Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

1.6 Å structure of an NAD$^+$-dependent quinate dehydrogenase from *Corynebacterium glutamicum*

Jan Schoepe,* Karsten Niefind and Dietmar Schomburg

Institute for Biochemistry, University of Cologne, Germany

Correspondence e-mail: j.schoepe@yahoo.de

Received 3 April 2008
Accepted 11 May 2008

PDB Reference: NAD$^+$-dependent quinate dehydrogenase, 2nlo, r2nlosf.

To date, three different functional classes of bacterial shikimate/quinate dehydrogenases have been identified and are referred to as AroE, SDH-L and YdiB. The enzyme AroE and the catalytically much slower SDH-L clearly prefer NADP$^+$/NADPH as the cosubstrate and are specific for (dehydro-)shikimate, whereas in YdiB the differences in affinity for NADP$^+$/NADPH *versus* NAD$^+$/NADH as well as for (dehydro-)shikimate *versus* (dehydro-)quinate are marginal. These three subclasses have a similar three-dimensional fold and hence all belong to the same structural class of proteins. In this paper, the crystal structure of an enzyme from *Corynebacterium glutamicum* is presented that clearly prefers NAD$^+$ as a cosubstrate and that demonstrates a higher catalytic efficiency for quinate rather than shikimate. While the kinetic constants for this enzyme clearly differ from those reported for AroE, SDH-L and YdiB, the three-dimensional structure of this protein is similar to members of these three subclasses. Thus, the enzyme described here belongs to a new functional class of the shikimate/quinate dehydrogenase family. The different substrate and cosubstrate specificities of this enzyme relative to all other known bacterial shikimate/quinate dehydrogenases are discussed by means of analyzing the crystal structure and derived models. It is proposed that in contrast to shikimate, quinate forms a hydrogen bond to the NAD$^+$. In addition, it is suggested that the hydroxyl group of a conserved active-site threonine hydrogen bonds to quinate more effectively than to shikimate. Also, the hydroxyl group of a conserved tyrosine approaches the carboxylate group of quinate more closely than it does the carboxylate group of shikimate. Taken together, these factors most likely lead to a lower Michaelis constant and therefore to a higher catalytic efficiency for quinate. The active site of the dehydrogenase reported here is larger than those of other known shikimate/quinate dehydrogenases, which may explain why quinate is easily accommodated within the catalytic cleft.

1. Introduction

After cellulose, lignin is the most abundant terrestrial biopolymer on the earth. Hence, the decomposition of this complex molecule is a vital component of the carbon cycle. An important degradation product of lignin is the hydroaromate quinate, which occurs as a free acid, derivatized or combined with alkaloids in amounts of ~2–10% of the dry weight of leaf litter (Dewick, 2001). The oxidation of this molecule to 3-dehydroquinate (Fig. 1*a*) is catalyzed by either pyrroloquinoline-quinone dehydrogenase (EC 1.1.99.25) or the NAD$^+$-dependent quinate dehydrogenase (QDH; EC

1.1.1.24). As early as 1911, M. Beijerinck noted several bacteria that were capable of using quinate as a sole carbon source (Beijerinck, 1911), but to date NAD^+-dependent quinate dehydrogenase activity has only been measured once for a bacterial protein. The primary structure of this specific enzyme was unknown at the time and hence its affiliation to a specific protein family remained unclear (Davis *et al.*, 1955).

The focus of this report is the NAD^+-dependent quinate dehydrogenase (*Cgl*QDH) from *Corynebacterium glutamicum*, which also displays significant affinity for shikimate. The three-dimensional fold of this enzyme is similar to those of bacterial (prokaryota and archaea) shikimate dehydrogenases, which together form a unique structural protein family, the shikimate/quinate dehydrogenase (S/QDH) family. Shikimate dehydrogenases are involved in the anabolic shikimate pathway. These enzymes catalyze the reversible reduction of dehydroshikimate to shikimate (Fig. 1*b*). Three functional subclasses of S/QDHs have been described to date in the literature. The first enzyme, AroE (EC 1.1.1.25), was discovered in *Escherichia coli*. This protein is highly specific for NADPH and shikimate (Yaniv & Gilvarg, 1955; Chaudhuri & Coggins, 1985). Orthologues of this enzyme have also been found in various other microorganisms. The second enzyme to be characterized was a shikimate dehydrogenase-like protein called SDH-L (EC 1.1.1.25) from *Haemophilus influenzae*. Strikingly, the catalytic efficiency (k_{cat}/K_M) of this enzyme was several orders of magnitude slower than that of AroE (Singh *et al.*, 2005). The third enzyme, an assumed bifunctional protein isolated from *E. coli*, was subsequently reported (YdiB; EC 1.1.1.282). It demonstrated nearly the same catalytic efficiencies for both quinate and shikimate as well as for $NADP^+$ and NAD^+ (Benach *et al.*, 2003).

