Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# The structure of *Sulfolobus solfataricus* 2-keto-3-deoxygluconate kinase

**Jane A. Potter,[a] Melina Kerou,[a] Henry J. Lamble,[b] Steven D. Bull,[c] David W. Hough,[b] Michael J. Danson[b] and Garry L. Taylor[a]***

[a]Centre for Biomolecular Sciences, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, Scotland, [b]Centre for Extremophile Research, Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, England, and [c]Department of Chemistry, University of Bath, Claverton Down, Bath BA2 7AY, England

Correspondence e-mail: glt2@st-andrews.ac.uk

Received 4 September 2008
Accepted 4 November 2008

**PDB References:** 2-keto-3-deoxygluconate kinase, apo, 2v78, r2v78sf; ternary complex, 2var, r2varsf.

The hyperthermophilic archaeon *Sulfolobus solfataricus* grows optimally above 353 K and utilizes an unusual promiscuous nonphosphorylative Entner–Doudoroff pathway to metabolize both glucose and galactose. It has been proposed that a part-phosphorylative Entner–Doudoroff pathway occurs in parallel in *S. solfataricus*, in which the 2-keto-3-deoxygluconate kinase (KDGK) is promiscuous for both glucose and galactose metabolism. Recombinant *S. solfataricus* KDGK protein was expressed in *Escherichia coli*, purified and crystallized in 0.1 *M* sodium acetate pH 4.1 and 1.4 *M* NaCl. The crystal structure of apo *S. solfataricus* KDGK was solved by molecular replacement to a resolution of 2.0 Å and a ternary complex with 2-keto-3-deoxygluconate (KDGlu) and an ATP analogue was resolved at 2.1 Å. The complex suggests that the structural basis for the enzyme's ability to phosphorylate KDGlu and 2-keto-3-deoxygalactonate (KDGal) is derived from a subtle repositioning of residues that are conserved in homologous nonpromiscuous kinases.

## 1. Introduction

The hyperthermophilic archaeon *Sulfolobus solfataricus* is an obligate aerobe that is able to grow either chemolithotrophically, oxidizing sulfur to sulfuric acid and fixing carbon dioxide, or chemoorganotrophically, oxidizing various organic compounds including $\alpha$- and $\beta$-linked glucose saccharides, glucose, galactose, pentose sugars and peptide-containing substrates. Its optimum growth conditions are 353–358 K and pH 2–4, maintaining an intracellular pH of 5–6. Its natural habitats are terrestrial volcanic hot springs (Grogan, 1989). The complete genome of *S. solfataricus* has been sequenced and published (She *et al.*, 2001), rendering it a model organism for physiological, biochemical and genetic studies of archaea and of thermoacidophiles in particular. Central carbohydrate metabolism in archaea is of particular interest as they utilize modifications of the classical Embden–Meyerhof (EM) and Entner–Doudoroff (ED) pathways for glucose degradation and a series of novel enzymes and enzyme families have been identified (reviewed by Danson *et al.*, 2007).

Sugar metabolism in *S. solfataricus* occurs through two variants of the ED pathway which proceed with no net production of ATP. In the nonphosphorylative ED pathway (DeRosa *et al.*, 1984), the oxidation of glucose to gluconate and its subsequent dehydration to 2-keto-3-deoxygluconate (KDGlu) is catalysed by glucose dehydrogenase and gluconate dehydratase, respectively. KDGlu is then cleaved by 2-keto-3-deoxygluconate aldolase (KDG aldolase) to produce glyceraldehyde and pyruvate. Glyceraldehyde is oxidized to

glycerate, probably by a glyceraldehyde dehydrogenase, and this is followed by phosphorylation of glycerate to 2-phosphoglycerate by glycerate kinase. Conversion of 2-phosphoglycerate to phosphoenolpyruvate and finally to a second molecule of pyruvate is achieved by the actions of enolase and pyruvate kinase. In the semi-phosphorylative ED pathway (Ahmed *et al.*, 2005), KDGlu is phosphorylated by 2-keto-3-deoxygluconate kinase (KDGK) to form KDPGlu (2-keto-3-deoxy-6-phosphogluconate), which is followed by aldol cleavage of KDPGlu to pyruvate and glyceraldehyde-3-phosphate (GAP) by KDG aldolase (bifunctional KDG/KDPG aldolase). The latter metabolite is then oxidized to 3-phosphoglycerate by an $NAD(P)^+$-dependent nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase and converted to 2-phosphoglycerate by phosphoglycerate mutase; finally, a second molecule of pyruvate is produced as described previously. This pathway has also been found to operate in the hyperthermophilic archaeon *Thermoproteus tenax* (Ahmed *et al.*, 2005).

