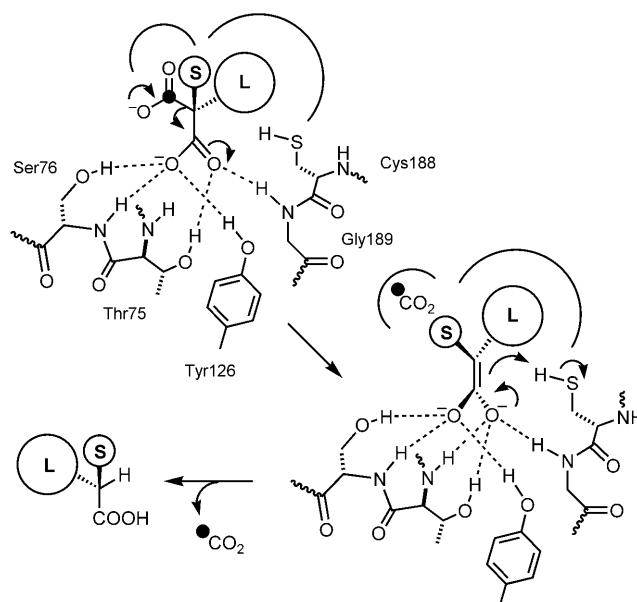


Structure-Guided Directed Evolution of Alkenyl and Arylmalonate Decarboxylases

Krzysztof Okrasa, Colin Levy, Matthew Wilding, Mark Goodall, Nina Baudendistel, Bernhard Hauer, David Leys,* and Jason Micklefield*

The asymmetric decarboxylation of prochiral α -disubstituted malonic acids is an attractive route to produce homochiral α -substituted carboxylic acids. Various metal-catalysed and organocatalytic methods have been explored for related enantioselective decarboxylative protonation (EDP) reactions.^[1] However, the synthetic EDP methods can only deliver chiral carboxylic acids or esters with moderate enantioselectivities (*ee*).^[1] On the other hand, enzymatic decarboxylation offers a cleaner, more sustainable and more efficient approach to deliver carboxylic acids of very high *ee*. Notably, the arylmalonate decarboxylase (AMDase) isolated from *Bordetella bronchiseptica*^[2] has been shown to be highly effective in the decarboxylation of α -arylmalonates to give enantiomerically pure α -arylcarboxylic acids (Scheme 1).^[3,4] The fact that the AMDase is highly robust and does not require co-factors further increases the potential of this enzyme for biocatalysis. Despite this, the pool of substrates accepted by the enzyme is limited to malonates that possess α -aryl substituents. Indeed a number of other bacterial AMDases have been discovered from sequence similarity searches,^[5] or selective enrichment experiments.^[6] While all these enzymes could catalyse the decarboxylation of α -arylmalonates, none have so far been demonstrated to decarboxylate malonates which do not possess α -aryl substituents.

In order to rationalise this, and guide the development of new enzymes which could catalyse a wider range of reactions, we solved the X-ray crystal structure of the *B. bronchiseptica* AMDase.^[5] Our first high-resolution structure^[7] reveals a “dioxanion hole” motif, which can donate six hydrogen bonds to potentially stabilize a putative high-energy enediolate intermediate, along with a small hydrophobic cavity likely favouring the formation of the neutral CO_2 . From this we were able to suggest a mechanism for the AMDase (Scheme 1). From comparison of the AMDase structure^[5]