Here, we present both enzyme-kinetic studies and a high-resolution X-ray structural analysis of the NAD^+-dependent quinate dehydrogenase from *C. glutamicum*. A bound glycerol in the active site facilitated modelling studies for substrate binding. Owing to their absence in mammalian cells, enzymes of the shikimate pathway, and in particular the shikimate dehydrogenases, with their highly conserved active-site architectures, are potential targets for drug design (Herrmann & Weaver, 1999; Lindner *et al.*, 2005). The altered substrate specificities of *Cgl*QDH reported here provide detailed information for the design of potentially new inhibitors directed at the shikimate dehydrogenases.

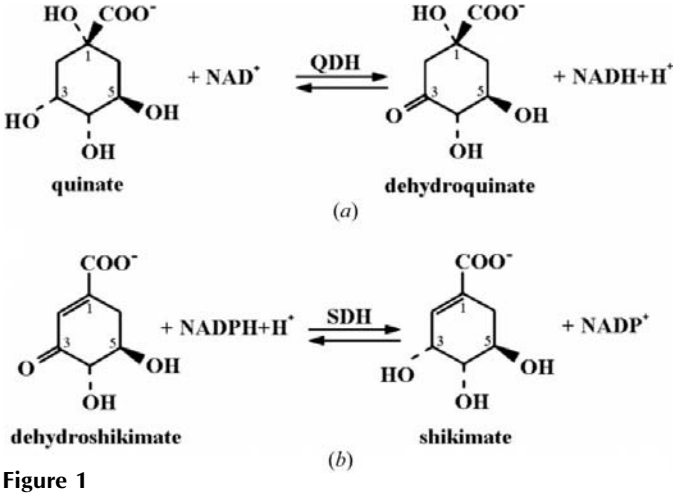

Figure 1
The oxidation of quinate to dehydroquinate by quinate dehydrogenase (QDH) is shown in (*a*). The reduction of dehydroshikimate to shikimate by shikimate dehydrogenase (SDH) is shown in (*b*).

Table 1
Summary of the refinement statistics for *Cgl*QDH.

Space group	*C*2
Resolution range (Å)	7.99–1.64
R factor (%)	14.8
R_{free} (%)	16.5
No. of unique reflections	29404
Missing residues	1–21
No. of protein atoms	2067
No. of water molecules	255
No. of ligand atoms	6
Average *B* values (Å^2)	
Protein	18.3
Solvent	30.4
Ligand	22.8
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.135
Ramachandran plot	
Most favoured (%)	95
Additional favoured (%)	5

2. Experimental procedures

The open reading frame *Cgl*0424 (nomenclature by Ikeda & Nakagawa, 2003) was cloned and overexpressed. The gene product was purified and crystallized and a native X-ray data set was collected as previously described (Schoepe *et al.*, 2006).

2.1. Phasing

The phase problem was solved by the molecular-replacement method. The structure of shikimate dehydrogenase from *Methanococcus jannaschii* (PDB code 1nvt; Padyana & Burley, 2003) was used as a search model. All residues apart from glycines were changed to alanines. The Patterson search was performed using *CNS* (Brünger *et al.*, 1998) in the a resolution range 15.0–6.0 Å. The rotation search was performed in real space. Despite the relatively low amino-acid sequence identity between the search model and *Cgl*QDH (∼32%), the correct solution appeared clearly at the top of the list of solutions after the translation search. The corresponding correlation coefficient between observed and calculated structure-factor amplitudes was 0.235 and had a significant offset from the second (0.185) and third solutions (0.162).

2.2. Model building, refinement and quality of the crystal structure

The model was built using the program *ARP*/*wARP* (Perrakis *et al.*, 2001). Completion of the structure required iterative cycles of refinement in *REFMAC*5 (Murshudov *et al.*, 1997) and manual rebuilding with the graphics program *O* (Jones *et al.*, 1991). The crystal structure was validated by means of a Ramachandran plot with the program *PRO-*

CHECK (Laskowski *et al.*, 1993). The refinement statistics are listed in Table 1.

