

DOI: 10.1002/cbic.201300321

# Structure and Activity of NADPH-Dependent Reductase Q1EQE0 from *Streptomyces kanamyceticus*, which Catalyses the *R*-Selective Reduction of an Imine Substrate

María Rodríguez-Mata,<sup>[a, b]</sup> Annika Frank,<sup>[a]</sup> Elizabeth Wells,<sup>[a]</sup> Friedemann Leipold,<sup>[c]</sup> Nicholas J. Turner,<sup>[c]</sup> Sam Hart,<sup>[a]</sup> Johan P. Turkenburg,<sup>[a]</sup> and Gideon Grogan<sup>\*[a]</sup>

NADPH-dependent oxidoreductase Q1EQE0 from *Streptomyces kanamyceticus* catalyzes the asymmetric reduction of the prochiral monocyclic imine 2-methyl-1-pyrroline to the chiral amine (*R*)-2-methylpyrrolidine with >99% *ee*, and is thus of interest as a potential biocatalyst for the production of optically active amines. The structures of Q1EQE0 in native form, and in complex with the nicotinamide cofactor NADPH have been solved and refined to a resolution of 2.7 Å. Q1EQE0 functions as a dimer in which the monomer consists of an N-terminal Rossmann-fold motif attached to a helical C-terminal domain

through a helix of 28 amino acids. The dimer is formed through reciprocal domain sharing in which the C-terminal domains are swapped, with a substrate-binding cleft formed between the N-terminal subunit of monomer A and the C-terminal subunit of monomer B. The structure is related to those of known  $\beta$ -hydroxyacid dehydrogenases, except that the essential lysine, which serves as an acid/base in the (de)protonation of the nascent alcohol in those enzymes, is replaced by an aspartate residue, Asp187 in Q1EQE0. Mutation of Asp187 to either asparagine or alanine resulted in an inactive enzyme.

## Introduction

Optically active amines constitute some of the most important synthetic intermediates in pharmaceuticals synthesis. Interest in developing sustainable chemical routes to optically active amines has stimulated research in various enzymatic methods for their production.<sup>[1]</sup> Many enzymes are capable of synthesizing chiral amines from prochiral or racemic precursors, with examples ranging from the (dynamic) kinetic resolution of racemic amines using N-acylases<sup>[2]</sup> or other hydrolases, transaminases for the quantitative conversion of prochiral ketones to amines at the expense of an ammonia donor,<sup>[3]</sup> and the deracemisation of racemic amines using flavin-dependent amine oxidases as part of a chemoenzymatic oxidation–reduction cycle with a chemical reductant.<sup>[4]</sup> In industry, optically active amines are also derived by chemical reduction from prochiral imines, themselves easily generated from precursor ketones through their reaction with amines. Many recent examples

exist of abiotic catalysts for the asymmetric reduction of prochiral imine substrates for this purpose.<sup>[5–8]</sup>

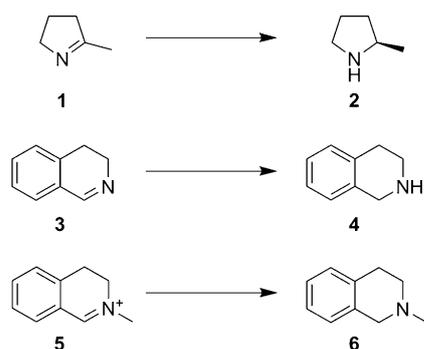
The enzymatic asymmetric reduction of imines has been less well-explored in the area of preparative biocatalysis, largely because of the aqueous lability of the imine substrates, but also because of the dearth of apparent “imine reductases” in the literature. The well-studied dihydrofolate reductase (DHFR) from *E. coli*<sup>[9]</sup> and related enzymes perform asymmetric hydrogenation of its imine substrate dihydrofolate, but the enzyme has not been the focus of study as a biocatalyst for chiral amine preparation, save for its application in the generation of its native product, (*S*)-tetrahydrofolate.<sup>[10]</sup> However, asymmetric imine reduction by whole-cell preparations of yeasts has recently been reported<sup>[11, 12]</sup> and imine reductase activity has also been attributed to some strains of anaerobic bacteria.<sup>[13]</sup> An NADPH-dependent reductase, PchG, from *Pseudomonas aeruginosa* and acting as a thiazonolonyl imine reductase has also been discovered in the biosynthetic pathway towards the siderophore pyochelin<sup>[14]</sup> and a related enzyme, Irp3 from *Yersinia enterocolitica*, has recently been the subject of structural studies.<sup>[15]</sup> However, the most interesting candidate enzymes possessing imine reductase activity with potential for application come from *Streptomyces* spp. Imine reductase activities dependent on the unusual deazaflavin cofactor F-420 have been implicated in the biosynthetic pathways of sibiromycin,<sup>[16]</sup> tomaymycin<sup>[17]</sup> and chlortetracycline,<sup>[18]</sup> amongst others, in various *Streptomyces* species. Additionally, a wide-ranging screen of *Streptomyces* organisms revealed asymmetric imine reductase activity in two strains: *Streptomyces* sp. GF3587 catalysed the reduction of 2-methyl-1-pyrroline (**1**, Scheme 1) to (*R*)-2-methylpyrrolidine (**2**) with 99.2% *ee*.<sup>[19]</sup> Another strain, *Streptomyces*

[a] Dr. M. Rodríguez-Mata, A. Frank, E. Wells, S. Hart, Dr. J. P. Turkenburg, Dr. G. Grogan  
York Structural Biology Laboratory  
Department of Chemistry, University of York  
Heslington, York, YO10 5DD (UK)  
E-mail: gideon.grogan@york.ac.uk

[b] Dr. M. Rodríguez-Mata  
Universidad de Oviedo  
Facultad de Química, Química Orgánica e Inorgánica  
Julian Clavería 8, C.P. 33006 Oviedo, Asturias (Spain)

[c] Dr. F. Leipold, Prof. N. J. Turner  
Manchester Institute of Biotechnology, University of Manchester  
131 Princess Street, Manchester, M1 7DN (U.K.)

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201300321>.