Scheme 1. Proposed mechanism of the AMDase.^[5] Preferred malonate substrates for the AMDase possess a larger substituent (L = aryl or alkenyl), which binds to a solvent-accessible cavity, and a smaller substituent (S = H, Me, OH, NH_2). The neutral leaving group CO_2 develops within the small hydrophobic pocket and the resulting enediolate is stabilised by the dioxanion hole. In the case of the model substrate 2-methyl-2-phenylmalonate (L = Ph, S = Me) decarboxylation has been shown to occur through loss of the pro-R carboxylate and subsequent protonation of the *si*-face by Cys 188, resulting in overall inversion of configuration.^[5]

with structures of Asp/Glu racemases,^[8] we proposed that conserved dioxanion hole motifs could have evolved to stabilise common enediolate intermediates in these mechanistically related enzymes. Indeed, up until our structure of the AMDase, there was no evidence to suggest how the AMDase or Asp/Glu racemases stabilised common putative high-energy enediolate intermediates.^[5] However, our proposal was based on a structure which only possessed phosphate as an active-site ligand.^[5] To lend further support to our proposal, we here present a second *B. bronchiseptica* AMDase structure with a mechanism-based inhibitor that more closely resembles the enediolate intermediate bound to the active site of the enzyme. This allowed us to rationalise the mechanism of this simple yet intriguing enzyme and design new and effective malonate substrates for the AMDase. Also, the new structure guided directed evolution of new AMDase mutants with significantly enhanced catalytic efficiency with a range of substrates.

[*] Dr. K. Okrasa, M. Wilding, M. Goodall, Prof. Dr. J. Micklefield
School of Chemistry, The University of Manchester
Manchester Interdisciplinary Biocentre
131 Princess Street, Manchester M1 7ND (UK)
E-mail: jason.micklefield@manchester.ac.uk

Dr. C. Levy, Dr. D. Leys
Faculty of Life Science, The University of Manchester
Manchester Interdisciplinary Biocentre
131 Princess Street, Manchester M1 7ND (UK)
E-mail: david.leys@manchester.ac.uk

Dr. N. Baudendistel, Prof. Dr. B. Hauer
BASF SE, GVF/E, 67056 Ludwigshafen (Germany)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/ange.200904112>.

The natural substrate of the *B. bronchiseptica* AMDase is not known, however phenylmalonic acid (**1a**) is the best substrate that has been described to date. Based on this, and the fact that phosphate was previously shown to bind to the dioxanion hole,^[5] benzylphosphonate (**2**) was selected as a potential mechanism-based inhibitor. Benzylphosphonate (**2**) is both sterically and electronically similar to the proposed enediolate intermediate **3** resulting from decarboxylation of phenylmalonate (Figure 1). In support of this, kinetic studies

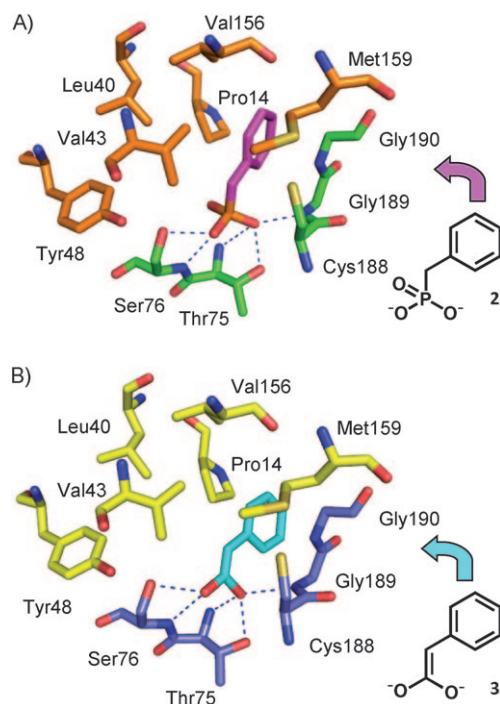


Figure 1. A) Active-site structure of the *B. bronchiseptica* AMDase with the benzylphosphonate inhibitor bound to the active site. This was derived by soaking crystals of the enzyme with benzylphosphonate. B) Model of the AMDase as derived from the original phosphate-bound structure^[5] by docking a putative enediolate intermediate **3** into the active site using the H-bonds of the dioxanion hole to establish orientation. Residues of the dioxanion hole are coloured in green in structure A and blue in structure B. The Tyr126, which forms a sixth H-bond of the dioxanion hole has been removed from these structures for clarity.