2.3. Modelling the ternary complexes

The glycerol was deleted from the model and manually substituted with quinate or shikimate according to the criteria described below. Both complexes were refined with *REFMAC*5 using ten cycles each. The X-ray term of the program was kept in order to conserve the structure of the enzyme. Thus, only the distances of the ligands to the atoms of the protein were optimized. To obtain the ternary complexes, the resulting models were structurally aligned with PDB entry 1nvt, which contains an NADPH molecule. The additional phosphoryl group of the NADPH was deleted to obtain NADH. The ternary complexes were again refined with *REFMAC5* as above.

The side chains of Thr88 and Thr221 of the ternary complex with quinate were rotated manually with the program *O* to optimize the distances to the quinate and to the NADH, respectively.

2.4. Figure preparation

All figures showing molecular structures were generated with either *PyMOL* (DeLano, 2002) or *BobScript* (Esnouf, 1999).

2.5. Biochemical characterization

After overexpression, the protein was N-terminally His_6-tagged. To ensure that the tag influenced neither the enzymatic activity nor the oligomeric state, it was cleaved with a derivative of the site-specific rTEV protease (Carrington & Dougherty, 1988) according to the protocol of Chatterjee *et al.* (2005). The dialysis buffer for refolding of the rTEV contained 100 m*M* Tris–HCl pH 8.5, 500 m*M* NaCl, 5 m*M* DTT, 0.5 m*M* EDTA and 50%(*v*/*v*) glycerol. Mass-spectrometric analysis showed that after cleavage the cloning artefacts Gly-Ala remained at the N-terminus (molecular weight 29.8 kDa).

To evaluate the oligomeric state of *Cgl*QDH, size-exclusion chromatography was accomplished with a HiLoad 16/60 Superdex 200 prep-grade column (Amersham Biosciences) using 50 m*M* Tris–HCl pH 7.5 buffer and about 2 mg protein diluted in 50 m*M* Tris–HCl, 100 m*M* NaCl pH 7.5. Calibration was performed using the reference protein mixture recommended by Amersham Biosciences.

The enzymatic activity was assayed at 303 K by monitoring the reduction of $NAD(P)^+$ to NAD(P)H at 340 nm (ε = 6220 M^{-1} cm^{-1}) in a 1 ml reaction volume. According to the Bradford method (Bradford, 1976) with BSA as the standard, the final enzyme concentration was 5 n*M* and the concentrations of quinate and shikimate were chosen to be between 0.4 and 0.8 m*M*. Initial steady-state rates were calculated from the linear portion of the reaction curve. Because of the observed substrate inhibition, the concentrations of the substrates were chosen to be below the Michaelis constant of the enzymes. Hence, the maximum velocity of the enzyme (V_{max}) had to be calculated. To measure the enzyme activity with quinate, 500 m*M* Tris–HCl buffer pH 9.0 was used. The measurements with shikimate were carried out in 100 m*M* sodium carbonate buffer pH 10.0. Given the fact that the enzyme has a slightly higher turnover number (k_{cat}) in Tris–HCl buffer, a correction factor of 1.6 was applied to correct for this error. The K_M value was not influenced by changing the buffer system.

Table 2
Enzyme-kinetic constants for *Cgl*QDH with NAD^+ as cosubstrate.

Substrate	k_{cat} (s^{-1})	K_M (m*M*)	k_{cat}/K_M (s^{-1} mM^{-1})
Quinate	61.9 ± 4.2	10.2 ± 2.0	6.1
Shikimate	85.2 ± 15.6	46.6 ± 10.7	1.8

3. Results and discussion

3.1. Enzymatic characterization of *C. glutamicum* QDH

Enzymatic assays show that *Cgl*QDH not only has affinity for quinate but also for shikimate. Under optimal *in vitro* conditions, the enzyme shows the highest k_{cat} value with shikimate (Table 2). On the other hand, the Michaelis constant with quinate as a substrate is significantly lower than with shikimate, resulting in a catalytic efficiency that is more than three times higher with quinate than with shikimate.

Moreover, with $NADP^+$ instead of NAD^+ as cosubstrate, the turnover number decreases by more than 300-fold with either shikimate or quinate as a substrate. Hence, we conclude that *Cgl*QDH plays a key role in the quinate-degradation pathway.

3.2. Homodimeric structure of *C. glutamicum* QDH

Prokaryotic members of the S/QDH family exist either as monomers or, more often, as homodimers in their active forms. In the latter case, the N-terminal domain is responsible for dimerization. According to hydrophobicity plots (Kyte & Doolittle, 1982), the corresponding residues of the potential 'dimerization domain' of *Cgl*QDH are on average at least as hydrophobic as those of the dimeric enzymes AroE and YdiB from *E. coli* and AroE from *M. jannaschii*. Gel-filtration studies performed on *Cgl*QDH with or without the His_6 tag indicate that this protein is indeed a homodimer in solution. According to the program *PISA* (Krissinel & Henrick, 2005), the dimer buries ∼1200 Å^2 of the solvent-accessible surface, which is at the lower end of the spectrum for 'standard-size' protein–protein interfaces (Lo Conte *et al.*, 1999).