**Scheme 1.** Biotransformation of imine substrates by imine reductase Q1EQE0 from *S. kanamyceticus*.

sp. GF3546, catalysed asymmetric reduction of **1** to the *S* enantiomer, with 92.3% *ee*. Further research by the Mitsukura group was performed in which the *R*-selective enzyme (termed RIR) was purified and characterised.<sup>[20]</sup> The enzyme appeared to be a dimer composed of monomer subunits of 32 kDa, and  $K_m$  and  $V_{max}$  values of 3.5 mM and  $10.2 \mu\text{mol min}^{-1} \text{mg}^{-1}$  respectively for substrate **1** were recorded. Cofactor requirement was limited to NADPH. The acquisition of N-terminal amino acid sequence data led to the cloning of the gene encoding RIR, which allowed the identification of a homologue, Uniprot accession number Q1EQE0, in the genome of *Streptomyces kanamyceticus*,<sup>[21]</sup> which was also shown to convert **1** to (*R*)-**2** with 99.6% *ee*.<sup>[22]</sup> A sequence alignment of RIR with Q1EQE0 revealed significant homology throughout the length of the protein chain, with 50% sequence identity and a further 19% strongly similar residues (Figure S1). Q1EQE0 is slightly longer (309 residues versus 295 for RIR), due to an insertion of 14 amino acids in the N-terminal region that is absent in RIR.

Genes encoding RIR and Q1EQE0,<sup>[22]</sup> and also the *S*-selective imine reductase (SIR) from *Streptomyces* sp. GF3546,<sup>[23]</sup> were cloned and expressed heterologously in *E. coli*, creating recombinant biocatalysts for the reduction of the imine substrate. No details concerning the purification and characterization of Q1EQE0 were reported.

As part of a screen of enzymes suitable for the reduction of synthetically relevant imines, we synthesised the gene encoding Q1EQE0 with the sequence codon-optimised for expression in *E. coli*. In this study, we report the structure of Q1EQE0 from *S. kanamyceticus* to 2.7 Å, and its complex with the cofactor NADPH, and describe structural homology with hydroxyisobutyrate dehydrogenases that may shed light on mechanism. It is envisaged that the structure may be used as the basis for rational engineering of enzymes catalysing the asymmetric reduction of imines for improved activity and wider substrate specificity.

## Results and Discussion

### Cloning and expression of the gene encoding Q1EQE0 and purification and assay of the enzyme

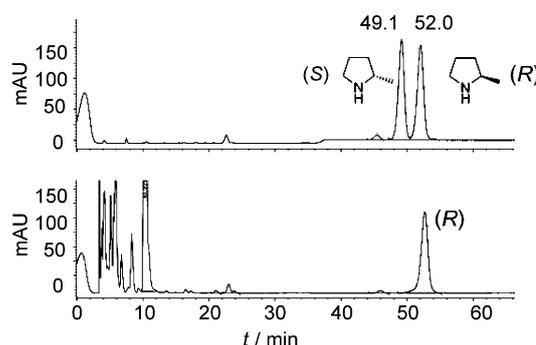
The gene encoding Q1EQE0 was synthesised with codons optimised for expression in *E. coli* BL21 (DE3). Good levels of solu-

ble expression were achieved and purification of the enzyme using nickel-affinity and size-exclusion chromatography yielded a protein which, confirmed by SDS-PAGE analysis, had an approximate molecular weight of 32 kDa (Figure S2). Using a standard spectrophotometric assay for the measurement of substrate-stimulated NADPH oxidation, a  $K_m$  value of  $8.21 \pm 1.07 \text{ mM}$  and a  $k_{cat}$  of  $(0.018 \pm 1.4) \times 10^{-3} \text{ s}^{-1}$  for substrate **1** were recorded (Figure S3, Table 1).<sup>[20]</sup> At concentrations of

**Table 1.** Kinetic constants for NADPH-dependent reductases Q1EQE0 acting on substrates **1**, **3** and **5**.

Substrate	$k_{cat}$ [ $\text{s}^{-1}$ ]	$K_m$ [mM]	$k_{cat}/K_m$ [ $\text{s}^{-1} \text{mM}^{-1}$ ]
<b>1</b>	$0.018 \pm 1.4 \times 10^{-3}$	$8.21 \pm 1.07$	$2.2 \times 10^{-3} \pm 3.0 \times 10^{-4}$
<b>3</b>	$9 \times 10^{-4} \pm 2.0 \times 10^{-5}$	$1.16 \pm 0.15$	$7.4 \times 10^{-4} \pm 1.0 \times 10^{-4}$
<b>5</b>	$3.8 \times 10^{-3} \pm 1.0 \times 10^{-4}$	$0.724 \pm 0.09$	$5.2 \times 10^{-3} \pm 1.0 \times 10^{-4}$

above 30 mM **1**, significant substrate inhibition was observed, with a calculated  $K_i$  value of  $56.5 \pm 10.7 \text{ mM}$ . No activity was detected using NADH as the cofactor. Biotransformations of 5 mM **1** using the pure enzyme and NADPH, in the presence of glucose-6-phosphate and glucose-6-phosphate dehydrogenase for NADPH recycling, gave the *R* product in >99% *ee* with 23% conversion (Figure 1). No significant reaction was observed in the absence of Q1EQE0.



**Figure 1.** Chiral HPLC chromatograms of GITC-derivatised amine **2**. The top trace shows the GITC-derivatised racemic amine **2**, with the *S* and *R* amines eluting at 49.1 and 52.0 min respectively; the bottom trace shows the GITC-derivatised (*R*)-**2** recovered from biotransformation of 2-methyl-1-pyrroline **1** (5 mM) by NADPH-dependent Q1EQE0.

Low conversions (<5%) of imine substrates 3,4-dihydroisoquinoline (**3**, Scheme 1, Figure S4A and B) and 2-methyl-3,4-dihydroisoquinolin-2-ium triflate (**5**, Scheme 1, Figure S5A and B) were also observed over a period of 24 h, during which negligible degradation of the substrates was observed in buffered solution.

The kinetic constants for the transformations of these substrates are presented in Table 1. Substrates **3** and **5** were not reported to be substrates for RIR or Q1EQE0 previously.

## Structure of Q1EQE0

The structure of Q1EQE0 was solved by molecular replacement. Data collection and refinement statistics are given in Table 2. For structure solution, the highest sequence homologue in the Protein Data Bank was an oxidoreductase from *Pseudomonas putida* KT2440 (PDB ID: 3L6D, 26% sequence identity to Q1EQE0), although the molecular replacement pipeline BALBES<sup>[24]</sup> selected an alternative model for molecular replace-

**Table 2.** Data collection and refinement statistics for oxidoreductase Q1EQE0 in native form and in complex with NADPH. Numbers in brackets refer to data for highest resolution shells.