Both pathways in *S. solfataricus* were found to be promiscuous for the metabolism of both glucose and galactose (Lamble *et al.*, 2003, 2004, 2005), which implies that the pathway enzymes exhibit catalytic and stereochemical promiscuity whilst maintaining their activity and stability at extreme temperatures. Our previous studies have elucidated the structural basis of substrate promiscuity for *S. solfataricus* glucose dehydrogenase (Milburn *et al.*, 2006) and KDG aldolase (Theodossis *et al.*, 2004).

KDG kinase is a key enzyme in the semi-phosphorylative variant of the ED pathway. The gene encoding the *S. solfataricus* KDG kinase (*Ss*KDGK) was initially identified by homology searches (Ahmed *et al.*, 2005; Lamble *et al.*, 2005). The *Ss*KDGK gene has been cloned and expressed and the recombinant enzyme has been purified, biochemically characterized (Lamble *et al.*, 2005) and found to follow Michaelis–Menten kinetics, with optimal activity between temperatures of 343 and 353 K and pH values of 7 and 8 (Kim & Lee, 2006). *Ss*KDGK exhibits a similar catalytic efficiency for KDGlu and 2-keto-3-deoxygalactonate (KDGal), suggesting that both sugars are natural substrates for the enzyme *in vivo* (Lamble *et al.*, 2005). The enzyme also requires divalent cations, preferably $Mg^{2+}$, for activity.

*Ss*KDGK belongs to the PfkB family of carbohydrate kinases, which also includes ribokinase, adenosine kinase, *Escherichia coli* 6-phosphofructokinase and archaeal ADP-dependent glucokinases and phosphofructokinases. Members of this family share a high degree of structural homology, as well as a common catalytic mechanism in which a conserved aspartate residue mediates the transfer of the ATP phosphate group to a substrate hydroxyl group.

In this study, we present the crystal structure of *Ss*KDGK in its apo form to a resolution of 2.0 Å and as a ternary complex with KDGlu and 5′-adenyl-$\beta,\gamma$-imidodiphosphate (AMP-PNP) to a resolution of 2.1 Å. Comparison with homologous but nonpromiscuous KDG kinases suggests structural reasons for the observed ability of *Ss*KDGK to phosphorylate KDGlu or KDGal.

**Table 1**
X-ray data collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

| | Apo | Ternary complex |
|---|---|---|
| Space group | $P6_522$ | $P6_522$ |
| Unit-cell parameters (Å) | $a = b = 104.9$, $c = 422.5$ | $a = b = 104.5$, $c = 420.8$ |
| X-ray source | ID14-1 | ID14-2 |
| Wavelength (Å) | 0.934 | 0.934 |
| Resolution range (Å) | 76.3–2.0 | 90.5–2.1 |
| No. of unique observations | 91205 | 80339 |
| Completeness (%) | 97 (81.4) | 99.8 (99.5) |
| Redundancy | 12.8 (3.9) | 6.2 (4.2) |
| $R_{\rm merge}$† | 0.063 (0.210) | 0.084 (0.309) |
| $\langle I/\sigma(I)\rangle$ | 26.8 (7.2) | 15.7 (4.0) |
| Refinement | | |
| No. of reflections | | |
| Working | 86496 | 80626 |
| Test | 4554 | 4008 |
| No. of protein atoms | 7636 | 7373 |
| No. of ligand atoms | 0 | 234 |
| No. of waters | 462 | 457 |
| Average $B$ factors (Å$^2$) | | |
| Protein | 21 | 23 |
| Ligand | — | 34 |
| Waters | 28 | 32 |
| $R_{\rm cryst}$‡ | 0.162 | 0.183 |
| $R_{\rm free}$‡ | 0.198 | 0.218 |
| R.m.s.d. bond distances (Å) | 0.020 | 0.017 |
| R.m.s.d. bond angles (°) | 1.79 | 1.70 |

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{\rm cryst}$ and $R_{\rm free} = \sum ||F_{\rm o}| - |F_{\rm c}|| / \sum |F_{\rm o}|$.