3.3. Tertiary and supersecondary structure of *C. glutamicum* QDH

The asymmetric unit contains one subunit with 281 residues and 255 water molecules. Residues 1–21 are disordered and therefore were not included in the model. These residues contain the His_6 tag, the rTev cleavage site, spacer amino acids and the first two amino acids of the native protein. *Cgl*QDH is composed of two $\alpha\beta\alpha$ domains (Fig. 2). These domains contain two discontinuous segments, Asp22–Asn127 and Gly287–Leu302. The N-terminal domain is formed by six β-strands

with strand order $\beta2$-$\beta1$-$\beta3$-$\beta5$-$\beta6$-$\beta4$ and $\beta5$ running anti-parallel to the others. This mixed β-sheet is surrounded by four α-helices, $\alpha5$, $\alpha6$, $\alpha7$ and $\alpha8$, as well as a 3_{10}-helix. The N-terminal domain forms an open $\alpha\beta\alpha$ sandwich, which is characteristic of enzymes of the S/QDH family but different from all other known proteins (Michel *et al.*, 2003; Ye *et al.*, 2003).

The C-terminal domain or cosubstrate-binding domain consists of all residues from Leu141 to Gly270. Six parallel but twisted β-strands with strand order $\beta9$-$\beta8$-$\beta7$-$\beta10$-$\beta11$-$\beta12$ are flanked by four α-helices, $\alpha5$, $\alpha6$, $\alpha7$ and $\alpha8$, and another 3_{10}-helix. This arrangement of secondary-structural elements represents the classical Rossmann or dinucleotide-binding fold (Holm & Sander, 1996).

The N- and C-terminal domains are linked together by α-helices 4 and 9, which form the bottom of a groove approximately 20 Å deep. The active site of the protein is located between the two domains.

3.4. Sequence and structural similarities to other bacterial shikimate/quinate dehydrogenases

To date, eight three-dimensional structures of S/QDH-family members have been determined and are available from the RCSB Protein Data Bank, including all four known subtypes. These are AroEs from *E. coli* (Michel *et al.*, 2003), *H. influenzae* (Ye *et al.*, 2003), *Thermus thermophilus* (Bagautdinov & Kunishima, 2007), *Aquifex aeolicus* (Gan *et al.*, 2007) and *M. jannaschii* (Padyana & Burley, 2003), Ydib from *E. coli* (Benach *et al.*, 2003; Michel *et al.*, 2003), SDH-L from *H. influenzae* (Singh *et al.*, 2005) and finally *Cgl*QDH (this work).

Comparisons of amino-acid sequence identities between these proteins as well as structural superpositions reveal that their three-dimensional architectures are highly conserved, in contrast to their primary sequences. Indeed, the superimposed protein structures show that they all have nearly identical $\alpha\beta\alpha$ domains with the same arrangements of β-strands and a deep cleft separating the N- and C-terminal domains. Conserved residue side chains extend into the deep cleft of the enzymes, where the active sites are presumably located (Laskowski *et al.*, 1996).

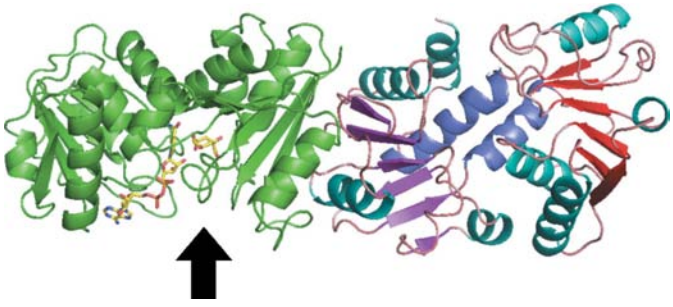

Figure 2
Homodimeric structure of *Cgl*QDH. One subunit is shown in green. The arrow marks the deep cleft between the N- and C-terminal domains. The N-terminal α-helices and β-strands of the second subunit are shown in cyan and magenta, whereas those of the C-terminal domain are coloured blue and red, respectively. The two α-helices that bridge the N- and C-terminal domains of the second subunit are displayed in slate. The order of the β-strands in the N-terminal domain is specific for proteins of the S/QDH family.