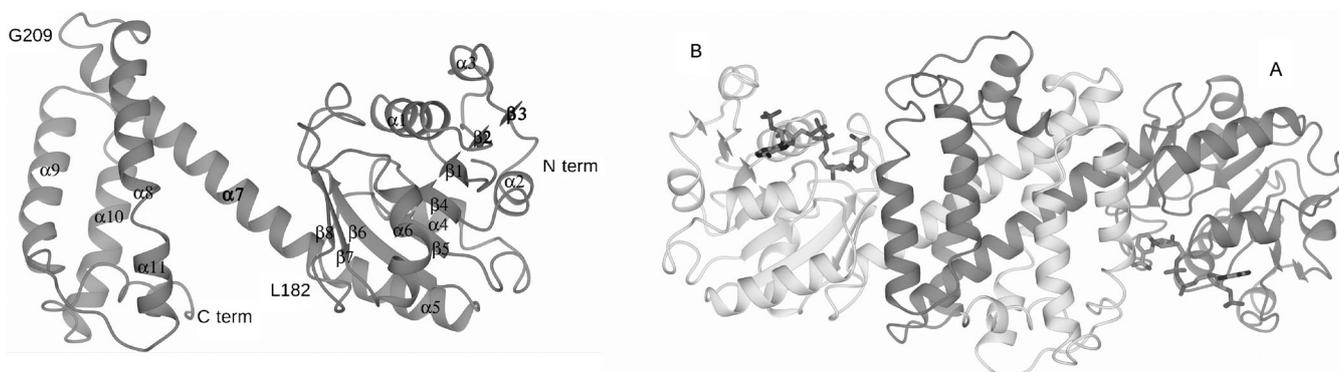
	Q1EQE0 native	Q1EQE0 NADPH
beamline	Diamond I03	Diamond I04-1
wavelength [Å]	0.97630	0.91999
resolution [Å]	108.68–2.71 (2.79–2.71)	74.27–2.73 (2.80–2.73)
space group	C121	C121
unit cell		
<i>a</i> , <i>b</i> , <i>c</i> [Å]	203.41, 131.11, 77.48	205.80, 130.30, 77.70
$\alpha$ , $\beta$ , $\gamma$ [°]	90.00, 107.22, 90.00	90.00, 107.10, 90.00
no. of molecules in the asymmetric unit	4	4
unique reflections	49 759	51 705
completeness [%]	99.3 (99.6)	99.4 (99.1)
<i>R</i> <sub>merge</sub> [%]	0.05 (0.52)	0.08 (0.60)
<i>R</i> <sub>p.i.m.</sub>	0.05 (0.49)	0.07 (0.50)
multiplicity	3.4 (3.5)	4.3 (4.3)
$\langle I/\sigma(I) \rangle$	13.9 (2.3)	12.2 (2.0)
CC <sub>1/2</sub>	0.99 (0.80)	1.00 (0.78)
overall <i>B</i> factor from Wilson plot [Å <sup>2</sup> ]	55	47
<i>R</i> <sub>cryst</sub> / <i>R</i> <sub>free</sub> [%]	18.5/21.1	18.5/21.6
no. protein atoms	7915	8094
no. water molecules	106	111
RMSD 1–2 bonds [Å]	0.014	0.015
RMSD 1–3 angles [°]	1.55	1.79
average main chain B [Å <sup>2</sup> ]	66	57
average side chain B [Å <sup>2</sup> ]	67	58
average water B [Å <sup>2</sup> ]	52	43
average ligand B [Å <sup>2</sup> ]	–	64

ment; the tartronic semialdehyde reductase from *Salmonella typhimurium* (1YB4, 19%).<sup>[25]</sup>

The solution contained four monomers (A–D), arranged as two pairs of dimers (A–B and C–D). The monomer, represented by subunit A, of Q1EQE0 consists of an N-terminal Rossmann-fold domain of approximately 181 amino acids (1–181) and a C-terminal helical domain (210–306) connected by a long interdomain helix from residues Leu182–Gly209 (Figure 2, left). The N-terminal domain consists of a  $\beta$ -sheet of a sheet of six parallel strands [ $\beta$ 1 (22–25),  $\beta$ 2 (45–48),  $\beta$ 3 (64–65),  $\beta$ 4 (78–81),  $\beta$ 5 (105–108),  $\beta$ 6 (130–136)] adjacent to two antiparallel strands [ $\beta$ 7 (149–153),  $\beta$ 8 (172–175)] interspersed in sequence and encapsulated by six well-defined  $\alpha$ -helices:  $\alpha$ 1 (29–42),  $\alpha$ 2 (56–60),  $\alpha$ 3 (69–74),  $\alpha$ 4 (86–93),  $\alpha$ 5 (114–127) and  $\alpha$ 6 (156–167). The C-terminal domain is comprised of the end of the interconnecting helix  $\alpha$ 7, and a further four helices:  $\alpha$ 8 (216–241),  $\alpha$ 9 (252–269),  $\alpha$ 10 (274–289) and  $\alpha$ 11 (297–303). In the native enzyme structure, the model covered residues 18–306 in subunits A–C, but it was also possible to model N-terminal residues 10–17 in subunit D. This was also the case for the NADP complex structure.

In the native structure, each subunit featured poorer density in the NADPH phosphate recognition loop between residues 51 and 65, meaning that residues 53 to 56 were not modelled in subunits A and D, and residues 51–65 were not modelled in subunits B and C. Near these regions of poorer density, Asn49 (B) and Ala52 (A) were noted as the only outliers in the Ramachandran plot. In the NADPH complex structure, the density in this loop was much improved, with only residues 55 in chains A–D and 54 in subunit B not modelled. Only two residues, Ala52 (C) and Ala53(D) were Ramachandran outliers in the NADPH complex structure.

An electrostatic surface of the monomer revealed the interconnecting helix  $\alpha$ 7 to be somewhat hydrophobic, which would prove to be significant in dimer formation (Figure S6). Analysis of the monomer structure using the DALI server,<sup>[26]</sup> showed that, of structures of enzymes for which an activity has been determined, Q1EQE0 is most similar to members of the gamma-hydroxybutyrate dehydrogenase (GHBDH) family



**Figure 2.** Left: Structure of Q1EQE0 monomer illustrating N-terminal Rossmann-fold domain made up of  $\alpha$ -helices  $\alpha$ 1–6 and  $\beta$ -strands  $\beta$ 1–8 and C-terminal helical domain containing  $\alpha$ -helices  $\alpha$ 8–11 connected by inter-domain helix  $\alpha$ 7. Helix  $\alpha$ 7 stretches from Leu182 to Gly209. Right: Structure of Q1EQE0 dimer, formed by reciprocal domain swapping between subunits A (dark grey) and B (light grey); NADPH is shown in cylinder format with the carbon atoms in grey. NADPH is bound at the interface formed between, in one case, the N-terminal domain of A and the C-terminal helical domain of B.

such as PDB structures 3PEF and 3PDU, or of the hydroxyisobutyrate dehydrogenase (HIBDH) family, such as 2CVZ, but there are significant differences in orientation of the C-terminal domain in Q1EQE0 that result in extensive domain swapping not observed in characterised HIBDHs as described below.