## 2. Materials and methods

### 2.1. Expression, purification and crystallization

*Ss*KDGK was expressed and purified as described previously, but with an additional heat-precipitation step (338 K, 10 min) prior to nickel-affinity column purification (Lamble *et al.*, 2005). For crystallization, the protein was dialysed into 20 m*M* Tris–HCl pH 8.5, 100 m*M* NaCl and concentrated to 3 mg ml$^{-1}$. The sitting-drop vapour-diffusion method was used to produce crystals. Hexagonal rod-shaped crystals appeared after 3 d at 293 K using a crystallization buffer consisting of 0.1 *M* sodium acetate pH 4.1 and 1.4 *M* NaCl. Crystals of a ternary complex of *Ss*KDGK were obtained by soaking crystals in mother liquor containing 20 m*M* KDGlu, 5 m*M* AMP-PNP (Sigma–Aldrich) and 5 m*M* $MgCl_2$. Crystals were frozen in a nitrogen-gas stream after being soaked in 10%(*v*/*v*) followed by 20%(*v*/*v*) glycerol in mother liquor for 30 s each. The crystals belonged to space group $P6_122$ or $P6_522$, with unit-cell parameters $a = b = 104.9$, $c = 422.5$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$.

### 2.2. Data collection and processing

For the apo structure, X-ray diffraction data were collected to 2.0 Å resolution on beamline ID14-1 of the European Synchrotron Radiation Facility (ESRF), Grenoble. Data for the ligand-bound structure were collected to 2.1 Å resolution on beamline ID14-2. Although the crystals diffracted further than these resolution limits, extending data to a higher resolution would have resulted in spatial overlap of reflections

owing to the large unit cell. The data were processed using *MOSFLM* (Leslie, 1992) and scaled with the program *SCALA* in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

### 2.3. Structure solution and refinement

The structure of the apo form of KDGK was determined by molecular replacement using *Phaser* (McCoy *et al.*, 2005) with a monomer of *S. tokodaii* KDGK (*St*KDGK) as the search model (PDB code 2dcn) and the space group was confirmed as $P6_522$. Three monomers were found in the asymmetric unit. Refinement and model building were performed using restrained maximum-likelihood refinement in *REFMAC* v.5.2 (Murshudov *et al.*, 1997) and *Coot* (Emsley & Cowtan, 2004). The data-collection and refinement statistics are given in Table 1 and the structures of the apo and ternary complex structures have been deposited in the Protein Data Bank under accession codes 2v78 and 2var, respectively.

## 3. Results and discussion

### 3.1. Structure of *Ss*KDGK

The asymmetric unit of the crystal contains three monomers, with the hexamer formed through the operation of a crystallographic twofold axis (Fig. 1*a*). Each monomer has a major $\alpha/\beta$ domain composed of 11 $\alpha$-helices positioned around a central $\beta$-sheet formed by eight parallel and one antiparallel $\beta$-strands, although differences are observed among the chains in the number and length of the $\beta$-strands. The lid subdomain comprises an antiparallel four-stranded $\beta$-sheet and a small $3_{10}$-helix. The monomer of *Ss*KDGK has an r.m.s.d. of 0.8 Å for 307 $C^{\alpha}$ atoms with *St*KDGK, with which it shares 57% sequence identity. Dimerization occurs *via* the lid subdomains as seen in *St*KDGK.

Structural homologues of *Ss*KDGK were identified using the *DALI* server. Significant matches include the KDG kinase from *Thermus thermophilus* (*Tt*KDGK), with which *Ss*KDGK shares 34% sequence identity (Ohshima *et al.*, 2004). The structure of a ternary complex of *Tt*KDGK with KDGlu and AMP-PNP showed that the KDGlu was bound predominantly in the open-chain form (Ohshima *et al.*, 2004). Several related structures have been deposited in the PDB but have not yet been described in publications. These include the apo structure of *S. tokodaii* KDG kinase (*St*KDGK) to a resolution of 2.8 Å (PDB code 1wye), a complex of *St*KDGK with the $\alpha$-furanose form of 2-keto-6-phosphogluconate and ADP to a resolution of 2.25 Å (PDB code 2dcn) and a putative 5-dehydro-2-deoxygluconokinase from *Bacillus hydurans* (PDB code 2qcv).