3.5. Potential binding site of NADH

The positions of the cosubstrates are similar in all the different S/QDHs (see, for example, Ye *et al.*, 2003). Thus, it was possible to position NADH into the *Cgl*QDH model with reasonable accuracy.

Depending on the substrate chosen, the enzymatic pH optimum of *Cgl*QDH was approximately between 9 and 10. Hence, we regarded all carboxylic groups in our studies as completely deprotonated to carboxylate groups.

On the basis of this modelling, several observations regarding *Cgl*QDH can be made. Firstly, the characteristic glycine-rich loop with general sequence G*X*/SGG*X*/A found in S/QDHs is present in *Cgl*QDH. This loop anchors the diphosphate of the cosubstrate and arranges a bend in the dinucleotide that is rather unusual for NAD(P)-dependent dehydrogenases (Ye *et al.*, 2003). Cosubstrates of the S/QDHs demonstrate an *anti* conformation for the nicotinamide ring and are A (pro-*R*) site stereospecific (see, for example, Benach *et al.*, 2003). The backbone O atom of the highly conserved Gly270 plays a role in maintaining the *anti* conformation of the nicotinamide ring by hydrogen bonding to the amide group of NADH. A hydrogen bond between the hydroxyl group of the conserved Thr221 and the oxygen in the adenosyl ribose ring as well as a weak hydrogen bond between the thiol group of the 'semiconserved' Met225 and the 2′-OH and 3′-OH of the ribose of the nicotinamide mononucleotide (NMN) can also be inferred. The distance for hydride transfer between C3 of the metabolite and C4 of the nicotinamide ring should be within the range 3–5 Å. In our model, this distance is 3.3 Å for quinate and 3.8 Å for shikimate. Furthermore, it is well known that in $NAD(P)^+$/NAD(P)H-dependent dehydrogenases only a few residues are responsible for cosubstrate recognition, namely arginine and threonine or serine for $NAD(P)^+$/NADPH-dependent enzymes and aspartate for NAD^+/NADH-dependent proteins (see, for example, Luyten *et al.*, 1989). The primary structure of *Cgl*QDH contains an Asp177-*X*-Asp179 motif. Modelling the NADH into the protein structure shows that Asp177 binds the 2′- and 3′-hydroxyl groups of the adenosylribose and thus is partially responsible for dinucleotide recognition. In the case of NADPH binding, there would be van der Waals as well as Coulombic repulsion of the negative charged phosphoryl group of the NADPH with the carboxylate group of Asp177.

3.6. Potential binding site of quinate and shikimate

Electron density corresponding to a glycerol molecule was found in the *Cgl*QDH active site (Fig. 3). Two of the hydroxyl groups of the ligand are hydrogen bonded to the completely conserved residues Lys92 and Asp102. These two residues are assumed to be essential for catalytic activity in members of the S/QDH family (Lindner *et al.*, 2005). Moreover, the third

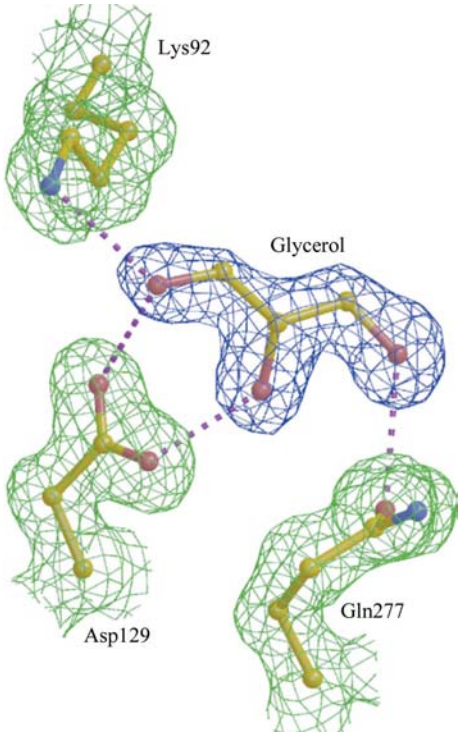

Figure 3
$2F_o - F_c$ electron densities for the glycerol (contoured at 1.5σ) and the three conserved amino-acid side chains of *Cgl*QDH which form hydrogen bonds to the ligand.