### Dimer formation

The active form of Q1EQE0 is a dimer, which is formed by reciprocal domain swapping between two subunits (Figure 2, right) and signified in the description below by the notation A and B. Analysis by PISA<sup>[27]</sup> indicated a total contact area of 3761 Å<sup>2</sup> between the subunits. The N-terminal Rossmann-fold domain of subunit A makes contacts with the C-terminal helical domain of subunit B forming a cleft which constitutes the active site of the enzyme (vide infra); this interface is stabilised by inter-subunit interactions including that between Ser114 (A) at the beginning of helix  $\alpha$ 4 and Glu265 (B) in helix  $\alpha$ 8. The N-terminal domains are at the periphery of the dimer. At the end of N-terminal domain A, the interdomain helix  $\alpha$ 7 protrudes and inserts through a hydrophobic channel in the C-terminal domain of B (Figure S7) to then emerge and continue as the C-terminal helical domain of A, which contacts the N-terminal domain of subunit B. Reciprocal salt bridges between Arg285 (A or B) and Glu274 (A or B) assist in stabilization of the dimer.

As described above, DALI-based analysis of the structure of Q1EQE0 revealed that the structures of known activity that most resemble Q1EQE0 are NADPH-dependent dehydrogenases acting on isobutyrate substrates of the GHBDH or HIBDH families. However, in representatives of each of those cases, although the structures are dimeric, extensive domain swapping is not observed. Superimposition of the Q1EQE0 monomer with a monomer of 2CVZ for example, a HIBDH from *Thermus thermophilus* HB8<sup>[28]</sup> (Figure S8A and S8B) shows that there is good fold conservation of the N-terminal domain and the interdomain helix  $\alpha$ 7 (RMSD 1.7 Å for residues 18–209 in Q1EQE0 and 2–190 for 2CVZ). There is also good fold conservation between the isolated C-terminal domains (2.0 Å for residues 232–303 in Q1EQE0 superimposed with residues 212–282 for 2CVZ). However a sharp  $\beta$ -turn at residue Ser204 in 2CVZ in the helix equivalent to  $\alpha$ 8 in Q1EQE0 results in the chain returning in the direction of the N-terminal domain of A, to continue and form the C-terminal helical bundle required to complete the active site cleft within the same monomer. In Q1EQE0, helix  $\alpha$ 8 continues at the equivalent point (Gln224) and travels away from the N-terminal domain to form the C-terminal helical bundle that will form the active site cleft with the N-terminal domain of subunit B.

Domain swapping of this kind in the dehydrogenases has not been reported extensively in the literature, although one further outcome of the DALI-based analysis was the identification of two structures that share more significant structural homology with Q1EQE0 throughout the length of the chain, and which also participate in domain swapping. These are 3QHA (from *Mycobacterium avium* 104) and 3L6D (from *P. putida* KT2440), although in each case annotation is restricted to putative oxidoreductases of as yet undetermined activity. A su-

perimposition of these structures with Q1EQE0 can be found in the Supporting Information (Figure S9A and S9B).

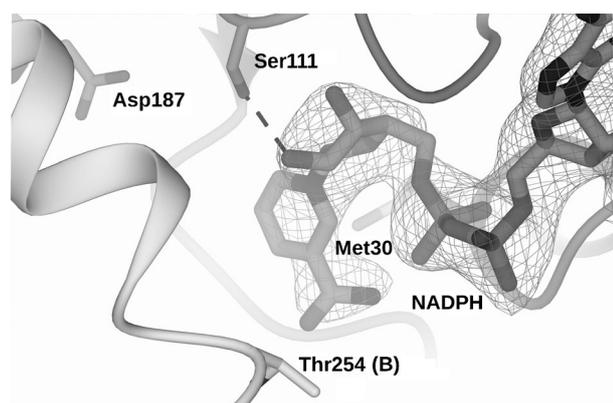
### NADPH binding in Q1EQE0 and the active site

Datasets resulting from crystals of Q1EQE0 that had been soaked in a solution of NADPH (10 mM) resulted in electron density maps in which the omit map at the subunit interface could be modelled unambiguously as the cofactor, with excellent occupancy in each of the four active sites found in the pair of dimers in the structure solution. The electron density of residues and their side chains in the region binding the NADPH ribose-2'-phosphate (Asn49–Lys54) was much improved in the NADPH complex than in the native structure. The cofactors (two per dimer) are bound in clefts formed by, in one instance, the N-terminal domain of subunit A, helices  $\alpha$ 7 and  $\alpha$ 8 of the C-terminal helical bundle of subunit B and the loop that connects them (Figure 2, right). NADPH only forms close interactions with the N-terminal domain however. The active site of Q1EQE0 is a large channel that traverses the enzyme structure at the surface of the dimer interface. Apart from residues that bind the phosphate residues of NADPH, the entrance to the channel and the channel itself are strikingly negatively charged, perhaps commensurate with its activity in recognising, binding, and perhaps stabilising, imine substrates or their iminium ions (Figure S10).

The characteristic GXGXXG consensus sequence for NADPH binding in Q1EQE0 is present as G(26) LGMLG (31). The ribose-2'-phosphate of NADPH is secured by interactions with the following residues: Arg50, which also makes  $\pi$ -stacking interactions with the adenine ring of the ADP moiety; Lys54; Asn49; and the side-chain and peptidic N–H of Thr51. The 2' and 3'-ribose hydroxyls interact with the side chain of Ser111 and the backbone of Ser111, Val83 and Ser84. The *re*-face of the nicotinamide ring stacks against the side chain of Met30; the *si*-face is presented to the active-site cavity (Figure 3). Protic residue side chains, which may have a role in mechanism (vide infra) within a 9 Å sphere of the nicotinamide ring include Thr254 from subunit B and Ser111 and Asp187 from subunit A.