revealed benzylphosphonate to be a competitive inhibitor of the AMDase ($K_i = 5.2$ mM). Consequently, co-crystallisation of the AMDase with the benzylphosphonate was undertaken. The corresponding crystal structure shows the phosphonate dianion **2** bound to the active site by the same six hydrogen bonds that make up the dioxanion hole (Figure 1 A). Moreover, the phenyl group of the phosphonate side chain is positioned in the larger solvent-accessible cavity, with the phenyl ring stacking, through van der Waals interactions, between the Gly189–Gly190 amide bond and the edge of the Pro14 residue. This is similar to a model that was generated, based on the first X-ray crystal structure of the AMDase (Figure 1 B and Figure S1 in the Supporting Information), with the enediolate intermediate **3** docked to the dioxanion hole.^[5] This structure thus provides additional evidence that the AMDase catalyses decarboxylation of arylmalonates

through formation of an enediolate intermediate, which is entirely consistent with the previous proposal.^[5]

Asymmetric decarboxylation of aliphatic α -disubstituted malonates would be highly desirable. To rationalise why these substrates are not accepted by the enzyme, a model of 2-methyl-2-propylmalonate bound to the AMDase active site was generated (Figure S3). This substrate fits well within the active site, suggesting that electronic rather than steric effects must account for the substrate selectivity in this case. Indeed, the AMDase structures (Figure 1) indicate that the stacking interaction between Pro14 and Gly189–Gly190 amide bond necessitate that the phenyl group must be coplanar with the π -system of the enediolate. Therefore, in addition to the H-bonding with the dioxanion hole, delocalisation of electron density into the phenyl group is likely to further help stabilise the enediolate. To explore this, malonate **5c** with a *trans*-propenyl substituent was prepared by modification of literature procedures (Scheme S1).^[9] Notably, this substrate **5c** is decarboxylated by the *B. bronchiseptica* AMDase to give the *R*-configured acid **6c**, with 99% *ee* (Table 1). Kinetic parameters were determined for this transformation, using the method described previously.^[5,10] The results revealed that the catalytic efficiency for conversion of the *trans*-propenyl derivative **5c** is lower than for the 2-methyl-2-phenylmalonate (**5a**) (Table 1).

In light of this promising result, a range of substrates possessing alkenyl substituents were prepared, through modification of literature procedures (Scheme S1).^[9,11] Remarkably, all of these compounds proved to be substrates for the AMDases giving decarboxylation products with very high enantioselectivities, albeit with slightly lower catalytic efficiencies compared with the corresponding phenylmalonates.

Table 1: Kinetic parameters for AMDase catalysed decarboxylations.

$ \begin{array}{l} \text{1 S = H} \\ \text{5 S = CH}_3 \\ \text{7 S = OH} \\ \text{9 S = NH}_2 \end{array} \begin{array}{c} \text{HO}_2\text{C} \\ \\ \text{C} \begin{array}{l} \text{S} \\ \text{L} \end{array} \\ \\ \text{CO}_2\text{H} \end{array} \longrightarrow \begin{array}{l} \text{4 S = H} \\ \text{6 S = CH}_3 \\ \text{8 S = OH} \\ \text{10 S = NH}_2 \end{array} \begin{array}{c} \text{L} \\ \\ \text{C} \\ \\ \text{CO}_2\text{H} \end{array} $						
$ \text{L} = \begin{array}{c} \text{a} \end{array} \begin{array}{c} \text{b} \end{array} \begin{array}{c} \text{c} \end{array} \begin{array}{c} \text{d} \end{array} \begin{array}{c} \text{e} \end{array} $						
Subst.	$K_m^{[a]}$ [mM]	$k_{cat}^{[a]}$ [s ⁻¹]	$k_{cat}/K_m^{[a]}$ [s ⁻¹ mM ⁻¹]	Product ^[c] (config.)	<i>ee</i> ^[b] [%]	
1a	10.7 ± 0.8 ^[a]	316 ± 36 ^[a]	29.6 ± 6.0 ^[a]	4a (–)	–	
5a	26.9 ± 2.1	279 ± 32	10.4 ± 2.1	6a (R)	99	
5b	7.8 ± 0.6	42.8 ± 4.9	5.5 ± 1.1	6b (R)	99	
5c	12.8 ± 1.0	46.2 ± 5.3	3.6 ± 0.8	6c (R)	99	
5d	14.4 ± 1.1	23.9 ± 2.7	1.7 ± 0.3	6d (R)	99	
5e	3.5 ± 0.3	5.2 ± 0.6	1.5 ± 0.3	6e (R)	99	
7a	30.2 ± 2.4	101 ± 12	3.3 ± 0.7	8a (R)	96	
7b	9.1 ± 0.7	47.6 ± 5.5	5.2 ± 1.1	8b (R)	98	
7e	7.3 ± 0.6	20.7 ± 2.4	2.8 ± 0.6	8e (R)	76	
9a	n.d. ^[d]	n.d. ^[d]	n.d. ^[d]	10a (R)	33 ^[d]	
9b	n.d. ^[d]	n.d. ^[d]	n.d. ^[d]	10b (R)	61 ^[d]	