### 3.2. KDGlu binding

The active site is located in a tight deep cleft formed by the main $\alpha/\beta$ domain and the lid subdomain (Fig. 1*b*). A difference electron-density map clearly shows the binding of AMP-PNP and KDGlu in the active site (Fig. 2*a*), with both the open-chain and $\beta$-furanose forms of KDGlu fitting the electron density (Figs. 2*b* and 2*c*). Slight differences in hydrogen-bond lengths and the conformation of the KDGlu exist between the three monomers in the asymmetric unit, but no additional interactions were observed when comparing the open-chain and furanose forms of the bound substrate. The carboxylate group of KDGlu is positioned within hydrogen-bonding distance of Tyr106 $O^{\eta}$ (2.5 Å in the open-chain and 2.8 Å in

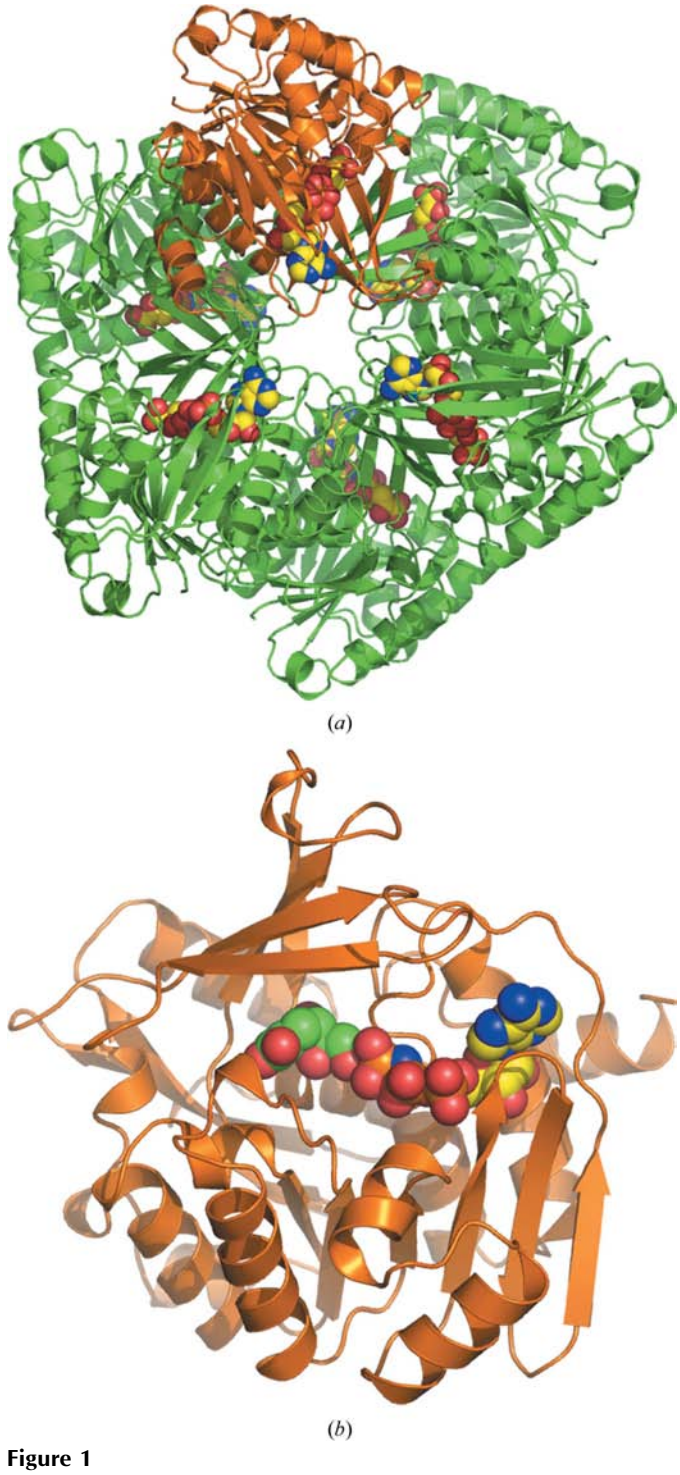
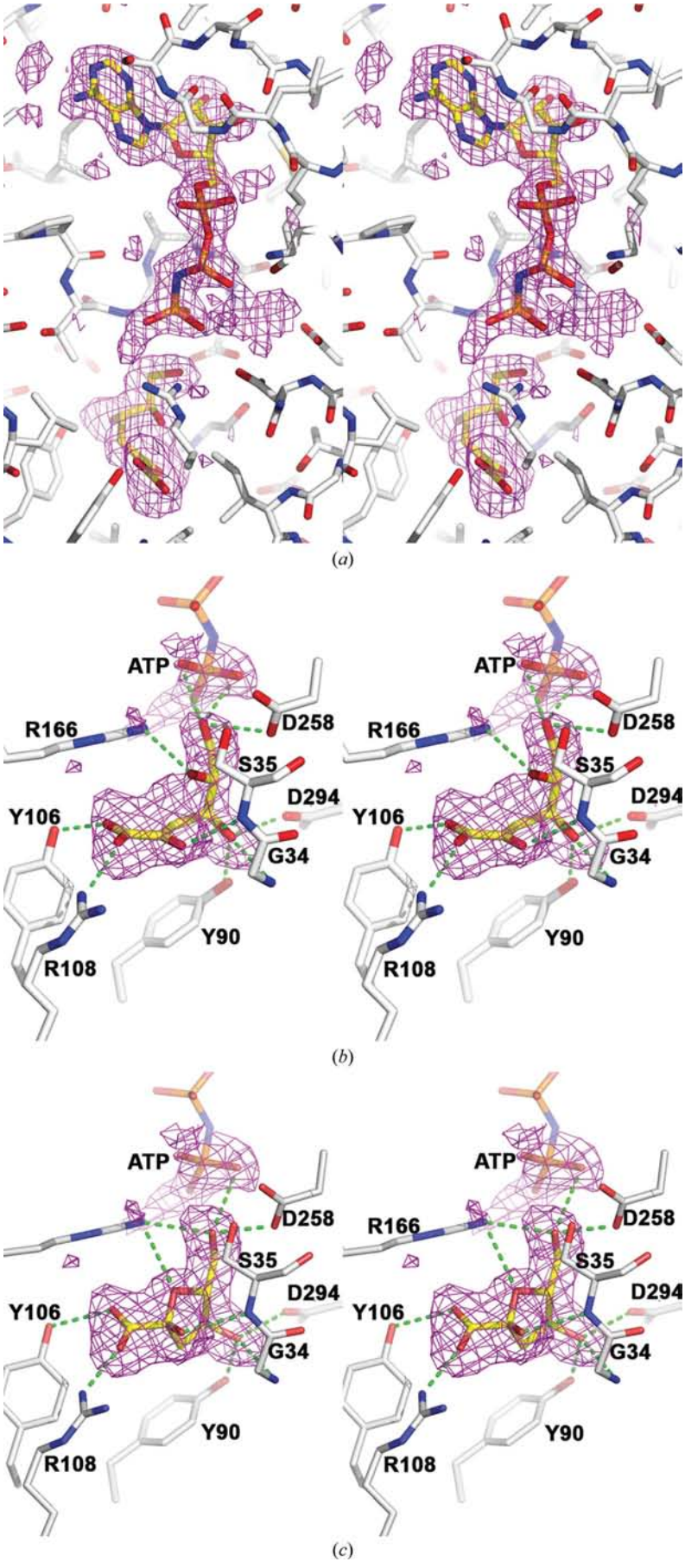



**Figure 1**
Structure of *Ss*KDGK. (*a*) Hexamer and (*b*) monomer of *Ss*KDGK, showing the location of AMP-PNP and KDGlu [green C atoms in (*b*)]. All figures were prepared with *PyMOL* (DeLano, 2007).