In NAD(P)H-dependent ketoreductases such as the HIBDH 2CVZ, a protic residue within the vicinity of the active site nicotinamide ring will be required to provide a proton to the nascent hydroxyl group in the reductive direction. In 2CVZ, this function is fulfilled by Lys165.<sup>[28]</sup> Superimposition of the active sites of Q1EQE0 and 2CVZ reveal that the lysine residue is replaced by the aspartate Asp187 (Figure S11). Mutation of this residue to either alanine (Asp187Ala) or asparagine (Asp187Asn) resulted in inactive enzyme variants. Additionally, the wild-type Q1EQE0 displayed no HIBDH activity towards commercially available 3-hydroxyisobutyrate when supplied with NADP<sup>+</sup>.

The biological reduction of imines by possible imine reductases has been the focus of some recent interest owing to the role of these activities in natural product biosynthesis, but also, increasingly because of the potential for enzymes capable of asymmetric imine reduction for preparative biocatalysis. Prior to either of those areas being explored, however, the best-



**Figure 3.** Active site of Q1EQE0. The peptide backbone is shown in ribbon format with subunits A in dark grey and B in light grey. Side chains of subunits A and B are shown in cylinder format with carbon atoms in dark grey and light grey respectively. NADPH is shown in cylinder format with carbon atoms in grey. The dashed black line represents hydrogen bonding interaction between Ser111 and NADPH. Electron density map represents the  $F_o - F_c$  omit map refined in the absence of NADPH, contoured at a level of  $3\sigma$ . The NADPH atoms from the refined complex have then been added for clarity.

studied imine reductase is undoubtedly DHFR, such as the enzyme from *E. coli*, and which catalyses the asymmetric reduction of dihydrofolate to tetrahydrofolate at the expense of NADPH.<sup>[9]</sup> Many structures of this enzyme and its mutants have been reported, in complex with both the NADPH cofactor,<sup>[29]</sup> substrate dihydrofolate,<sup>[30]</sup> methotrexate<sup>[30]</sup> and other inhibitors.<sup>[31]</sup> Recently, the structure of an imine reductase (Irp3) acting on a thiazoline imine attached to an acyl-carrier protein from *Yersinia enterocolitica*, and with homology to the PchG enzyme from the pyochelin biosynthetic pathway in *P. putida*, has also been presented.<sup>[15]</sup> None of these enzymes has yet been shown to possess asymmetric imine reductase activity towards imine substrates of synthetic interest however. Q1EQE0 and the RIR and SIR from *Streptomyces* spp. described by Mitsukura and co-workers represent the first examples of isolated enzymes capable of such reductions, and the structure of Q1EQE0 allows these reductive reactions to be put into an enzyme structural context for the first time. The low values of catalytic efficiency of Q1EQE0 towards **1**, **3** and **5** (Table 1) clearly suggest that these are not the natural substrates for the enzyme. The identities of genes surrounding that which encodes Q1EQE0 in the genome of *S. kanamyceticus* are also not informative in this context. The active site of Q1EQE0 is large however, hinting perhaps at a larger natural substrate than those studied herein. It is also possible that the reduction of imines represents a promiscuous activity for these oxidoreductases of otherwise uncharacterised activity.

It is interesting to compare the structure and proposed mechanism of DHFR and Irp3 with Q1EQE0. While each of these reductase enzymes shares the Rossmann fold associated with binding the ADP moiety of the NADP cofactor, both Q1EQE0 and Irp3 possess a separate C-terminal domain that is absent in DHFR. DHFR binds both cofactor and substrate within a discrete single domain that shares some structural homology with both other imine reductases, notably of the core  $\beta$ -sheet region (Figure S12A). In DHFR, however, the nicotina-

mid ring is bound within the core of the Rossmann fold, within which it is brought into contact with the DHFR substrate as observed in PDB structure 7DFR.<sup>[30]</sup> In Irp3, the nicotinamide ring is presented at the surface of the N-terminal domain to the active site (Figure S12B), as a result of rotation around the bond connecting the ribose to the diphosphate, relative to the case with DHFR. In this respect, Q1EQE0 is more similar to Irp3 except that the active site is formed within one subunit in the latter enzyme.

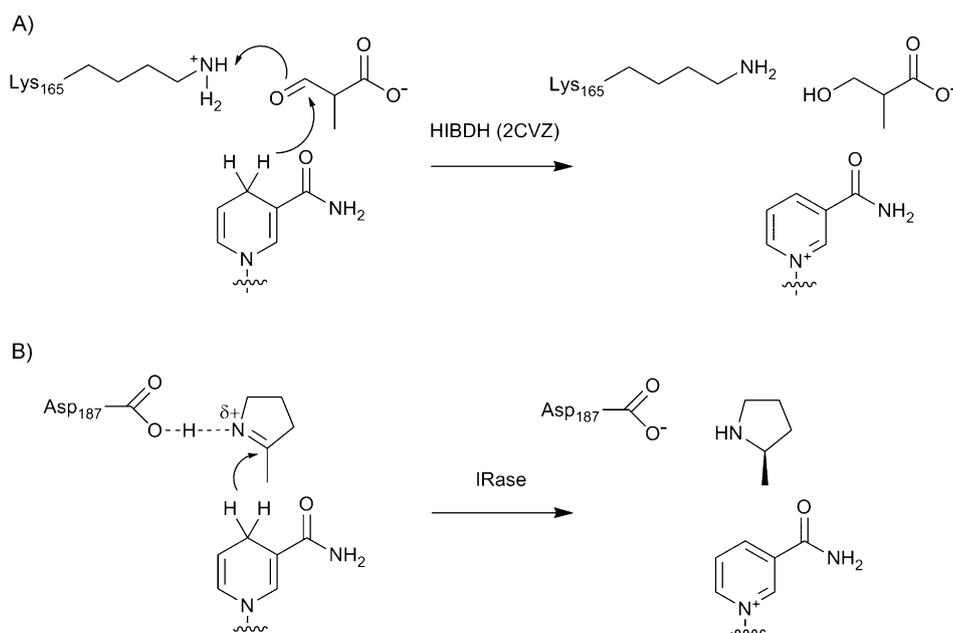
The putative mechanisms of C=X (where X=N or O) bond reduction in DHFR, Irp3 and the ketoreductase 2CVZ, which Q1EQE0 resembles structurally, appear to be distinct. DHFRs from *E. coli* and other bacteria are thought to first protonate the imine through a water molecule, to form an iminium ion that is reduced by hydride.<sup>[32]</sup> In Irp3, hydride transfer is thought to be followed by proton transfer from either a histidine (His101) or a tyrosine (Tyr128) residue, both of which may be expected to be in the protonated form at physiological pH.<sup>[15]</sup> In 2CVZ, a proton is thought to be delivered to the nascent alcohol product by the protonated lysine residue Lys165,<sup>[28]</sup> which superimposes well with Asp187 in Q1EQE0. Mutation of Asp187 in Q1EQE0 to either asparagine or alanine yields an inactive enzyme. In the absence of a structure of Q1EQE0 in complex with both NADPH and **1** or other substrate, it is difficult to precisely determine the role of this residue in catalysis, although its role as a proton donor to the imine may be facilitated by an increase in  $pK_a$  as a result of being sandwiched between the hydrophobic residues Leu191 and Leu137; an environment not provided in 2CVZ, in which a serine (Ser117) and an asparagine (Asn169) interact with Lys165. Protonation would result in an iminium ion that would be the substrate for reduction by hydride (Scheme 2). The stabilisation of such an iminium ion may be fostered by the predominantly negatively charged active-site channel.

