. a, Overlay of the structures of wild-type MAL (cyan ribbons) and the Gln73Ala (green ribbons) and Leu384Ala (orange ribbons) mutants. The structures of the Gln73Ala (PDB code 3ZVH) and Leu384Ala (PDB code 3ZVI) mutants can be superimposed on the wild-type MAL structure[12] (PDB code 1KCZ) with a root-mean-square deviation (RMSD) of 0.53–0.74 Å, considering all atoms in the two subunits that occupy the different crystallographic asymmetric units. When considering only the atoms of residues that are within 5 Å of the residues at positions 73 and 384, the RMSD is 0.19–0.50 Å. **b**, Overlay of the active sites of wild-type MAL and the Gln73Ala mutant. Colour scheme as in **a**. **c**, Overlay of the active sites of wild-type MAL and the Leu384Ala mutant. Colour scheme as in **a**. In all three figures, the residues at positions 73 and 384 are shown as sticks. For clarity, these sticks are shown in grey for wild-type MAL. The figures were prepared with Discovery Studio 2.5.

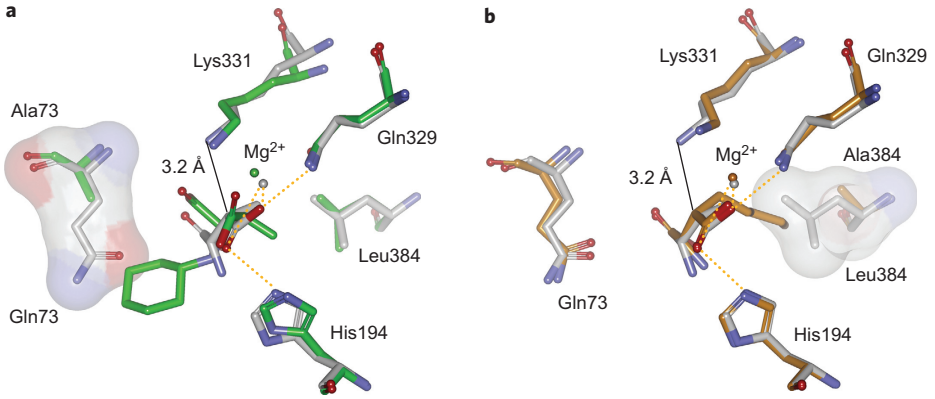


Figure 4 | Molecular docking of substrates in the active sites of wild-type MAL and the two MAL mutants. a, Overlay of the most favourable docked conformations of *threo*-**3a** and **3k** in the active sites of wild-type MAL (grey) and the Gln73Ala mutant (green), respectively. Residue Gln73 in wild-type MAL is represented by a solvent-accessible surface calculated with a probe radius of 1.4 Å. **b**, Overlay of the most favourable docked conformations of *threo*-**3a** and **4f** in the active sites of wild-type MAL (grey) and the Leu384Ala mutant (orange), respectively. Residue Leu384 in wild-type MAL is represented by a solvent-accessible surface calculated with a probe radius of 1.4 Å. In both figures, interactions between active-site residues or the Mg^{2+} ion and substrate *threo*-**3a** in wild-type MAL are displayed as orange dashed lines. The distance between the substrate's C3 atom and the side chain of Lys331 is shown in ångstroms (atoms connected by a solid line). The figures were prepared with Discovery Studio 2.5.

the Q73A–substrate complex. The observation that residue Q73 occludes the amine binding pocket thus explains why cyclohexylamine **2k** and other large amines are not accepted as substrates by wild-type MAL. An overlay of the docking model of wild-type MAL complexed with *threo*-**3a** and that of mutant L384A in complex with the unnatural substrate *threo*-(2*S*,3*S*)-3-pentylaspartic acid (**4f**) is shown in Fig. 4b. This comparison shows that the mutation of residue L384 to the shorter A384 prevents a clash between the substrate's pentyl group and the large leucine residue found at position 384 in wild-type MAL, thus rationalizing the ability of the L384A mutant to convert large electrophiles.

Discussion

Wild-type MAL accepts only a few amines as alternative nucleophiles for ammonia and a few fumarate derivatives as alternative electrophiles for mesaconate[17–19]. Therefore, the goal of this study was to expand both the nucleophile and electrophile scope of MAL. This goal was achieved by using an engineering strategy that involves randomization at carefully chosen sites lining the substrate binding pocket of the enzyme. The major benefit of this structure-guided engineering strategy lies in the reduction of the sequence space that has to be covered to identify valuable mutants[20–24].