[a] Kinetic parameters for *B. bronchiseptica* AMDase.^[5] [b] Product *ee* determined by chiral-phase GC. [c] Configuration was confirmed by comparison with available homochiral standards or optical rotations from the literature. [d] The absence of kinetic parameters and the low *ee* observed are due to competing non-enzymatic decarboxylation.

nate derivatives. Interestingly, the alkenyl substrates with α -methyl (**5b–e**) or α -hydroxy groups (**7b, 7e**) possess lower K_m values than the corresponding phenyl derivatives (**5a, 7a**). This suggests alkenyl substrates, particularly those with more substituted double bonds (**5e, 7e**), have higher affinity for the AMDase. On the other hand, the turnover number is lower for the alkenyl substrates compared with corresponding phenyl derivatives. Given that the decarboxylation step to generate the enediolate is most likely rate-limiting, this could be explained by the greater delocalisation that is afforded by phenyl compared with alkenyl substituents. Indeed the lowest turnover numbers are observed for the substrates with more electron-donating alkyl-substituted double bonds (e.g. **5e, 7e**), which would be predicted to result in greater destabilisation of the transition state leading to the enediolate. Similar results were also obtained with the AMDase from *Mesorhizobium* sp.^[5] (Table S1).

Previously the AMDase-catalysed decarboxylation of 2-amino-2-phenylmalonate was reported to give (*R*)-phenylglycine in 96% *ee*.^[12] In our hands, this proved difficult to reproduce as the background non-enzymatic decarboxylation was found to compete with the enzyme-catalysed reaction, which is in line with literature observations.^[13] Despite this, 2-amino-2-vinylmalonate (**9b**) was prepared (Scheme S1) and incubated with the AMDase. Whilst significant background decarboxylation was again observed, it was possible to obtain (*R*)-vinylglycine (**10b**) in 61% *ee*. It is likely that the *ee* could be improved using a milder and more efficient synthesis of **9b** that does not involve aggressive base-catalysed hydrolysis in the final step. It may also be possible to increase the rate of the enzymatic decarboxylation, through the directed evolution of the AMDases.

Based on the AMDase structure, five key active-site residues were identified, and divided into two groups, for saturation mutagenesis. The first group includes residues from the large solvent-exposed aryl/alkenyl-binding pocket (Pro14, Pro15 and Gly190) and the second group from the smaller, more hydrophobic cavity (Val43 and Met159) (Figure S4A). The two groups of active-site residues were separately mutated, using the NNK degenerate codon, following a strategy that is similar to iterative saturation mutagenesis (ISM).^[14] Following this strategy, mutation of the aryl binding pocket group (Pro14, Pro15 and Gly190) led in all cases to a first generation of single mutants of each residue (Figure S4B). Those single mutants, which proved to be active decarboxylases, were then subjected to a second round of mutagenesis. Similarly, the second generation of active double mutants were mutated to generate a library of triple mutants. Phenylmalonic acid was used for initial screening, because turnover of this substrate by AMDases is considerably faster than any of the disubstituted malonates. Using this substrate, mutant AMDases can be screened rapidly in 96-well plates, using the pH indicator bromothymol blue (BTB). This assay allows mutants to be selected visually through colour change, and also accurate kinetic parameters can be determined as described previously.^[5]