## Conclusions

The structure of Q1EQE0 represents the first opportunity to study the active-site determinants of mechanism and selectivity in NADPH-reductases that might be of use in industrial biocatalytic processes for the reduction of imine substrates. The exact mechanism of the imine reduction remains to be elucidated. Whilst the substrate specificity of this and related wild-type enzymes appear to be narrow, the structures will provide a basis for protein engineering experiments targeted at improving enzyme activity and altering characteristics such as stability, process suitability, substrate range and enantioselectivity.

## Experimental Section

**Gene synthesis, cloning, expression and protein purification:** The gene encoding Q1EQE0 was synthesised by GeneArt (Invitrogen), with codons optimised for expression in *E. coli*. The full gene sequence can be found in the Supporting Information (Figure S13). The gene was amplified by PCR from the commercial template using the following primers: 5'-CCAGGG ACCAGC AATGCC GGATAA



**Scheme 2.** Contrasting mechanisms proposed for carbonyl and imine group reduction in HIBDH and Q1EQE0, respectively. A) Lys165 protonates the nascent alcohol in the reduction of hydroxyisobutyric acid by 2CVZ.<sup>[28]</sup> B) Asp187 of Q1EQE0 provides stabilization and protonation to a substrate iminium form of **1**, which is reduced by hydride from NADPH to yield the *R* product.

TCCGAG CACCAA AGGTC-3' (forward) and 5'-GAGGAG AAGCGG CGTTAT TATTTA CCGCTA TGGGTA CGAAAC TGTC-3' (reverse). After gel analysis of the PCR product, the relevant band was eluted from the gel using a PCR cleanup kit (Qiagen). The gene was then sub-cloned into the pET-YSBL-LIC-3C vector using a published ligation-independent cloning procedure.<sup>[33]</sup> The resultant plasmid MRM1 was then used to transform *E. coli* XL1-Blue cells (Novagen), yielding colonies which in turn gave plasmids using standard miniprep procedures that were sequenced to confirm the identity and sequence of the gene. The native gene was used as a template for the creation of mutants Asp187Ala and Asp187Asn using a Quick-change site-directed mutagenesis kit from Stratagene, using the manufacturer's protocol. The primers for the Asp187Ala were: 5'-CAAGCC TGTATG CGGCCG CAGGTC TGG-3' (forward) and 5'-CCAGAC CTGCGG CCGCAT ACAGGC TTG-3' (reverse). For Asp187Ala, the primers used were: 5'-GGCAAG CCTGTA TAATGC CGCAGG TC-3' (forward) and 5'-GACCTG CGGCAT TATACA GGCTT CC-3' (reverse). Following cloning, the presence of the designed mutation sites was confirmed by sequencing.

The recombinant vector(s) containing the Q1EQE0 gene and its Asp187Ala and Asp187Asn mutants were used to transform *E. coli* BL21 (DE3) cells using kanamycin (30  $\mu\text{g mL}^{-1}$ ) as antibiotic marker on Luria-Bertani (LB) agar. A single colony from an agar plate grown overnight was used to inoculate LB broth (5 mL), which was then grown overnight at 37 °C with shaking at 180 rpm. The starter culture served as an inoculum for a culture of LB broth (500 mL) in which cells were grown until the optical density ( $\text{OD}_{600}$ ) had reached a value of 0.6. Expression of Q1EQE0 was then induced by the addition of isopropyl  $\beta$ -D-1-thiogalactopyranoside (final concentration of 1 mM). The culture was then incubated at 18 °C in an orbital shaker overnight at 180 rpm for approximately 18 h. The cells were harvested by centrifugation (4225 *g*, 15 min) in a Sorvall RC5B Plus centrifuge with a Sorvall G53 rotor and then resuspended in Tris-HCl buffer (50 mM, pH 7.5, 50 mL) containing NaCl (300 mM). The cells were then sonicated for 3  $\times$  30 s bursts at 4 °C

with 1 min intervals and the soluble and insoluble fractions were separated by centrifugation (26892 *g*, 30 min) in a Sorvall SS34 rotor. The clear supernatant was loaded onto a 5 mL His-Trap chelating HP nickel column. After washing with Tris-HCl buffer (ten column volumes) containing imidazole (30 mM), the Q1EQE0 protein was eluted with a gradient of imidazole (30–500 mM) over 20 column volumes. Column fractions containing Q1EQE0 (as determined by SDS-PAGE analysis) were pooled and then concentrated using a 10 kDa cut-off Centricon filter membrane. The concentrated enzyme was then loaded onto a pre-equilibrated S75 Superdex 16/60 gel-filtration column, and eluted with the Tris-HCl buffer (120 mL) at a flow rate of 1 mL min<sup>-1</sup>. Fractions containing pure Q1EQE0, as determined by SDS-PAGE analysis were pooled and stored at 4 °C.