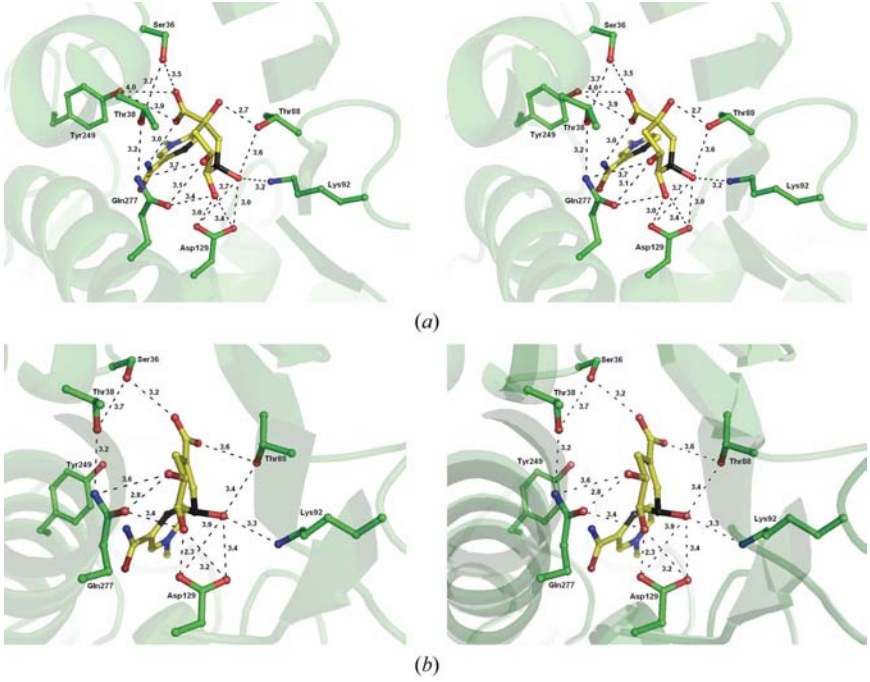

Figure 4
Stereoviews of the potential substrate-binding mode with quinate (*a*) or shikimate (*b*). The side chains of the active-site residues are shown in green; the substrates and the nicotinamide ring of the NAD^+ are coloured yellow. The C atoms C3-quinate (*a*), C3-shikimate (*b*) and C4-NAD^+ (both) where oxidation/reduction occurs are shown in black. The distances between these C atoms amount to 3.3 Å for quinate binding and 3.8 Å for shikimate binding, respectively. Dashed lines represent distances between the corresponding atoms of up to 4 Å, but do not necessarily imply the presence of a hydrogen bond.

remaining hydroxyl group of glycerol interacts with Gln277. This conserved residue is also part of the active site.

We propose that the three hydroxyl groups at the C3, C4 and C5 positions (Fig. 1) of the substrates are bound to the enzyme in a similar manner. Figs. 4(*a*) and 4(*b*) show the results of our modelling. Both quinate and shikimate are located in similar positions and surrounded by the same conserved residues (Fig. 4). The locations of shikimates bound in the active sites of the AroEs from *T. thermophilus* (Bagautdinov & Kunishima, 2007) and *A. aeolicus* (Gan *et al.*, 2007) are similar to that shown in Fig. 4(*b*).

According to our model, in the case of quinate binding the NAD^+ is closer to the carboxylate group of the substrate than in the case of shikimate, resulting in a hydrogen bond between an O atom of the carboxylate group of the quinate and a hydrogen of the amino group of the nicotinamide ring (3.0 *versus* 5.8 Å). In addition, the side chain of Thr88 is freely rotatable. In the ideal conformation, the hydroxyl group of this amino acid interacts with the additional C1 hydroxyl group of quinate much more closely (2.7 Å) than it does with the carboxylate group of shikimate (3.6 Å). Also, as a result of the tetrahedral C1 atom of quinate (Fig. 1), the hydroxyl group of the conserved Tyr249 is located much closer to the two O atoms of the carboxylate group of the ligand (3.9 and 4.0 Å) compared with that observed with shikimate bound (5.5 and 6.1 Å). Taking into account substrate movement, the plausibility of building at least temporal hydrogen bonds with Tyr249 is much higher with quinate than with shikimate. Hence, the stronger binding of quinate compared with shikimate could result in a lower K_M value for quinate (Table 2).

Substrate specificity in members of the S/QDH family generally depends upon the size of their respective active sites. Quinate, for example, requires more space to bind in the active site than shikimate does owing to its additional hydroxyl group (Fig. 1). Also, in *Cgl*QDH the side chain of Thr88 extends more deeply into the active-site pocket than observed for other S/QDH-family members, with the net effect of decreasing the size of the catalytic cleft (Fig. 5). Moreover, in *Cgl*QDH there is another residue across from Thr88, namely Ala274. In other S/QDH-family members, this position is generally occupied by larger side chains such as leucine or phenylalanine. Importantly, the crystal structures of the AroEs from *T. thermophilus* (Bagautdinov & Kunishima, 2007) and *A. aeolicus* (Gan *et al.*, 2007) with and without shikimate show that substrate binding does not significantly change the distance between these two space-filling residues.