the furanose form) and Arg108 $N^{\eta 2}$ (2.7 and 2.6 Å, respectively); O2 of KDGlu interacts with the backbone amide of Ser35 (3.2 and 2.9 Å, respectively); O4 interacts with the amide of Gly34 (3.1 and 3.1 Å, respectively), Tyr90 $O^{\eta}$ (2.8 and 3.0 Å, respectively) and Asp294 $O^{\gamma 2}$ (3.1 and 2.8 Å, respectively); O5 interacts with Arg166 $N^{\eta 2}$ (3.0 and 3.1 Å, respectively), which also interacts with O6 (2.5 and 2.9 Å, respectively); Asp258 $O^{\gamma 2}$ (3.1 and 2.6 Å, respectively) also interacts with O6. Leu11, Leu104, Tyr90, Thr254 and Ile137 also participate in van der Waals interactions with the substrate. For the open-chain form of KDGlu the P atom of the $\gamma$-phosphate of AMP-PNP is 2.7 Å from O6, and in the furanose form it is 3.6 Å from O6, in both cases poised for phosphoryl transfer.

### 3.3. Potential structural basis for promiscuity

*Ss*KDGK is able to phosphorylate KDGlu and KDGal with equal efficiency and we have shown the structural details of how KDGlu is recognized by the enzyme. Attempts to obtain a structure of the *Ss*KDGK–KDGal or *Ss*KDGK–KDPGal complex using similar soaking conditions to those described for the AMP-PNP–KDGlu complex proved unsuccessful; however, given the organization of the active site and the possible favourable orientations the substrate could assume, we can model the binding of KDGal. The different orientation of the C4 hydroxyl in KDGal does not appear to affect the general binding pattern in the active site, as the positions of the other carbon-substituent groups remain unchanged and are predicted to form the same hydrogen bonds as KDGlu.

**Figure 2**
Stereo diagrams of the active site of *Ss*KDGK showing the $F_o - F_c$ difference electron density for the ternary complex contoured at 2.5$\sigma$ (magenta). (*a*) The final refined AMP-PNP and KDGlu (open chain) coordinates superimposed. (*b*) The hydrogen-bonding interactions made by the open-chain form of KDGlu are shown as green dotted lines. (*c*) Hydrogen-bonding interactions made by the furanose form of KDGlu.

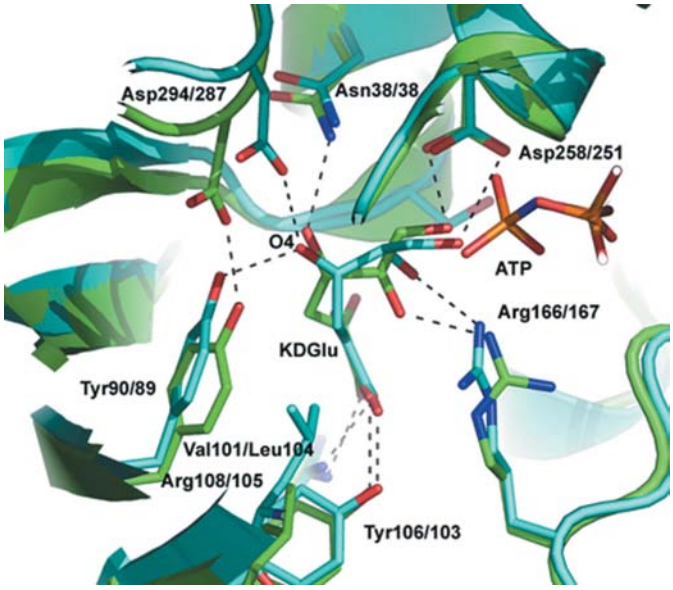



**Figure 3**
A comparison of the binding of KDGlu to *Ss*KDGK (cyan) and *Tt*KDGK (green). Hydrogen-bonding interactions are shown as dashed lines.

However, O4 in KDGal is predicted to form only two hydrogen bonds to Tyr90 and Asp294, losing the third interaction observed between KDGlu and the backbone amide of Gly34. This additional bond may account for the differences in the kinetic parameters for the two substrates and explain the slightly lower $K_M$ value for KDGlu: 3.6 m*M* compared with 8.1 m*M* for KDGal (Lamble *et al.*, 2005).