**Protein crystallisation:** Crystallisation conditions for Q1EQE0 were determined using a range of commercially available trial screens in sitting-drop format 96-well plates with 300 nL drops. The best initial crystals were obtained in conditions containing 35% (*v/v*) tascimate at pH 7.0 and protein at a concentration of 10 mg mL<sup>-1</sup> (using a 1:1 ratio of protein and precipitant solution). Larger crystals for diffraction analysis using optimised conditions were prepared using the hanging-drop method in 24-well plate Linbro dishes with 2  $\mu\text{L}$  drops consisting of 1  $\mu\text{L}$  of protein at a concentration of 10 mg mL<sup>-1</sup> and 1  $\mu\text{L}$  of the reservoir. The best crystals were obtained in 35% (*v/v*) tascimate with 2.5% (*v/v*) ethylene glycol. Cocrystallisation with NADPH (10 mM) did not yield crystals under the same buffer conditions. In order to obtain the NADP(H) complex, native crystals were transferred from the growth drop into a cryogenic solution which consisted of the mother liquor containing 10% (*v/v*) glycerol and NADPH (10 mM), and incubated for 1 min, after which they were immediately flash-cooled in liquid nitrogen. Native crystals were flash-cooled in a cryogenic solution containing the only the mother liquor and 10% (*v/v*) glycerol. All crystals were tested for diffraction in-house using a Rigaku Micromax-007HF fitted with Osmic multilayer optics and a Marresearch MAR345 imaging plate detector. Those crystals that diffracted to greater than 3 Å resolution were retained for full dataset collection at the synchrotron.

**Data collection, structure solution, model building and refinement of Q1EQE0:** Complete datasets for native Q1EQE0 and the NADPH complex were collected on beam lines I03 and I04-1 respectively at the Diamond Light Source (Didcot, UK) The data were processed and integrated using XDS<sup>[34]</sup> and scaled using SCALA<sup>[35]</sup> as included within the Xia2 processing system.<sup>[36]</sup> Data collection statistics are given in Table 1. The crystals of each complex were isomorphous and were in space group C121. The structure of the enzyme was solved using BALBES,<sup>[24]</sup> which selected a truncated dimer model of the tartronic semialdehyde reductase from *S. typhimurium* LT2 (PDB ID: 1YB4; 19% amino acid sequence identity to

QE1QE0 as a search model). The solutions each contained four molecules in the asymmetric unit, representing two dimers. The solvent content in each case was 69%. The structures were built using Autobuild in the Phenix suite of programs<sup>[37]</sup> and Coot<sup>[38]</sup> and refined using REFMAC<sup>[39]</sup> employing local NCS restraints. For the NADPH complex, following building and refinement of the protein and water molecules, the omit maps were observed to contain clear residual density at the subunit–dimer interface, which was modelled and refined as NADPH. The final structures exhibited  $R_{\text{cryst}}$  and  $R_{\text{free}}$  values of 18.5 and 21.1% (native) and 18.5 and 21.6% (NADPH). Each structure was finally validated using PROCHECK.<sup>[40]</sup> Refinement statistics are presented in Table 2. The Ramachandran plot for the native enzyme showed 97.6% of residues to be situated in the most favoured regions, 1.6% in additional allowed and 0.8% outlier residues. For the NADPH complex, the corresponding values were 96.6, 2.7 and 0.7% respectively. The coordinates and structure factors for the native imine reductase and the NADPH complex have been deposited in the Protein Data Bank with the PDB IDs 3ZGY and 3ZHB, respectively.

**Enzyme assays:** The NADPH-dependent imine reductase activity of recombinant QE1QE0 and its Asp187Ala and Asp187Asn mutants was assessed using UV spectrophotometry.<sup>[20]</sup> Activity was determined on a Spectramax M2 spectrophotometer/plate reader (Molecular Devices, Sunnyvale, CA) by monitoring the decrease of NADPH either at 340 nm ( $\epsilon = 6.22 \text{ mm}^{-1} \text{ cm}^{-1}$ ) for substrate **1** or at 370 nm ( $\epsilon = 2.216 \text{ mm}^{-1} \text{ cm}^{-1}$  for **3** and **5** as these substrates absorbed strongly at 340 nm. Activity was determined using microtiter plates (Sterilin, Newport, UK). Reaction mixtures (0.2 mL) contained pure enzyme (0.3 mg), NADPH (750  $\mu\text{M}$ ) and the appropriate amount of substrate (2  $\mu\text{L}$ ) from a 100-fold concentrated stock in DMSO. The reaction was started by adding the enzyme to the mixture and the reaction followed for 5 min. One unit of imine reductase is defined as the amount of protein that oxidised 1  $\mu\text{mol}$  NADPH per minute. Kinetic parameters were determined using purified enzyme and analysed by nonlinear regression analysis based on Michaelis–Menten kinetics using the program QtiPlot.

**Biotransformations:** In order to confirm the enantioselectivity of the reduction of **1** by QE1QE0, biotransformations were performed using the purified enzyme with NADPH-recycling by using the method of Faber and co-workers.<sup>[41]</sup> Reactions were performed in a 12 mL total reaction volume containing pure QE1QE0 (0.25  $\text{mg mL}^{-1}$ ) and substrate **1** (5 mM) from a stock solution (0.25 M) in DMF. The final concentration of dimethylformamide in the reaction was 2% (v/v). For cofactor recycling NADPH (5 mg, 0.006 mmol), glucose-6-phosphate (3.5 mg, 0.013 mmol) and glucose-6-phosphate dehydrogenase suspension (12  $\mu\text{L}$ ; G8404 from Sigma–Aldrich) was added per mL of total reaction volume. After overnight reaction at 30 °C, samples from biotransformations with 2-methyl-1-pyrroline (**1**), an aqueous sample of 100  $\mu\text{L}$  was mixed with triethylamine (100  $\mu\text{L}$  of a 0.8% (w/v) solution, equivalent to 11.00  $\mu\text{L mL}^{-1}$ ) in acetonitrile and 2,3,4,6-tetra-*O*-acetyl- $\beta$ -D-glucopyranosyl isothiocyanate (GITC, 200  $\mu\text{L}$  of a 0.8% (w/v) solution equivalent to 8  $\text{mg mL}^{-1}$ ). The mixture was incubated at 37 °C for 30 min. Insoluble fractions were removed by centrifugation and the supernatant analysed by HPLC.

Reversed-phase HPLC for samples of GITC-derivatised **2** was performed on an Agilent system equipped with a G1322A degasser, G1311A binary pump, a G1329A well-plate autosampler unit, a G1316A temperature-controlled column compartment and a G1315B diode array detector. The column Luna 3  $\mu\text{m}$  C18(2) 100 Å (Phenomenex 150 mm length, 4.6 mm diameter, 3  $\mu\text{m}$  particle size) was used. A typical injection volume was 10  $\mu\text{L}$  and chromatograms were monitored at 254 nm. The eluant was potassium phosphate buffer (10 mM, pH 2.5) and methanol in a 55:45 ratio. The elution times of the *S* and *R* enantiomers of **1** were 49.1 and 52.0 min respectively.