Product-inhibition as well as isotope-exchange studies have shown that the

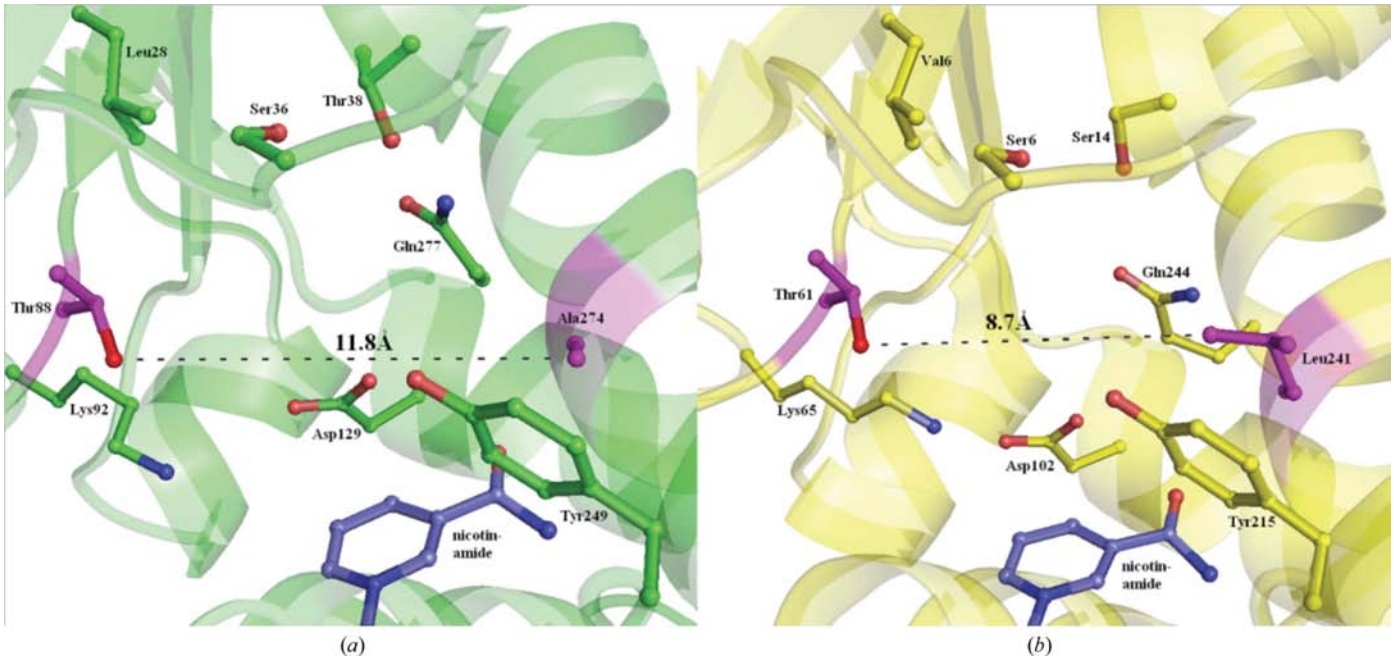

Figure 5
Possible reasons for the different substrate specificities of *Cgl*QDH (*a*) and AroE from *E. coli* (*b*). Leu241 in AroE restricts the space of the active site more than its analogue Ala274 in *Cgl*QDH. Furthermore, the corresponding distances are on average 9.9 Å in YdiB and 9.6 Å in SDH-L.

mechanism of the shikimate dehydrogenase in *Pisum sativum* is an ordered bi-bi mechanism (Balinsky *et al.*, 1971; Cleland, 1963). Moreover, kinetic experiments demonstrate that in SDH-L from *H. influenzae* a functional group with a pK_{acid} of ∼8 is involved in the catalytic reaction, which needs to be deprotonated for activity (Singh *et al.*, 2005). We assume that this is also the case for *Cgl*QDH owing to its similar pH rate profiles (not shown). On the basis of our modelling of quinate into the active site of *Cgl*QDH, we predict that the active-site base is Lys92, which is completely conserved in the active site of all S/QDHs. The deprotonated form of Lys92 would serve as an excellent proton acceptor. Therefore, under the assumption that quinate assumes a chair-like conformation in order to bind *Cgl*QDH, we propose the following reaction mechanism. After the binding of quinate and NAD^+, the oxygen of the C3-OH of quinate forms a hydrogen bond to the side chain of the conserved Thr88 and Lys92 functions as the active-site base to remove the proton on the C3-OH. Simultaneously, a hydride is transferred from C3 of quinate to NAD^+.