Engineering of the amine binding pocket (Fig. 1a) yielded a mutant (Q73A) that has a very broad nucleophile scope and excellent regio- and stereoselectivity, producing the *threo* isomers of the *N*-substituted 3-methylaspartic acids **3b–3u** with very high diastereomeric and enantiomeric excess (Table 1). These results demonstrate that the high regio- and stereoselectivity of the Q73A-catalysed addition reactions is not influenced by the different substituents on the amine substrate. This is of high synthetic significance, especially when compared to the Q73A-catalysed ammonia addition to mesaconate (the natural reaction), which shows low diastereoselectivity (Table 1, entry 1). Recently, a mechanism was proposed for MAL[8,12,13] in which the proton abstraction/addition steps are effected by K331 and H194 as the (*S*)- and (*R*)-specific general base/acid catalysts, respectively, which are juxtaposed on either side of the chiral C3 carbon of the amino-acid substrate. In view of this mechanism, the high diastereoselectivity of the Q73A-catalysed amine addition reactions could be explained by assuming that the

substituents on the amine substrate sterically hinder protonation of the C3 carbon (numbering according to the amino-acid product) of mesaconate by H194, thereby preventing the formation of *erythro* isomers. Structural analysis of the Q73A mutant showed that this mutant enzyme has an enlarged amine binding pocket, without changes in the orientation of active-site residues, thus rationalizing its ability to convert the new amine substrates.

Engineering of the alkyl binding pocket (Fig. 1b) yielded a mutant (L384A) that has a very broad electrophile scope and excellent regio- and enantioselectivity in the amination reaction, producing (*S*)-aspartic acid (**4a**) and its 3-substituted derivatives **4b–4n** (Table 2). The structure of L384A showed that this mutant enzyme has an enlarged alkyl binding pocket, with an otherwise unchanged active site geometry, thus rationalizing its ability to convert fumarate derivatives with large substituents. In contrast to the Q73A-catalysed amine additions, the L384A-catalysed amination of the fumarate derivatives gives varying degrees of diastereoselectivity. This irregular pattern of diastereoselectivities is an interesting phenomenon and suggests that some of the fumarate derivatives, like the natural substrate mesaconate, are optimally positioned to undergo protonation by both K331 and H194 from either face of the substrate, producing the undesired mixture of *threo* and *erythro* isomers, whereas others are bound in a position that does not favour protonation by H194, yielding only the desired *threo* isomers. Nonetheless, the synthetic potential of the L384A mutant is convincingly demonstrated by its use as catalyst in the stereoselective synthesis (diastereomeric excess (d.e.) values are >95%) of *threo*-3-benzyloxyaspartate (**4j**) and *threo*-3-benzylaspartate (**4n**), which are important compounds in studying the role of glutamate transporters in the brain[26–28].

In summary, we provide evidence that a single amino acid substitution is sufficient to dramatically expand the nucleophile or electrophile spectrum of MAL. The results provide support for the notion that structure-guided saturation mutagenesis of sites lining the substrate binding pocket is a powerful strategy to create enzymes that can accommodate non-native substrates and perform reactions not observed in nature[22]. The new designer enzymes catalyse attractive and straightforward asymmetric transformations in water, making use of readily available starting substrates, without the need for organic (co)solvents or the implementation of dynamic kinetic resolution strategies. These are very desirable features in view of the quest for environmentally benign and atom efficient processes. The potential of the two engineered MALs for application in the selective synthesis of new and valuable aspartic acid derivatives has been demonstrated, given that the product yields may be further improved by optimizing the reaction conditions and purification protocols. The Q73 and L384 residues represent good targets for future combinatorial mutagenesis experiments (Supplementary Results and Discussion), potentially yielding new biocatalysts with an even larger substrate scope.

Methods

Mutant library preparation and screening. Site-saturation mutagenesis libraries were generated by overlap extension polymerase chain reaction[33] using plasmid pBAD(MAL-His)[8] as the template and the primers listed in Supplementary Table S4. Detailed procedures for library preparation and screening are provided in the Supplementary Methods.

Enzyme expression, purification and kinetic characterization. Wild-type MAL and MAL mutants were overproduced in *E. coli* TOP10 cells and purified to homogeneity as described before[8]. Procedures for the kinetic characterization of MAL (wild-type or mutants) are provided in the Supplementary Methods.