In the case of the small binding pocket, we were able to generate single mutants of Met159, but no mutants of Val43 could be isolated. Single mutants of Met159 were subjected to

further mutagenesis of Val43, but no active double mutants were isolated. Despite this, we were able to generate a library of single, double and triple mutants with significantly improved efficiency for phenylmalonate decarboxylation (Table 2). In light of this, successful mutants were further screened with the pool of aryl and alkenylmalonates that had been synthesised. Notably, the M159V mutant shows a 51-fold improvement in activity with phenylmalonate **1a**. Such a large

Table 2: Kinetic parameters of the selected mutant AMDases.

Subst.	Mutants ^[b]	K_m [mM]	k_{cat} [s ⁻¹]	k_{cat}/K_m [s ⁻¹ mM ⁻¹]	Rel. act. ^[a]
1a	M159V	0.3 ± 0.02	450 ± 52	1499 ± 290	51
	M159G	0.43 ± 0.03	246 ± 28	570 ± 115	19
	P14V + P15G	3.5 ± 0.3	1143 ± 131	326 ± 62	11
	M159S	5.0 ± 0.4	386 ± 44	77.2 ± 16	2.6
	M159C	12.1 ± 1.0	489 ± 56	40.4 ± 8.7	1.4
5a	M159C	3.2 ± 0.3	47.8 ± 5.5	14.8 ± 3.6	1.4
5b	G190A	0.8 ± 0.1	20.8 ± 2.4	25.2 ± 4.9	4.6
5e	M159V	1.6 ± 0.1	4.6 ± 0.5	2.9 ± 0.5	2.1
7a	G190S + P14V + P15G	14.3 ± 1.1	94.4 ± 11	6.6 ± 1.4	1.9
	P14V + P15G	15.3 ± 1.2	99.8 ± 11	6.5 ± 1.4	1.9
	M159C	4.1 ± 0.3	23.4 ± 2.7	5.8 ± 1.1	1.7
7e	P14V + P15G	8.0 ± 0.6	34.4 ± 4.0	4.3 ± 0.9	1.5

[a] Relative activity was calculated using: Rel. act. = (k_{cat}/K_m of mutant) / (k_{cat}/K_m of wild-type). [b] *B. bronchiseptica* AMDase mutants were characterised by DNA sequencing and ESI-MS.

increase in activity, from a single-point mutant, suggests that phenylmalonate is not the natural substrate for the *B. bronchiseptica* AMDase. Indeed, nature has presumably sampled this mutation as the *Mesorhizobium* sp. AMDase possesses V rather than M at the position corresponding to 159, the only difference between the active sites of both enzymes. The M159V mutant also shows a 2-fold increase in catalytic efficiency with the new alkenyl malonate **5e**. Modelling of M159V and also M159G mutations suggests these appear to generate more space around the catalytic C188 residue, which can result in elevated flexibility of this residue and closer interaction with bound substrate and/or enediolate. The double mutant P14V + P15G proved to be one of the most versatile variants, since it has enhanced activity towards three different substrates, phenylmalonate (**1a**) (11-fold), 2-hydroxy-2-phenylmalonic acid (**7a**) (1.9-fold) and the corresponding alkenyl derivative **7e** (1.5-fold). This double mutation provides more space and flexibility in the large aryl-binding pocket, which can better accommodate the larger alkenyl side chain of **7e** as well as the phenyl groups of **1a** and **7a**. Another notable mutant, G190A, showed a 4.6-fold increase in activity towards the smallest aliphatic substrate, 2-vinyl-2-methylmalonic acid (**5b**). Presumably the G190A mutation decreases the size of the aryl-binding pocket consequently providing tighter binding to the smaller substrate **5b**. Indeed, the K_m of the G190A mutant for this substrate is 0.8 mM, compared with 7.8 mM observed for the wild-type enzyme. Finally, analysis of the chiral products obtained from the decarboxylation with the mutant enzymes showed that there was no change in *ee* compared with the reaction catalysed by the wild-type AMDase.