This work was funded by the Bundesministerium für Bildung und Forschung (BMBF), Berlin, Germany. The assistance of the staff of the EMBL Outstation, Hamburg, Germany is gratefully acknowledged.

References

Bagautdinov, B. & Kunishima, N. (2007). *J. Mol. Biol.* **373**, 424–438.
Balinsky, D., Dennis, A. W. & Cleland, W. W. (1971). *Biochemistry*, **10**, 1947–1952.
Beijerinck, M. W. (1911). *Proc. R. Acad. Sci. (Amsterdam)*, **13**, 1066–1077.
Benach, J., Lee, I., Edstrom, W., Kuzin, A. P., Chiangi, Y., Acton, T. B., Montelione, G. T. & Hunt, J. F. (2003). *J. Biol. Chem.* **278**, 19176–19182.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Carrington, J. C. & Dougherty, W. G. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 3391–3395.
Chatterjee, S., Schoepe, J., Lohmer, S. & Schomburg, D. (2005). *Protein Expr. Purif.* **39**, 137–143.
Chaudhuri, S. & Coggins, J. R. (1985). *Biochem. J.* **226**, 217–223.
Cleland, W. W. (1963). *Biochim. Biophys. Acta*, **67**, 104–137.
Davis, B. D., Gilvarg, C. & Mitsuhashi, S. (1955). *Methods Enzymol.* **2**, 300–311.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, USA.
Dewick, P. M. (2001). *Medicial Natural Products – A Biosynthetic Approach*, 2nd ed. Chichester: John Wiley & Sons.
Esnouf, R. M. (1999). *Acta Cryst.* D**55**, 938–940.
Gan, J., Wu, Y., Prabakaran, P., Gu, Y., Li, Y., Andrykovitch, M., Liu, H., Gong, Y., Yan, H. & Ji, X. (2007). *Biochemistry*, **46**, 9513–9522.
Herrmann, K. M. & Weaver, L. (1999). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **25**, 473–503.
Holm, L. & Sander, C. (1996). *Science*, **273**, 595–602.
Ikeda, M. & Nakagawa, S. (2003). *Appl. Microbiol. Biotechnol.* **62**, 99–109.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Krissinel, E. & Henrick, K. (2005). *Detection of Protein Assemblies in Crystals.* Berlin/Heidelberg: Springer-Verlag.
Kyte, J. & Doolittle, R. F. (1982). *J. Mol. Biol.* **157**, 105–132.
Laskowski, R. A., Luscombe, N. M., Swindells, M. B. & Thornton, J. M. (1996). *Protein Sci.* **5**, 2438–2452.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lindner, H. A., Nadeau, G., Matte, A., Michel, G., Ménard, R. & Cygler, M. (2005). *J. Biol. Chem.* **280**, 7162–7169.
Lo Conte, L., Chothia, C. & Janin, J. (1999). *J. Mol. Biol.* **285**, 2177–2198.
Luyten, M., Bur, D., Wynn, H., Parris, W., Gold, M., Friesen, J. D. & Jones, J. B. (1989). *J. Am. Chem. Soc.* **111**, 6800–6804.
Michel, G., Roszak, A., Sauve, V., Maclean, J., Matte, A., Coggins, J. R., Cygler, M. & Lapthorn, A. J. (2003). *J. Biol. Chem.* **278**, 19463–19472.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Padyana, A. K. & Burley, S. (2003). *Structure*, **11**, 1005–1013.
Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). *Acta Cryst.* D**57**, 1445–1450.
Schoepe, J., Niefind, K., Chatterjee, S. & Schomburg, D. (2006). *Acta Cryst.* F**62**, 635–637.
Singh, S., Korolev, S., Koroleva, O., Zarembinski, T., Collart, F., Joachimiak, A. & Christendat, D. (2005). *J. Biol. Chem.* **280**, 17101–17108.
Yaniv, H. & Gilvarg, C. (1955). *J. Biol. Chem.* **213**, 787–795.
Ye, S., Delft, F., Brooun, A., Knuth, M. W., Swanson, R. V. & McRee, D. E. (2003). *J. Bacteriol.* **185**, 4144–4151.