General procedure for monitoring the addition of various amines to mesaconate. The ability of wild-type MAL and the Q73A mutant to add different amines (**2a–2u**) to mesaconate (**1b**) was examined using ^{1}H NMR spectroscopy. Reaction mixtures consisted of 500 μl of a 1 M stock solution of amine in water (pH 9.0, containing 20 mM $MgCl_2$), 100 μl of 500 mM **1b** in water (pH 9.0, containing 20 mM $MgCl_2$) and 50 μl of D_2O. In the case of amines **2f**, **2g** and **2m**, 500 μl of a 250 (**2f**) or 125 mM (**2g** and **2m**) stock solution was added to the reaction mixtures to avoid protein precipitation. Reactions were started by the addition of 200 μg of freshly purified enzyme, and reaction mixtures were incubated at 22 °C. ^{1}H NMR spectra were recorded 3 h, 3 days and 7 days after the addition of enzyme. Product amounts were estimated by integration of the respective product signals, if present.

General procedure for monitoring the amination of various electrophiles. The ability of wild-type MAL and the L384A mutant to add ammonia to different electrophiles (**1a–1n**) was examined using ^{1}H NMR spectroscopy. Reaction mixtures consisted of 500 μl of a 1 M stock solution of NH_4Cl in water (pH 9.0, containing 20 mM $MgCl_2$), 50 μl D_2O, and 200 μl of a 250 mM stock solution (or 100 μl of a 500 mM stock solution) of electrophile in water (pH 9.0, containing 20 mM $MgCl_2$). Reactions were started by the addition of 200 μg of freshly purified enzyme, and reaction mixtures were incubated at 22 °C. ^{1}H NMR spectra were recorded 3 h, 3 days and 7 days after addition of enzyme. Product amounts were estimated by integration of the respective product signals, if present.

Starting substrates. Amines **2a–2u** and unsaturated acids **1a** and **1b** were purchased from Sigma-Aldrich or Merck. Unsaturated acids **1c–1n** were prepared according to experimental procedures described in the Supplementary Methods.

Identification of the amino-acid products. Procedures for the enzymatic synthesis and purification of amino acids **3n–3u** and **4a–4n** are described in the Supplementary Methods. The relative configuration of products **4b–4d**, **4h**, **4j** and **4n** was determined by comparison of their ^{1}H NMR signals to those of commercially available or wild-type MAL-synthesized[17] authentic standards with known configuration, and/or by referring to the literature[29–31]. Compound *threo*-DL-**4j** was purchased from Tocris Bioscience. The procedures to establish the relative configuration of products **4k–4m** are provided in the Supplementary Methods. The e.e. of products **4a** and **4j** was determined by chiral HPLC using a Chirex 3126-(D)-penicillamine column (250 mm × 4.6 mm, Phenomenex) or a Nucleosil chiral-1 column (250 mm × 4 mm, Macherey-Nagel), respectively. Authentic standards L-**4a** and D-**4a** were purchased from Sigma-Aldrich. HPLC conditions and observed retention times are given in the Supplementary Methods.

Identity and configuration of amino-acid products 3a–3m. The identity and relative configuration of products **3a–3m** were determined by comparative ^{1}H NMR analysis of the crude reaction mixtures and authentic compounds with known relative configuration. The experimental procedures for the synthesis of these authentic standards are provided in the Supplementary Methods. Products **3c–3e**, **3h**, **3i** and **3l** were also purified from preparative scale (0.15 g **1b**) reactions and their e.e. was determined by chiral HPLC. HPLC conditions, procedures for the synthesis of the *threo*-DL standards, and the observed retention times are provided in the Supplementary Methods.

Optimization of reaction conditions for amino acid synthesis. By increasing both the amount of biocatalyst and the molar excess of amine over mesaconate, excellent conversions were achieved for several Q73A-catalysed amine additions within short reaction times. The optimized reaction conditions were as follows. Reaction mixtures (30 ml) consisted of 3 M amine (either **2b** or **2o**) and 26 mM **1b** in 20 mM $MgCl_2$ (final pH 9). In the case of amine **2i**, a lower concentration (2 M) was used to avoid protein precipitation. The reactions were started by the addition of freshly purified enzyme (0.05 mol%), and reaction mixtures were incubated at 25 °C. Aliquots of each reaction mixture (5 ml) were withdrawn 5 min, 15 min, 45 min, 90 min, 180 min and 24 h after the addition of enzyme. For each sample, the reaction was stopped by incubating the reaction mixture at 100 °C for 5 min.

Optimized conditions for several L348A-catalysed amination reactions were as follows. Reaction mixtures (6–12 ml) consisted of 5 M NH_4Cl and 30 mM **1a**, **1g** or **1j** in 20 mM $MgCl_2$ (final pH 9). The reactions were started by the addition of freshly purified enzyme (0.05 mol%), and reaction mixtures were incubated at 25 °C. Aliquots (1–2 ml) of the reaction mixtures were withdrawn 5 min, 15 min, 45 min, 90 min, 180 min and 24 h after the addition of enzyme. For each sample, the reaction was stopped by incubating the reaction mixture at 100 °C for 5 min.