In summary, we have obtained a new crystal structure of the *B. bronchiseptica* AMDase with benzylphosphonate bound in the active site, suggesting that an enediolate is the most probable intermediate formed during decarboxylation. Although the presence of a dioxyanion hole stabilises the enediolate, this appears insufficient for decarboxylation to ensue, and viable substrates also need either an aryl or alkenyl group in conjugation with the enediolate π -system to further stabilise the developing negative charge concomitant with decarboxylation. Other substituents possessing a π -system, such as alkynyl groups, would presumably facilitate enzymatic decarboxylation. These findings allowed us to design a library of novel α -alkenyl malonate substrates which were decarboxylated with very high *ee* by the AMDase. Many of the corresponding products are difficult to prepare, in such high *ee*, using non-enzymatic asymmetric synthesis methods. Furthermore, the double bond in these products is an ideal handle for further synthetic modification, with the newly installed α -stereogenic centre providing a means of controlling the configuration of additional stereocentres.^[15–17] The α -aryl and α -alkenylmalonate AMDase substrates are also very cheap and readily available from natural sources. In addition to malonic acid, many of the substrates can be derived from ketomalonate (Scheme S1) which is available from oxidation of glycerol, a by-product of biodiesel production.^[18] Coupled with the key environmentally benign enzymatic decarboxylation, this makes for a more sustainable route to these key chiral precursors. Finally, we used our refined inhibitor-bound structure of the AMDase (PDB code 3IP8)^[19] to guide the directed evolution of new decarboxylases, with up to 50-fold increase in activity. Moreover, mutants were selected which showed improved catalytic activity with a range of α -aryl and α -alkenyl malonates. Indeed, we have demonstrated that AMDase displays considerable plasticity in the substrate binding pocket, and is thus ideal for further rational redesign and directed evolution.

Received: July 24, 2009

Published online: September 8, 2009

Keywords: decarboxylase · directed evolution · enediolates · enzyme mechanisms · inhibitors