## Acknowledgements

We thank the Spanish Ministry of Science and Innovation for a FPU fellowship to M.R.M.; the industrial affiliates of the Centre of Excellence for Biocatalysis, Biotransformations and Biomanufacture (CoEBio3) for awarding a studentship to E.W.; the Marie-Curie ITN programme BIOTRAINS (FP7-PEOPLE-ITN-2008-238531) for awarding a studentship to A.F.; F.L. received funding from the European Union's Seventh Framework Programme FP7/2007-2013 under grant agreement n° 266025 (BIONEXGEN).

**Keywords:** amines • biocatalysis • imines • NADPH • oxidoreductases

- [1] M. Höhne, U. T. Bornscheuer, *ChemCatChem* **2009**, *1*, 42–51.
- [2] S. Buchholz, H. Gröger in *Biocatalysis in the Pharmaceutical and Biotechnology Industries* (Ed.: R. N. Patel), CRC Press, Boca Raton, **2006**, pp. 829–847.
- [3] S. Mathew, H. Yun, *ACS Catal.* **2012**, *2*, 993–1001.
- [4] R. Carr, M. Alexeeva, M. J. Dawson, V. Gotor-Fernández, C. E. Humphrey, N. J. Turner, *ChemBioChem* **2005**, *6*, 637–639.
- [5] K. A. Nolin, R. W. Ahn, F. D. Toste, *J. Am. Chem. Soc.* **2005**, *127*, 12462–12463.
- [6] M. Rueping, E. Sugiono, C. Azap, T. Theissmann, M. Bolte, *Org. Lett.* **2005**, *7*, 3781–3783.
- [7] Z. Wang, X. Ye, S. Wei, P. Wu, A. Zhang, J. Sun, *Org. Lett.* **2006**, *8*, 999–1001.
- [8] A. V. Malkov, K. Vranková, S. Stončius, P. Kočovský, *J. Org. Chem.* **2009**, *74*, 5839–5849.
- [9] D. R. Smith, J. M. Calvo, *Nucleic Acids Res.* **1980**, *8*, 2255–2274.
- [10] T. Eguchi, T. Oshiro, Y. Kuge, K. Mochida, T. Uwajima (Kyowa Hakko Kogyo Co., Ltd.), EP0356934A2, **1990**.
- [11] T. Vijayanthi, A. Chadha, *Tetrahedron: Asymmetry* **2008**, *19*, 93–96.
- [12] M. Espinoza-Moraga, T. Petta, M. Vasquez-Vasquez, V. F. Laurie, L. A. B. Moraes, L. S. Santos, *Tetrahedron: Asymmetry* **2010**, *21*, 1988–1992.
- [13] H. Li, P. Williams, J. Micklefield, J. M. Gardiner, G. Stephens, *Tetrahedron* **2004**, *60*, 753–758.
- [14] C. Reimann, H. M. Patel, L. Serino, M. Barone, C. T. Walsh, D. Haas, *J. Bacteriol.* **2001**, *183*, 813–820.
- [15] K. M. Meneely, A. L. Lamb, *Biochemistry* **2012**, *51*, 9002–9013.
- [16] W. Li, A. Khullar, S. Chou, A. Sacramo, B. Gerratana, *Appl. Environ. Microbiol.* **2009**, *75*, 2869–2878.
- [17] W. Li, S. Chou, A. Khullar, B. Gerratana, *Appl. Environ. Microbiol.* **2009**, *75*, 2958–2963.
- [18] T. Nakano, K. Miyake, H. Endo, T. Dairi, T. Mizukami, R. Katsumata, *Biosci. Biotechnol. Biochem.* **2004**, *68*, 1345–1352.
- [19] K. Mitsukura, M. Suzuki, K. Tada, T. Yoshida, T. Nagasawa, *Org. Biomol. Chem.* **2010**, *8*, 4533–4535.
- [20] K. Mitsukura, M. Suzuki, S. Shinoda, T. Kuramoto, T. Yoshida, T. Nagasawa, *Biosci. Biotechnol. Biochem.* **2011**, *75*, 1778–1782.
- [21] K. Yanai, T. Murakami, M. Bibb, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9661–9666.
- [22] T. Nagasawa, T. Yoshida, K. Ishida, H. Yamamoto, N. Kimoto (Daicel Chemical Industries, Ltd.), US 2011028749A, **2011**.
- [23] T. Nagasawa, T. Yoshida, K. Ishida, H. Yamamoto, N. Kimoto (Daicel Chemical Industries, Ltd.), US 20110262977, **2011**.
- [24] F. Long, A. A. Vagin, P. Young, G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2008**, *64*, 125–132.
- [25] J. Osipiuk, M. Zhou, A. Moy, F. Collart, A. Joachimiak, *J. Struct. Funct. Genomics* **2009**, *10*, 249–253.
- [26] L. Holm, C. Sander, *Science* **1996**, *273*, 595–560.

- [27] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774–797.
- [28] N. K. Lokanath, N. Ohshima, K. Takio, I. Shiromizu, C. Kuroishi, N. Okazaki, S. Kuramitsu, S. Yokoyama, M. Miyano, N. Kunishima, *J. Mol. Biol.* **2005**, *352*, 905–917.
- [29] M. R. Sawaya, J. Kraut, *Biochemistry* **1997**, *36*, 586–603.
- [30] C. Bystroff, S. J. Oatley, J. Kraut, *Biochemistry* **1990**, *29*, 3263–3277.
- [31] R. L. Summerfield, D. M. Daigle, S. Mayer, D. Mallik, D. W. Hughes, S. G. Jackson, M. Sulek, M. G. Organ, E. D. Brown, M. S. Junop, *J. Med. Chem.* **2006**, *49*, 6977–6986.
- [32] C. M. Czekster, A. Vandemeulebroucke, J. S. Blanchard, *Biochemistry* **2011**, *50*, 367–375.
- [33] K. E. Atkin, R. Reiss, N. J. Turner, A. M. Brzozowski, G. Grogan, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2008**, *64*, 182–185.
- [34] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.
- [35] P. Evans, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62*, 72–82.
- [36] G. Winter, *J. Appl. Crystallogr.* **2010**, *43*, 186–190.
- [37] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echou Is, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.
- [38] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- [39] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53*, 240–255.
- [40] R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
- [41] M. Hall, C. Stueckler, W. Kroutil, P. Macheroux, K. Faber, *Angew. Chem.* **2007**, *119*, 4008–4011; *Angew. Chem. Int. Ed.* **2007**, *46*, 3934–3937.

---

Received: May 20, 2013

Published online on June 28, 2013