Each sample was dried *in vacuo*, followed by resuspending the resulting residue in ~10 ml of water for lyophilization. This process was repeated 5 to 10 times to remove the excess of amine/ammonia to enable analysis by ^{1}H NMR spectroscopy. The final residue was dissolved in 800 μl D_2O. The conversion of substrate to product was estimated by integration of the respective substrate and product signals.

Crystallization, structure determination and molecular docking. The procedures used for crystallization, structure determination and comparative docking simulations are described in the Supplementary Methods.

Received 16 January 2012; accepted 21 March 2012; published online 29 April 2012

References

1. Sonke, T., Kaptein, B. & Schoemaker, H. E. Use of enzymes in the synthesis of amino acids, in *Amino Acids, Peptides and Proteins in Organic Chemistry* Vol. 1 (ed. Hughes, A. B.) 77–117 (Wiley-VCH, 2009).

2. Wohlgemuth, R. Biocatalysis—key to sustainable industrial chemistry. *Curr. Opin. Biotechnol.* **21,** 713–724 (2010).
3. Panke, S. & Wubbolts, M. Advances in biocatalytic synthesis of pharmaceutical intermediates. *Curr. Opin. Chem. Biol.* **9,** 188–194 (2005).
4. Panke, S., Held, M. & Wubbolts, M. Trends and innovations in industrial biocatalysis for the production of fine chemicals. *Curr. Opin. Biotechnol.* **15,** 272–279 (2004).
5. Turner, N. J. Ammonia lyases and aminomutases as biocatalysts for the synthesis of α-amino and β-amino acids. *Curr. Opin. Chem. Biol.* **15,** 234–240 (2011).
6. Barker, H. A., Smyth, R. D., Wilson, R. M. & Weissbach, H. The purification and properties of β-methylaspartase. *J. Biol. Chem.* **234,** 320–328 (1959).
7. Goda, S. K., Minton, N. P., Botting, N. P. & Gani, D. Cloning, sequencing, and expression in *Escherichia coli* of the *Clostridium tetanomorphum* gene encoding β-methylaspartase and characterization of the recombinant protein. *Biochemistry* **31,** 10747–10756 (1992).
8. Raj, H. *et al.* Alteration of the diastereoselectivity of 3-methylaspartate ammonia lyase by using structure-based mutagenesis. *ChemBioChem* **10,** 2236–2245 (2009).
9. Bright, H. J. Divalent metal activation of β-methylaspartase. The importance of ionic radius. *Biochemistry* **6,** 1191–1203 (1967).
10. Bright, H. J. On the mechanism of divalent metal activation of β-methylaspartase. *J. Biol. Chem.* **240,** 1198–1210 (1965).
11. Kato, Y. & Asano, Y. 3-Methylaspartate ammonia-lyase as a marker enzyme of the mesaconate pathway for (*S*)-glutamate fermentation in Enterobacteriaceae. *Arch. Microbiol.* **168,** 457–463 (1997).
12. Asuncion, M., Blankenfeldt, W., Barlow, J. N., Gani, D. & Naismith, J. H. The structure of 3-methylaspartase from *Clostridium tetanomorphum* functions via the common enolase chemical step. *J. Biol. Chem.* **277,** 8306–8311 (2002).
13. Levy, C. W. *et al.* Insights into enzyme evolution revealed by the structure of methylaspartate ammonia lyase. *Structure* **10,** 105–113 (2002).
14. Bright, H. J., Ingraham, L. L. & Lundin, R. E. The mechanism of the methylaspartate ammonia–lyase reaction: deuterium exchange. *Biochim. Biophys. Acta* **81,** 576–584 (1964).
15. Bright, H. The mechanism of the β-methylaspartase reaction. *J. Biol. Chem.* **239,** 2307–2315 (1964).
16. Babbitt, P. C. *et al.* The enolase superfamily: a general strategy for enzyme-catalyzed abstraction of the α-protons of carboxylic acids. *Biochemistry* **35,** 16489–16501 (1996).
17. Akhtar, M., Botting, N. P., Cohen, M. A. & Gani, D. Enantiospecific synthesis of 3-substituted aspartic acids via enzymic amination of substituted fumaric acids. *Tetrahedron* **43,** 5899–5908 (1987).
18. Botting, N. P., Akhtar, M., Cohen, M. A. & Gani, D. Substrate specificity of the 3-methylaspartate ammonia-lyase reaction: observation of differential relative reaction rates for substrate-product pairs. *Biochemistry* **27,** 2953–2955 (1988).
19. Gulzar, M. S., Akhtar, M. & Gani, D. Preparation of *N*-substituted aspartic acids via enantiospecific conjugate addition of *N*-nucleophiles to fumaric acids using methylaspartase: synthetic utility and mechanistic implications. *J. Chem. Soc. Perkin Trans.* **1,** 649–656 (1997).
20. Lutz, S. & Bornscheuer, U. T. *Protein Engineering Handbook*. Vols 1–2 (Wiley-VCH, 2009).
21. Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. & Vogel, A. Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. *Angew. Chem. Int. Ed.* **44,** 4192–4196 (2005).
22. Morley, K. L. & Kazlauskas, R. J. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **23,** 231–237 (2005).
23. Arnold, F. H. & Georgiou, G. (eds) *Methods in Molecular Biology (Directed Enzyme Evolution)* Vol. 230 (Humana Press, 2003).
24. Turner, N. J. Directed evolution drives the next generation of biocatalysts. *Nature Chem. Biol.* **5,** 567–573 (2009).
25. Shimamoto, K. *et al.* Characterization of novel L-*threo*-β-benzyloxyaspartate derivatives, potent blockers of the glutamate transporters. *Mol. Pharmacol.* **65,** 1008–1015 (2004).
26. Shimamoto, K. Glutamate transporter blockers for elucidation of the function of excitatory neurotransmission systems. *Chem. Rec.* **8,** 182–199 (2008).
27. Esslinger, C. S. *et al.* The substituted aspartate analogue L-β-*threo*-benzyl-aspartate preferentially inhibits the neuronal excitatory amino acid transporter EAAT3. *Neuropharmacology* **49,** 850–861 (2005).
28. Bridges, R. J. & Esslinger, C. S. The excitatory amino acid transporters: pharmacological insights on substrate and inhibitor specificity of the EAAT subtypes. *Pharmacol. Ther.* **107,** 271–285 (2005).
29. Mavencamp, T. L., Rhoderick, J. F., Bridges, R. J. & Esslinger, C. S. Synthesis and preliminary pharmacological evaluation of novel derivatives of L-β-*threo*-benzylaspartate as inhibitors of the neuronal glutamate transporter EAAT3. *Bioorg. Med. Chem.* **16,** 7740–7748 (2008).
30. Shimamoto, K. *et al.* Syntheses of optically pure β-hydroxyaspartate derivatives as glutamate transporter blockers. *Bioorg. Med. Chem. Lett.* **10,** 2407–2410 (2000).
31. Spengler, J., Pelay, M., Tulla-Puche, J. & Albericio, F. Synthesis of orthogonally protected L-*threo*-β-ethoxyasparagine. *Amino Acids* **39,** 161–165 (2010).
32. LaBean, T. H. & Kauffman, S. A. Design of synthetic gene libraries encoding random sequence proteins with desired ensemble characteristics. *Protein Sci.* **2,** 1249–1254 (1993).
33. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77,** 51–59 (1989).