- [1] a) J. Blanchet, J. Baudoux, M. Amere, M.-C. Lasne, J. Rouden, *Eur. J. Org. Chem.* **2008**, 5493–5506; b) M. Amere, M.-C. Lasne, J. Rouden, *Org. Lett.* **2007**, 9, 2621–2624; c) J. T. Mohr, T. Nishimata, D. C. Behenna, B. M. Stoltz, *J. Am. Chem. Soc.* **2006**, 128, 11348–11349.
- [2] K. Miyamoto, H. Ohta, *J. Am. Chem. Soc.* **1990**, 112, 4077–4078.
- [3] a) K. Miyamoto, H. Ohta, *Biocatalysis* **1991**, 5, 49–60; b) K. Miyamoto, H. Ohta, *Appl. Microbiol. Biotechnol.* **1992**, 38, 234–238; c) K. Miyamoto, H. Ohta, *Eur. J. Biochem.* **1992**, 210, 475–481; d) K. Matoishi, M. Ueda, K. Miyamoto, H. Ohta, *J. Mol. Catal. B* **2004**, 27, 161–168.
- [4] a) Y. Ijima, K. Matoishi, Y. Terao, N. Doi, H. Yanagawa, H. Ohta, *Chem. Commun.* **2005**, 877–879; b) Y. Terao, Y. Ijima, K. Miyamoto, H. Ohta, *J. Mol. Catal. B* **2007**, 45, 15–20.
- [5] K. Okrasa, C. Levy, B. Hauer, N. Baudendistel, D. Leys, J. Micklefield, *Chem. Eur. J.* **2008**, 14, 6609–6613.
- [6] a) K. Miyamoto, Y. Yatake, K. Tamura, Y. Terao, H. Ohta, *J. Biosci. Bioeng.* **2007**, 104, 263–267; b) Y. Yatake, K. Miyamoto, H. Ohta, *Appl. Microbiol. Biotechnol.* **2008**, 78, 793–799.
- [7] Another *B. bronchiseptica* AMDase structure was also reported at around the same time: E. B. Kuettner, A. Keim, M. Kircher, S. Rosmus, N. Sträter, *J. Mol. Biol.* **2008**, 377, 386–394. This structure is proposed to possess four independent subunits and differs considerably from our structure. Attempts to align the two structures proved problematic. Also, there is no ligand bound to any of the subunits, in the structure from Kuettner et al., and none of the key active-site residues, apart from Cys188 are identified. Moreover, the Cys188 residue had been modified by mercaptoethanol, presumably an artefact of the crystallisation process employed by Kuettner et al. From our work, the conditions of crystallisation were shown to be very important. The use of mercaptoethanol, in addition to modifying Cys188 which would completely abolish the ability of this residue to function as a general acid, may also have effected the internal stabilisation of the protein leading to the four misfolded subunits. As a result, no reasonable structural or mechanistic conclusions can be drawn from this.
- [8] T. Lundqvist, S. L. Fisher, G. Kern, R. H. Folmer, Y. Xue, D. T. Newton, T. A. Keating, R. A. Alm, B. L. de Jonge, *Nature* **2007**, 447, 817–822.
- [9] S. Tsuboi, K. Muranaka, T. Sakai, A. Takeda, *J. Org. Chem.* **1986**, 51, 4944–4946.
- [10] a) L. E. Janes, A. C. Lowendahl, R. J. Kazlauskas, *Chem. Eur. J.* **1998**, 4, 2324–2331; b) A. Banerjee, P. Kaul, R. Sharma, U. C. Banerjee, *J. Biomol. Screening* **2003**, 8, 559–565.
- [11] a) D. Grée, L. Vallerie, R. Gre, L. Toupet, I. Washington, *J. Org. Chem.* **2001**, 66, 2374–2381; b) M. Calí, M. Begtrup, *Synthesis* **2002**, 63–66.
- [12] K. Tamura, Y. Terao, K. Miyamoto, H. Ohta, *Biocatal. Biotransform.* **2008**, 26, 253–257.
- [13] J. W. Thanassi, *Biochemistry* **1970**, 9, 525–532.
- [14] a) M. T. Reetz, J. D. Carballeira, A. Vogel, *Angew. Chem.* **2006**, 118, 7909–7915; *Angew. Chem. Int. Ed.* **2006**, 45, 7745–7751; b) M. T. Reetz, J. D. Carballeira, *Nat. Protoc.* **2007**, 2, 891–903.
- [15] a) E. J. Corey, T. Hase, *Tetrahedron Lett.* **1979**, 20, 335–338; b) J.-M. Garnier, S. Robin, R. Guillot, G. Rousseau, *Tetrahedron: Asymmetry* **2007**, 18, 1434–1442.
- [16] G. T. Crisp, P. T. Glink, *Tetrahedron* **1992**, 48, 3541–3556.
- [17] G. T. Pearce, W. E. Gore, R. M. Silverstein, *J. Org. Chem.* **1976**, 41, 2797–2803.
- [18] a) Y. Zheng, X. Chen, Y. Shen, *Chem. Rev.* **2008**, 108, 5253–5277; b) R. Ciriminna, M. Pagliaro, *Adv. Synth. Catal.* **2003**, 345, 383–388.
- [19] The atomic coordinates for *B. bronchiseptica* AMDase with benzylphosphonate bound to the active site dioxyanion hole has been deposited in the Protein Data Bank (accession number 3IP8).