There is a high conservation of the amino acids in the active sites of *Ss*KDGK and *Tt*KDGK, so why is *Ss*KDGK promiscuous for KDGlu and KDGal? Fig. 3 shows a superposition of the active sites of *Ss*KDGK and *Tt*KDGK with KDGlu bound. It is clear that although there is a slight variation in the conformation of KDGlu, the carboxylate O atoms of KDGlu, as well as O2, O5 and O6, make the same interactions in both structures. The main variation is around O4, where there is a shift in the position of the loop carrying Asp294 in *Ss*KDGK (equivalent to Asp287 in *Tt*KDGK). In *Ss*KDGK the position of this aspartate would allow O4 to interact in either C4 epimer, whereas the position in *Tt*KDGK would disfavour binding of KDGal. Tyr90 in *Ss*KDGK can interact with O4 in KDGlu or KDGal, whereas its equivalent Tyr89 in *Tt*KDGK does not interact with KDGlu but interacts with Asp287. Furthermore, in *Tt*KDGK the O4 of KDGlu interacts with Asn38, an interaction that we would predict to be lost if KDGal were able to bind. Finally, Leu104 in *Ss*KDGK is replaced by the shorter Val101 in TtKDGK and this may have an effect on the positioning of the ligand, in particular C3, perhaps favouring the ability of *Ss*KDGK to bind both diastereomers. We can therefore conclude that promiscuity in *Ss*KDGK may be achieved through subtle positional differences of the active-site residues that are responsible for binding and stabilization of the substrate or the transition state of the reaction. These differences can be attributed to differences in sequence in other parts of the polypeptide chain distant from the active site, which can cause general rearrangement of the various secondary-structure elements and subsequent movement in the chains comprising the active site. Mutagenesis experiments are in progress to test this hypothesis.

We would like to thank the staff at the European Synchrotron Radiation Facility and the European Union for funds to access the facility. We would also like to thank the St Andrews BMS mass-spectrometry and proteomics facility for mass-spectrometric verification of the identity of the purified/crystallized protein. JAP and HJL were supported by a Biotechnology and Biological Sciences Research Council (BBSRC) grant to MJD, DWH, SDB and GLT.

## References

Ahmed, H., Ettema, T. J., Tjaden, B., Geerling, A. C., van der Oost, J. & Siebers, B. (2005). *Biochem. J.* **390**, 529–540.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Danson, M. J., Lamble, H. J. & Hough, D. W. (2007). *Archaea: Molecular and Cellular Biology*, edited by R. Cavicchioli, pp. 260–287. Washington: ASM Press.
DeLano, W. L. (2007). *The PyMOL Molecular Graphics System.* De Lano Scientific LLC, Palo Alto, California, USA. http://www.pymol.org.
DeRosa, M., Gambaccorta, A., Nicolaus, B., Giardina, P., Poerio, E. & Buonocore, V. (1984). *Biochem. J.* **224**, 407–414.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Grogan, D. W. (1989). *J. Bacteriol.* **171**, 6710–6719.
Kim, S. & Lee, S. B. (2006). *Biosci. Biotechnol. Biochem.* **70**, 1308–1316.
Lamble, H. J., Heyer, N. I., Bull, S. D., Hough, D. W. & Danson, M. J. (2003). *J. Biol. Chem.* **278**, 34066–34072.
Lamble, H. J., Milburn, C. C., Taylor, G. L., Hough, D. W. & Danson, M. J. (2004). *FEBS Lett.* **576**, 133–136.
Lamble, H. J., Theodossis, A., Milburn, C. C., Taylor, G. L., Bull, S. D., Hough, D. W. & Danson, M. J. (2005). *FEBS Lett.* **579**, 6865–6869.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Milburn, C. C., Lamble, H. J., Theodossis, A., Bull, S. D., Hough, D. W., Danson, M. J. & Taylor, G. L. (2006). *J. Biol. Chem.* **281**, 14796–14804.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Ohshima, N., Inagaki, E., Yasuike, K., Takio, K. & Tahirov, T. H. (2004). *J. Mol. Biol.* **340**, 477–489.
She, Q. *et al.* (2001). *Proc. Natl Acad. Sci. USA*, **98**, 7835–7840.
Theodossis, A., Walden, H., Westwick, E. J., Connaris, H., Lamble, H. J., Hough, D. W., Danson, M. J. & Taylor, G. L. (2004). *J. Biol. Chem.* **279**, 43886–43892.