Acknowledgements

The authors thank B. Wu, B. Kaptein and O. May for helpful discussions. This research was financially supported by VENI grant 700.54.401 and ECHO grant 700.59.042 (both to G.J.P.) from the Division of Chemical Sciences of the Netherlands Organisation of Scientific Research (NWO-CW), and by the Netherlands Ministry of Economic Affairs and the B-Basic partner organizations (www.b-basic.nl) through B-Basic, a public-private NWO–ACTS programme (ACTS, Advanced Chemical Technologies for Sustainability). Financial support from the Royal Netherlands Academy of Arts and Sciences (KNAW to W.S. and B.L.F.) is also gratefully acknowledged.

Author contributions

H.R. performed mutagenesis, library screening and activity measurements. W.S. and J.d.V. synthesized starting substrates and reference compounds. H.R., W.S., M.d.V. and F.J.D. performed preparative biocatalysis. H.J.R. performed X-ray crystallography experiments. V.P.V. performed chiral HPLC. C.R.R. performed the molecular docking experiments. S.d.W., W.J.Q., A.M.W.H.T., B.L.F., D.B.J. and G.J.P. supervised scientific work. All authors contributed to writing the paper.

Additional information

The authors declare no competing financial interests. Supplementary information and chemical compound information accompany this paper at www.nature.com/naturechemistry. Reprints and permission information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.J.P.